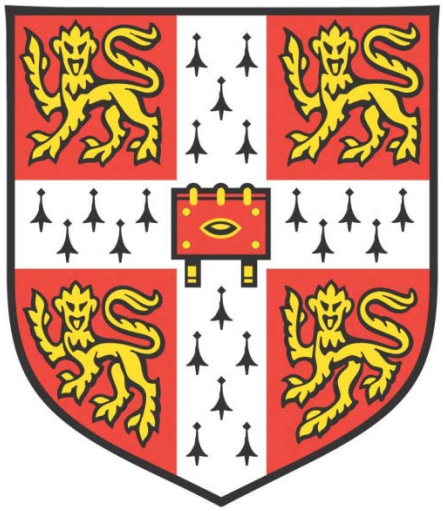
Leverhulme Centre for Human Evolutionary Studies

Department of Archaeology and Anthropology

University of Cambridge



DIET ASSESSMENT IN TROPICAL AFRICAN POPULATIONS

THE IMPLICATIONS OF DETECTING BIOLOGICAL SIGNALS IN CURRENT DIETS TO THE STUDY OF PAST DIETS

Maria Ana Moreira Afonso Correia

Robinson College

February 2018

This dissertation is submitted for the degree of Doctor of Philosophy

DECLARATION OF ORIGINALITY

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Introduction, and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

STATEMENT OF LENGTH

The length of this dissertation does not exceed the 80 000 limit set by the Degree Committee for the Department of Archaeology and Anthropology.

ABSTRACT

East Africa is central to many aspects of human evolution and diversification. At the same time, diet is a key aspect of the ecology of any population. Therefore, one is often interested in the diets of past populations. To assess human diet in the past, stable isotope ratio and dental microwear analyses are often perceived as the only semi-quantitative and objective techniques. However, there are still many unknowns on how isotopic and microwear signals change in response to dietary variation, because few controlled studies have been carried out in modern populations. To investigate this issue, this study targeted living humans from African ethnic groups (El Molo, Turkana, Luhya, and Luo, from Kenya, and Baka, from Cameroon) that practise a wide range of traditional subsistence strategies (pastoralism, fishing, and agriculture), with the objective of building a framework in which to consider past diet in an East African context. This study analysed human hair (n = 143), nail (n = 83), and breath (n = 186) for δ^{13} C and δ^{15} N from the six different communities, and dental moulds (n = 150) from five of those communities (no moulds were collected from the Baka), and related the findings to dietary information. Dental microwear analyses had a low success rate because microwear features were obscured by the biofilm produced by mouth bacteria. Nevertheless, a visual analysis of the results suggested that the El Molo have the hardest and the toughest diet among all the groups studied, possibly through the inclusion of abrasives in the diet during food processing. In turn, the isotopic analyses revealed the ways in which agriculturalists and hunter gatherers differ from pastoralists and fishers in their isotopic values, although the variation in δ^{13}C and δ^{15}N did not distinguish between pastoralists and fishers. The results emphasise recent changes in the diet of these groups, the importance of local factors in isotope values, and the variable sensitivity of isotopes to dietary practices. In conclusion, although each technique could provide complementary data that would contribute to a more inclusive view of diet, dental microwear analyses are not easily applied to modern human groups, due to the difficulty in acquiring comparative *in vivo* data, and in distinguishing between patterns caused by food items, or food processing techniques.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the many people and organizations that made this thesis possible. This study was carried out as part of the IN-AFRICA Project, funded through a European Research Council Advanced Award to Marta Mirazón Lahr (ERC 295907), and with permission from the Office of the President of Kenya (MML//NCST/5/002/R/419), the Turkana County Government, in collaboration with the National Museums of Kenya and the Turkana Basin Institute.

I am grateful to my supervisor, Prof. Rob Foley, for his careful guidance and for teaching me to see the forest through the trees; to Prof. Marta Mirazón Lahr, leader of the IN-AFRICA project, for her continued support and endless enthusiasm; and to both for inviting me to join the IN-AFRICA project and for allowing me the incredible experience of digging in Turkana. I am forever indebted to Dr. Tamsin O'Connell, one of the most accomplished human beings I have ever met, for the constant scientific and emotional help – I believe I am a better person for knowing you. Furthermore, I am thankful to Dr. Fernando Ramírez-Rozzi, for his generous collaboration and for collecting hair and breath among the Baka in Cameroon.

I also want to thank the people of El Molo, Turkana, Webuye, and Port Victoria for permission to work in their area, and in particular, all of the participants that agreed to take part in this project, and put up with my strange questions and puzzling requests with grace and patience. Moreover, none of the data collection in Kenya would have been possible without the field assistance of Julius Akolong, Jeremiah Lesirau, Joseph Wamocho, Simon, Damaris Wanyama, Moses, James, and Rosemary Ochieno.

In Cambridge, I would like to thank the other members of the IN-AFRICA team: Dr. Aurélien Mounier, Dr. Frances Rivera, Dr. Alex Wilshaw, Dr. Ann Van Balen, Dr. Chicca Crivellaro, Herman Muwonge, Pete Griffith, but especially Joe Jeffery, for sharing the experience of fieldwork with me, as well as some of his data. In addition, I want to acknowledge other current and past members of the Department: all the Isotope Gang, for many fruitful discussions on data analysis and interpretation;

Dr. Denis Misiko for helping me organise my fieldwork using his extensive Kenyan network; Marjolein Bosch, for distracting me when I most needed it; and to Nicole Gunstra and Dr. Julie Smith, for a lunch in New Orleans that gave me the needed encouragement to finish writing.

My thanks extend to: Prof. Peter Ungar, for his advice on dental microwear texture analysis; Sarah Livengood, for generously sharing her method on collecting dental moulds from living participants; James Rolfe, for his help with isotopic analysis; Catherine Kneale, for helping me in sample preparation and analysis; Maggie Belatti, Emma Devereux, and Fabio Lahr, for their technical support; and Sylvia Hogg, for all the biscuits.

For extra funding, I would like to thank the Department of Archaeology and Anthropology, Robinson College, and the Cambridge Philosophical Society. But, most of all, I want to thank the European Union, for funding this PhD through the European Research Council, but also the Erasmus programme that set me on the path that led me here. May you live long and prosper.

Finally, I want to thank all my friends in Cambridge and beyond for believing in me. *Para a minha Mãe*, my most devoted supporter – being your daughter is the one thing I am most proud of. *Para Josie, o melhor presente que Cambridge me deu* – thank you for sharing this journey with me.

TABLE OF CONTENTS

LIST OF APPENDICES (ON SD CARD ATTACHED TO BACK COVER)

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This study looks at the diet of living tropical African populations using carbon and nitrogen isotope ratio analysis and dental microwear texture analysis. By investigating this, one aims to better understand the signatures of present-day diets that are preserved by the body, and thus, to better contextualise dietary assessment in the East African past. A better understanding of past diets in this area should be useful to the reconstruction of human evolution and diversification. This chapter describes the design of this research: it includes the background motivation of the research, it delimits the main aim of the study, it explains the research design followed, it enumerates the research questions, and the consequent objective and contributions, and, finally, it outlines the structure of the thesis.

1.1. BACKGROUND

When studying past human populations, one is often interested in the diet they consumed as a component of the ecological niche those people occupied. In other words, the diet of a population can provide information on the role it plays within its environment. However, past ecological niches are often elusive because these are difficult entities to measure (Newsome, Martinez del Rio, Bearhop, & Phillips, 2007), and because the amount of information is restricted by the sliding scale of resolution inherent to archaeology, i.e. the older the site, the worse its preservation (Wynn & Coolidge, 2010).

One may use many perspectives to approach these two issues, i.e. the difficulty in measuring the ecological niche, and the sliding scale of resolution. On one hand, stable isotope ratio analysis and dental microwear texture analysis offer ways to measure this ecological space. In particular, both methods quantify the diet an organism consumes throughout its life, but whereas the former measures the chemical signals of food uptake into the body, the latter looks at the dental wear caused by food and consumed abrasives (DeNiro & Epstein, 1978, 1981; Scott et al., 2006). On the other hand, reliable diet assessment that counteracts the sliding scale of resolution depends on the quality of the fossil evidence, on the methods used, and on the adequacy of developed models (Ungar & Sponheimer, 2013).

Regarding these new aspects of diet assessment in the past, i.e. the quality of the fossil evidence, the methods used, and the adequacy of developed methods, one must consider three points.

First, there is a surprising lack of fossil and chronologically-controlled archaeological evidence of the events that took place in Africa between the appearance of *Homo sapiens*, 200 000 years ago (200 Ka), until the present day (Mirazón Lahr & Foley, 2016). Nevertheless, this region – and particularly East Africa – is central to all aspects of human modern evolution and diversification. Indeed, it is known that human populations expanded and contracted out of refugia in the area, as environmental

conditions fluctuated (Mirazón Lahr, 2016). During these demographic movements, modern humans evolved, as well as the social behaviours associated with modern hunter gatherers (Barham & Mitchell, 2008). Subsequently, throughout the Holocene, humans widened the breadth of resources they consumed, leading to the emergence of new subsistence strategies, such as fishing, pastoralism, and agriculture (Kusimba, 2013; Lane, 2013). These new strategies overtook the dominant hunting and gathering, although this system survives in some areas to this day. In fact, subsistence patterns fluctuated through time, as food-producing populations struggled to adapt to the diverse environment of East Africa (Mirazón Lahr, 2016; Robbins, 2006). Thus, it is clear that the period since the appearance of *Homo sapiens* is a key time in human history, and that diet played a crucial role in the human stage. The IN−AFRICA Project is one of the few projects addressing this wide gap in the archaeological record, by investigating the role of East Africa in the origin and evolution of human diversity. The European Science Foundation (ERC 295907) funds this project, a five year programme (2012-2017) that builds on previous research by the project Later Quaternary Archaeology and Human Palaeontology of the Kerio River and its Environs (2009−2011). The IN−AFRICA Project is based at the Leverhulme Centre for Human Evolutionary Studies (LCHES), University of Cambridge, and is directed by Prof. Marta Mirazón Lahr, with the contribution of Prof. Robert Foley.

Second, by combining the stable isotope ratio and dental microwear texture analyses, one may partly overcome a shortcoming common to both methods: their myopia (Newsome et al., 2007, p. 434). Both methods can only distinguish between resources with distinctive isotopic or microwear signatures: foods with a similar isotopic composition cannot be differentiated, and neither can foods that leave similar or no marks on teeth (Newsome et al., 2007; Sanson, Kerr, & Gross, 2007; Ungar & Sponheimer, 2013). Therefore, by combining methods, one hopes to achieve a higher resolution in diet assessment, since each method may potentially detect foods that the other cannot (Lee-Thorp & Sponheimer, 2006).

Third, both isotopic and microwear studies on archaeological material are based on comparisons with other archaeological faunal and human data, and on theories drawn from laboratory-based animal studies (O'Connell & Hedges, 1999a; Ungar, Scott, Scott, & Teaford, 2008). Therefore, by studying the effect of modern diverse diets on human isotopic values and microwear features, one may (1) understand the impact of diet variation on these variables, and (2) test present palaeodietary theories (O'Connell & Hedges, 1999a; O'Connell, Hedges, Healey, & Simpson, 2001). Once again, this research is particularly necessary in a tropical African context, considering the lack of modern human isotopic data for the continent (Hülsemann et al., 2015). Notwithstanding, one must acknowledge the extensive archaeological work done in recent populations in this continent (Ambrose, 1986; Sealy, 2006, 2010), as well as some previous isotopic work in living East African groups (Kiura, 2005). This body of work has made significant contributions to the research gap indentified here, but much remains to be done.

Therefore, just as the IN−AFRICA Project addresses the lack of archaeological data in Kenya for the described time period, this thesis within the Project addresses the lack of modern dietary data by combining chemical – isotopic – and morphological – dental microwear – methods to study how these signals vary within populations with known subsistence patterns. In other words, this research design addresses most of the issues inherent to diet assessment in the past, namely the difficulty in measuring ecological niches, and the sliding scale of resolution, which itself depends on the quality of the fossil evidence, on the methods used, and on the adequacy of developed models.

1.2. THE AIM

The reasoning behind this project is that by understanding how the isotopic and dental microwear dietary signatures vary among living people whose diets are known to differ, one may create a framework within which to consider older or less

contextualized samples, such as archaeological samples from East Africa, in general, and from the IN-AFRICA Project, in particular.

1.3. RESEARCH DESIGN

This section outlines the research study plan, by laying out the structure used to achieve the aim of this thesis. It includes an overarching methodology, within which this study fits, and a general framework that explains how each step contributes to answering the research questions posed.

1.3.1. OVERARCHING METHODOLOGY

This research is both deductive and inductive. It is deductive in the sense that it builds on pre-existing theory to develop hypotheses that are then tested in a semi-experimental setting. The literature (as reviewed in Chapter 2) shows that human and animal isotope and microwear data vary according to diet in a predictable way. However, one is still largely ignorant on how these data differ across many modern human populations and diets. Thus, I posit that the same principles apply and aim to test this in a semi-experimental setting. Concurrently, this research is inductive because it looks to improve our understanding of diet in the past by studying modern human populations with known diets. In which case, the observations included here are the starting point to a more general theory. Multiple quantitative methods are used to collect and analyse the data, including diet questionnaires, field sampling, dental microwear texture analysis, and carbon and nitrogen stable isotope ratio analysis.

1.3.2. FRAMEWORK

The following framework guides the reasoning behind the steps taken within this thesis. First, the choice of sampled populations focused on three aspects: that they were of tropical African origin, that they maintained relatively traditional subsistence patterns and that, together, they represented a high diversity of these patterns.

Therefore, this study looks at the diets of: the El Molo, fishers from Lake Turkana; the Turkana, pastoralists from the same area; the Luhya, agriculturalists in Webuye, and fishers in Port Victoria; the Luo, fishers in Port Victoria; and finally, the Baka, forest hunter gatherers in Moangé-Le-Bosquet. These populations are all Kenyan, except for the Baka, who are from Cameroon. I did all fieldwork and data collection, except amongst the Baka, where Dr. Fernando Ramirez-Rozzi generously collaborated by collecting the necessary information and samples.

The data collection was threefold: gathering of dietary information, collection of body samples for isotopic analysis, and the taking of dental moulds. The first step was to characterise these populations' diets: applied diet questionnaires described the populations' current diet vis-à-vis the ethnographically described one. After this initial step, one had to consider the feasibility of applying the chosen techniques to the chosen populations. On the one hand, conducting diet questionnaires and dental microwear analyses on live participants is viable, although the latter can be challenging as will be discussed later. On the other hand, it is unfeasible, not to say unethical, to collect bone collagen and bioapatite samples from a live individual, the most commonly used tissues in archaeological studies. Instead, hair and nail were collected to characterise the protein part of the diet, and thus act as proxy for collagen, while breath was collected to characterise the overall diet, and thus act as proxy for bioapatite (Ambrose & Norr, 1993; DeNiro & Epstein, 1978, 1981; Hedges, Rush, & Aalbersberg, 2009; O'Connell et al., 2001; Tieszen & Fagre, 1993). Yet, by using these proxies, it becomes necessary to carefully consider the isotopic relationships between body pools, since the synthetic and metabolic pathways of the body produce distinct isotopic signals. In other words, further steps may be needed to apply the isotopic results of hair, nail, and breath to an archaeological context. In addition, some samples of consumed fish and plants were collected. Even though this is not the main goal of this project, it is important to understand the food web that sustains an organism of interest. However, given legal restrictions on the import of non-aquatic animal products from outside the European Union into the United Kingdom, no samples of this kind were taken.

Finally, Dr. Fernando Ramírez-Rozzi was unable to collect diet questionnaires or dental moulds from the Baka, and so this group was compared to previously published work for the analyses of the relevant data (Romero, Ramírez-Rozzi, De Juan, & Pérez-Pérez, 2013; Sato, Kawamura, Hayashi, Inai, & Yamauchi, 2012; Yamauchi, Sato, & Kawamura, 2000).

1.4. RESEARCH QUESTIONS

Within this framework, the important research questions tackled by this thesis are:

- Do microwear analyses distinguish between modern populations consuming different diets?
- Do carbon and nitrogen isotope analyses of body tissues and breath samples distinguish between modern populations consuming different diets?
- Using the combined techniques, can one answer the following:
    - Do pastoralists consume significantly more animal products than other populations?
    - Do hunter gatherers rely mostly on C_3 forest produce?
    - Can one distinguish fishers from other populations?
    - Are fishers from different lakes distinguishable?
    - Is ethnic affiliation more important than diet or location?
- Are isotope and microwear values consistent with the information from diet questionnaires?
- Are combined isotopic and microwear data from modern human populations useful for diet assessment in the past?

1.5. OBJECTIVES AND CONTRIBUTIONS

Considering all of the above, the main objectives of this research are:

- To understand how well carbon and nitrogen isotopic analyses or microwear analyses distinguish between well-known modern diets;
- To find out whether combining the two types of data improves the resolution of the diet assessment;
- To detect any patterns that may be particular to a tropical African context, and thus relevant to diet assessment in archaeological populations;
- To evaluate how well these methods complement each other.

Consequently, the contributions towards achieving these objectives are:

- A contribution to research into modern human diets, by providing new data from a part of the world (Africa) where data are sorely lacking;
- A contribution to research into past human diets, by building a useful framework based on tropical African diets associated with more traditional subsistence strategies;
- A contribution to research methodology, by understanding the extent to which the three methods used co-vary and complement each other, and by identifying specificities of the methods that may be relevant specifically to the tropical African context.

1.6. GENERAL STRUCTURE

This thesis is divided into eight chapters. The present chapter lays out the reasoning used throughout this work, and enumerates the research questions.

Chapter 2 addresses the role of diet and of the East African enviroment within the ecological niche framework in which human evolution and diversification takes place.

This chapter also reviews the scientific background of dental microwear and stable isotope ratio analysis.

Chapter 3 encompasses the ethnographic description of the studied groups, with an emphasis on diet.

Chapter 4 describes the material and methods used, with the corresponding supplementary material in Appendix A.

Three chapters contain the description of the data gathered and all data analyses and interpretation. Chapter 5 discusses the results from the semi-quantitative food frequency questionnaires, and compares these results to the ethnographic diet described in Chapter 3. Appendix B includes the individual responses to the questionnaires, as well as the details of data analysis. Chapter 6 focuses on the results of dental microwear texture analyses, as well as on the issues faced throughout the use of this technique, particularly the difficulty in removing bacterial biofilm from tooth surfaces. Furthermore, it discusses the results in view of the diet questionnaires and previous research. Appendix C contains all supplementary information for this chapter. The last data chapter, Chapter 7, considers the results from the stable isotope ratio analyses, with additional information in Appendix D. As for the previous the chapter, the results from this technique are discussed vis-à-vis the results from the diet questionnaires and previous research. In addition, this chapter debates aspects specific to the study design used, such as the differences between the isotopic signals of analysed tissues.

Finally, Chapter 8 outlines the main conclusions from this work, in view of the research questions enumerated here, and examines the implications of these findings to the understanding of past diets.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews the main themes of this thesis. It starts with a discussion of diet and of East Africa as two dimensions that play a role in human evolution and diversification, aiming to explore the key relationship between subsistence and environment, but also to contextualize this thesis within the IN-AFRICA Project it belongs to. Next, it reviews approaches used for diet assessment in the past, building an argument for the methods used here. And last, it discusses the physical and biological principles of stable isotope ratio analysis and dental microwear texture analysis. It is necessary to understand these principles in order to interpret the methods successfully.

2.1. THE ROLE OF DIET

A need to ingest food defines all animals as an entity while diet flexibility is a hallmark of primates. In fact, diet is recognized as the key aspect underlying behavioural and ecological differences in living primates (Fleagle, 2013). Considering this, it comes as no surprise that studying the diet of a species or of a population helps in understanding its ecology. In fact, what, where, and when an individual eats influences its reproductive success and the relationships it establishes with its environment and with other organisms, leading to diversification and ultimately to evolution (Lambert & Rothman, 2015). In other words, diet acts as a driver of human diversification and evolution. In human populations, diet assessment allows for a more comprehensive grasp of social structures, mobility patterns, and overall health. This becomes particularly relevant in the study of past populations with no written record or with scarce contextual information. For these populations, diet becomes central to understanding their idiosyncrasies, ecological resilience, decision-making processes and consequently their place in History (Alemseged & Bobe, 2009; Kelly, 2013; Ungar & Sponheimer, 2011, 2013; Wood & Schroer, 2012).

While diet acts as a driver of human diversification and evolution, diet itself depends on the availability of food resources and on an individual's ability to exploit them (Lambert & Rothman, 2015). Consequently, diet must be understood within its environmental and geographical context. Hence, the next section explores the role of East Africa in human diversification and evolution as mediated by diet.

2.2. THE ROLE OF EAST AFRICA

Africa is central to human evolution and diversification. In particular, East Africa, and the East African Rift System, loom over all aspects of this process from the late Miocene to the present day (Darwin, 1871/2009; Foley, 2013; Kusimba, 2013; Semaw, 2000; White et al., 2009; Woodward, 1921). The arguments that explain this role lie in the complex

ecological and topographic structure of East Africa, which permits the formation of niche subdivision and hence, of a large diversity of environments, and consequently of diets (Kusimba, 2013; Mirazón Lahr, 2013a; Mirazón Lahr & Foley, 2016).

2.2.1 AFRICAN POPULATIONS EVOLUTION AND DIVERSIFICATION

This section reviews the evolutionary trajectory of recent hominins and humans, as well as the diversification in Africa since the origin of *Homo sapiens*, with a focus on dietary adaptations. This exercise illustrates the range of diet diversity in modern humans, and how diet mediates the complex interactions between humans and the environment.

Homo sapiens emerges at the latest between 350 Ka and 260 Ka (Schlebusch et al., 2017) during the late Middle Pleistocene – 780 Ka to 125 Ka. Despite its scarce fossil record, this is an intriguing period in human evolutionary history: it is argued that this is the period in which social behaviours typical of modern hunter gatherers appear, such as food sharing, sexual division of labour, and use of base camps. (Barham & Mitchell, 2008).

At the time of the emergence of modern humans, genetic data show that the ancestors of the modern Khoisan from southern Africa diverged from the remaining modern humans. After this, but before 200 Ka, another split event gave rise to the lineage that today includes Central African foragers, such as the western (Baka and Biaka) and eastern (Mbuti and Efe) Pygmies from Central Africa (Schlebusch et al., 2017). Later, one of the existing populations expanded once again, becoming the ancestors of the Bantu, and eventually, of all non-African peoples of the world. These signals are confounded by significant admixture in later periods of human history, and by frequent population extinction events, sometimes associated with food-producing people overtaking existing hunter gatherers (Mirazón Lahr, 2013a; Schlebusch et al., 2017; Tishkoff et al., 2009). Nevertheless, they prove an early differentiation of African populations (Mirazón Lahr, 2016).

In the Upper Pleistocene – 125 Ka to 10 Ka –, the first unequivocal signs of symbolic behaviour appear and stone tool technology develops to encompass new strategies of preparing tool blanks and shaping bifaces, and later on, microlith and composite tool technology (Barham & Mitchell, 2008; Mirazón Lahr, 2012). This technological transition may represent consumption of a wider variety of resources. This increase in dietary breadth supported larger and denser human populations, and was key to the expansion of modern humans out of Africa (Steele & Klein, 2009). However, the Last Glacial Maximum, peaking at around 18 Ka, was a particularly arid period in the tropics that led to the expansion of dry habitats, although highlands, lowland tropical forests close to the eastern edge of the Congo Basin and lakes and rivers probably still offered some refuge (Bonnefille, Roeland, & Guiot, 1990; Kusimba, 2013). Important adaptations include a high mobility and the ability to exploit lower ranked resources, such as fish (as reviewed in Kusimba, 2013).

Moving into the Holocene – circa 10 Ka to the present day –, warm and moist conditions punctuated by episodes of aridity characterised the East African environment (Kiage & Liu, 2006; Timm, Köhler, Timmermann, & Menviel, 2010). Present-day environmental features in this region became well established after 6 Ka, including the bimodal rainfall pattern and drought-resistant vegetation zones (Kiage & Liu, 2006; Mirazón Lahr, 2016; Thompson et al., 2002). This allowed hunter gatherers to expand from localized refugia, establishing an extensive population or a trading network across most of Africa (Mirazón Lahr, 2016). Aquatic resource exploitation intensified during the early Holocene (known as the African Humid Period), and pottery started to appear (Kusimba, 2013; Robbins, 2006). Due to the unpredictability of food resources, hunting and gathering required small groups spread across large territories, although fisher-forager populations probably lived semi-sedentary lives. Plant and animal husbandry offered a solution to this problem and led to higher population densities (Mirazón Lahr, 2016). East Africa is one of the few places in the world where the adoption of domestic livestock likely occurred before plant cultivation. Animal husbandry started around 4.5 Ka, more or less uniformly across the region and possibly following a southern route (Lane, 2013; Robbins, 2006). Evidence

suggests that first pastoralists, and later farmers, gradually replaced or assimilated hunter gatherer populations (Mirazón Lahr, 2016; Tishkoff et al., 2009). However, in East Africa, subsistence patterns fluctuated during the Holocene (Robbins, 2006). This is partly explained by challenges to the cultural and biological adjustments needed after the introduction of domesticated species in the region. These challenges, coupled with low population density, might explain the survival of hunter gatherers until recent times in regions not amenable to farming, such as deserts, forests, and lakeshores (Barham & Mitchell, 2008; Kusimba, 2013).

Reviewing the hominin narrative allowed us to clarify how humans have established complex interactions with the environment where ecology, and hence diet, are key to human behaviour. Over time, dietary breadth has increased, together with enhanced technological skills applied to food acquisition and, in a final stage, production. Nonetheless, there are still many gaps in the narrative. Namely, it is still unclear through which mechanisms have humans evolved and adapted. Hence, the next section addresses a persuasive model for these processes, while the following outlines the IN-AFRICA project, a research project that studies the role of East Africa in the evolution of human diversity.

2.2.2 AFRICA AS A REFUGIA NETWORK

After clarifying *how* humans evolved and diversified in Africa, one must now try to understand *why*. Mirazón Lahr (2013b) proposes a *refugia network* system where independent, potentially asynchronous refugia are connected by geographic corridors. This idea builds on previous research that advocates that East Africa and the southern Cape acted as refugia for early humans (Basell, 2008; Compton, 2011; Foley, 2013; Mirazón Lahr & Foley, 1994). Relevant to this work, this wide range in environments also implies dietary diversity. The refugia hypothesis is consistent with genetic evidence, which shows a dynamic demography driven by climate change and human adaptation – both biological and cultural – that led to a pattern of population expansion, followed by fragmentation and partial isolation. Such a model would create

a fluid mosaic of population bottlenecks and expansions, stimulating complex contexts of competition, resulting first in population diversity – both behaviourally and morphologically – and eventually in species' evolution (Mirazón Lahr, 2013a, 2013b).

Despite providing an appealing and elegant model to contextualize human evolution, few studies have investigated where such refugia were located, how did they work, and how do they translate in the archaeological record (Basell, 2008). As mentioned above, the proposed refugia locations consist of tropical forests, highlands, coastal plains, and lakes and river margins. In East Africa, refugia were likely located on the Ethiopian and Kenyan Highlands, and around Lake Turkana, Lake Victoria, Lake Eyasi, and Lake Nakuru. The Highlands were probably persistently covered with montane forest and canopy woodland (Basell, 2008). Three of the lakes (Turkana, Eyasi and Nakuru) were surrounded by savannah, while Lake Victoria would have had savannah on its eastern margin, and open woodland on the southern and western margins. These ecological niches would contract and expand with climate variations and with it, so would human populations. In short, these niches were areas with some tree cover, which provided a refuge from the extended areas of open grassland, semi desert, or desert that separated them (Basell, 2008).

Whilst the likely location of refugia is well understood, little is known about refugia activity and processes, i.e. more research is needed on the use of these spaces by different species, and on the relationships established between refugia. To achieve this, archaeological studies must be carried out at a regional scale and within an interdisciplinary framework, so that they may later be integrated at a global level. In addition, studies on modern examples of refugia are also necessary (Basell, 2008), since the substantial recent research on forest fragmentation caused by human environmental degradation tackles the problem at a different geographical scale.

2.2.3 *THE* IN-AFRICA *PROJECT*

The IN-AFRICA Research Project investigates the role of East Africa on the evolution of human diversity, building on previous research by the Later Quaternary Archaeology and Human Palaeontology of the Kerio River and its Environs Project.

The Project focuses on two areas of the Rift Valley of Kenya: the Nakuru-Naivasha Basin and West Turkana. Work in the first aims to bridge temporal gaps between Acheulean sites and Holocene ones, while work in the latter aims to identify new archaeological sites from the late Middle and Upper Pleistocene and early Holocene. So far, survey and excavation in West Turkana has focused on an area to the southwest of Lake Turkana, on the western margin of the Natome River – a tributary of the Kerio – and along the predicted shores of the late Pleistocene and early Holocene palaeolake Turkana (Figure 2.1) (Mirazón Lahr & Foley, 2014). Here, several archaeological sites have been identified, ranging in date from 12 Ka to 7 Ka (Mirazón Lahr & Foley, 2012), as well as a number of earlier sites that reflect human and hominin exploitation of the lake shore during previous interglacials (Lahr & Foley, 2012). Most of the early Holocene sites have comparatively few archaeological remains (as distinct from fossil remains), but at least two indicate intense use, possible occupation, and human remains. One site – Nataruk, found in 2012 – contained several completely or partially exposed skeletons that showed evidence of a violent death. This is one of the earliest cases of intergroup violence (Mirazón Lahr et al., 2016). Figure 2.1 illustrates the geographical distribution of these sites.

Incidentally, one of the key questions surrounding the Nataruk site concerns whether the conflict resulted from a raid for resources (e.g. territory, women, or food), or from an antagonistic response between social groups. In this context, it would be important to determine if these individuals belonged to a food-storing society or not. In the same way, and considering the role that diet plays in diversification and evolution, most archaeological research would benefit from a better understanding of the feeding habits of the targeted population.

Therefore, the IN-AFRICA Project provides a well grounded and interdisciplinary context on which to develop my research. By studying diet variation in modern populations, I aim to better contextualize populations in the past and hence to contribute to projects such as this one. In other words, by understanding how diets affect the chemical and morphological make-up of modern humans, one may look at the past through a new lens.

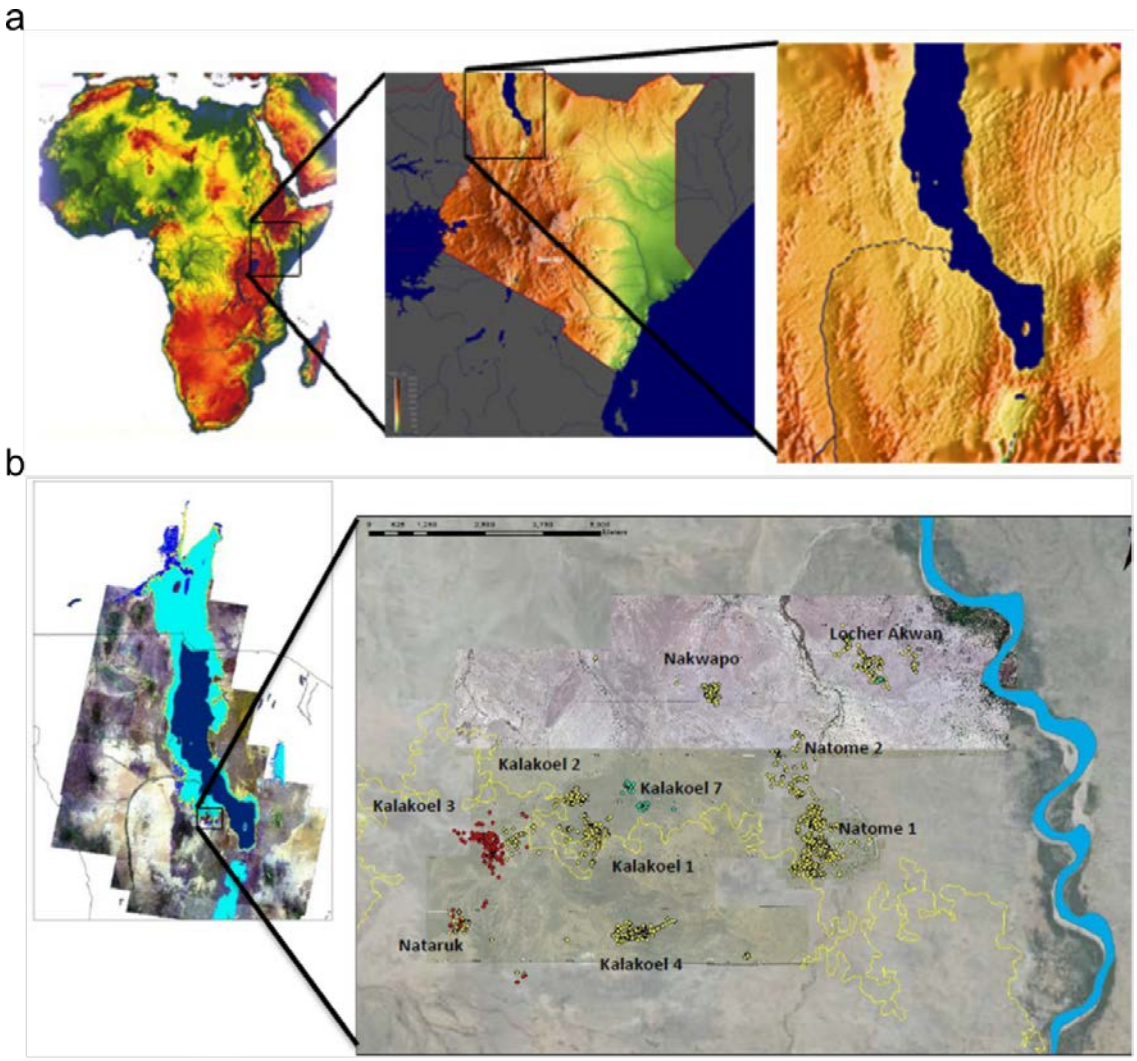


Figure 2.1 IN-AFRICA Project geographical location: (a) Lake Turkana within Africa ; (b) IN-AFRICA Project main paleontological and archaeological sites in West Turkana, with the Kerio River in light blue (Mirazón Lahr & Foley, 2014, p. 32).

2.3 DIET ASSESSMENT IN THE PAST

This section focuses on diet assessment in the past, namely the methods developed so far, and the resultant contributions to the subject. By reviewing this research, I aim to argue for the choice of methods used in this thesis – dental microwear texture analysis, and carbon and nitrogen isotope analysis. The theoretical background of these two methods will be further developed in the following sections.

In general, methods to assess diet in the past are divided into three categories (Ungar & Sponheimer, 2013):

- those used on the fossil evidence itself;
- those based on evidence found in context with the fossils; and,
- those derived from observations of living people and other animals, often involving some form of model construction.

The present section is organised according to this categorization into fossil evidence, contextual evidence, and modelling approaches.

2.3.1 FOSSIL EVIDENCE

Diet assessment methods used on fossil evidence comprise the study of tooth size, tooth shape, and tooth structure, as well as craniomandibular biomechanics, tooth macro- and microwear, tooth and bone chemistry, and dental and skeletal pathology.

Tooth Size

Particularly among hominins, large incisors are seen as an adaptation to foods requiring incisal preparation such as husked fruits, while large molars with some sharp edges are seen as an adaptation to an omnivorous diet (Kay, 1975; Ungar, 2012, 2014). Conversely, the decrease in molar size observed in the genus *Homo.* is traditionally explained by a relaxing of selective pressures on large molar size due to the development of food processing and cooking (Ungar, 2012).

This interpretation, however, is far too simplistic. At this point, researchers still do not fully understand the relationships between diet, tooth function, and tooth size (Kay, 1977; Ungar & Lucas, 2010; Ungar & Sponheimer, 2013). For instance, while frugivorous New World monkeys have larger molars than folivorous ones, the inverse is true for Old World monkeys (Kay, 1977; Ungar, 2012; Ungar & Sponheimer, 2013). As suggested by Ungar (2014), dental allometry can be used as a starting point to generate hypotheses that should subsequently be tested using other lines of evidence.

Tooth Shape

It has been proposed that primates that consume tough leaves or insects with a tough exoskeleton have teeth with long, sharp blades, and primates that consume hard objects have the flattest occlusal crowns with blunt cusps (Ungar & Lucas, 2010). However, Berthaume et al. (2013) suggest that, in fact, a combination of both blunt and sharp cusps might be the best solution to process hard items, since it produces high stress concentrations in the food item while minimizing stresses on the enamel, with some cusps initiating the fracture while others stabilize the food item.

An important limitation to the study of tooth shape is tooth wear. Most studies use unworn teeth, which is problematic because (1) it decreases the sample size and (2) the ways the organism copes with dental wear are also relevant to diet. Dental topographic analysis offers a solution to this problem since it does not depend on specific landmarks that might change with wear (Berthaume, Dumont, Godfrey, & Grosse, 2013; Ungar, 2004, 2007; Ungar & Sponheimer, 2013). Ungar (2004) used this technique to examine teeth from early hominins and extant primates, and found that *Australopithecus afarensis*, chimpanzees, *Homo spp.*, and gorillas, lie, in that order, on a continuum from low to high occlusal relief. Based on these results, the researcher suggests that *Homo spp.* were more adapted to consume tough, elastic foods, such as meat, than chimpanzees or *A. afarensis*, while *A. afarensis* was more suited to eat hard foods. Unfortunately, dental topographic analysis is still not widely used, and so, few comparative samples are available.

Tooth Structure

Although teeth are composed of different materials, the main components are dentin – that forms the core of the tooth – and enamel – that coats the dentine (Teaford, 2007b). Having thick enamel has been considered a diagnostic hominin trait among African apes (Pampush et al., 2013; Smith et al., 2012; Teaford, 2007b; Ungar & Lucas, 2010). There are two main explanatory hypotheses for thick enamel, not mutually exclusive. The first proposes that having thicker enamel is an adaptation to prevent tooth fracture and chipping resulting from feeding on hard foods. The second argues that thick enamel aids in maintaining tooth function in the face of gradual dental wear from grit, phytoliths and acid (Pampush et al., 2013; Ungar & Lucas, 2010).

Craniomandibular Biomechanics

Cranial morphology is largely explained by neutral evolution; that is, it is the result of stochastic processes of mutation and genetic drift (Betti, Balloux, Hanihara, & Manica, 2010; Nicholson & Harvati, 2006; Roseman & Weaver, 2004; Smith, 2009; von Cramon-Taubadel, 2009). However, skull morphology remains integral to multifunctional processes, including protection of the brain and sensory organs, food processing, and maintenance of the respiratory airways (Daegling, 2010). As such, certain regions of the skull also reflect these functions to a certain extent (Noback & Harvati, 2015; von Cramon-Taubadel, 2014a). For instance, the mandible is likely influenced by diet (Harvati & Weaver, 2006; von Cramon-Taubadel, 2011), which may explain why hunter gatherers have longer and narrower mandibles than agriculturalists (von Cramon-Taubadel, 2011). This pattern may be due to the development of food processing associated with agriculture and consequent decrease in masticatory stresses (González-José et al., 2005; Lieberman, 2008; von Cramon-Taubadel, 2011). However, considering that non-agriculturalists populations also process food, other factors, such as longer weaning periods in hunter gatherers, might have also contributed to this pattern (von Cramon-Taubadel, 2011, 2014b).

Dental Macro- and Microwear

Whilst tooth shape offers evidence on what an animal evolved to eat, dental macro and microwear studies provide direct evidence on the diet of an individual during its lifetime (Teaford, 2007a).

Macrowear is gross dental occlusal wear, observable at the macroscopic level (Schmidt, 2010). Such macrowear may eventually lead to changes in mandible shape, which links dental wear with craniomandibular mechanics (Margvelashvili, Zollikofer, Lordkipanidze, Peltomäki, & Ponce de León, 2013). So far, macrowear has been successfully used to distinguish between hunter gatherer populations who consume carnivorous and mixed diets (Fiorenza et al., 2011). However, the same study could not distinguish between marine versus terrestrial animal-based diets, which brings to attention the lack of resolution of this technique.

A plausible alternative to the study of macrowear is microwear analysis, which looks at the microscopic patterns of use wear left on teeth by different food types, or by the abrasives contained in them (Grine, Sponheimer, Ungar, Lee-Thorp, & Teaford, 2012; Teaford, 2007a; Ungar, 2012; Ungar & Lucas, 2010; Ungar & Sponheimer, 2013). Specifically, microwear studies inform on the material properties of the items an individual consumes (Scott et al., 2006; Ungar & Sponheimer, 2013). Clearly, some materials, like very soft foods, will not leave traces on the teeth. Instead, materials that produce microwear fall along a hard versus tough continuum; in which hard items have stress-limited characteristics that prevent a crack from starting, while tough items have displacement limited characteristics that halt a crack from spreading (Ungar & Lucas, 2010). As a result, the teeth of hard food consumers have a high prevalence of pits in their microwear surfaces – formed through the crushing of food by opposing teeth – while teeth of tough food consumers have parallel scratches – formed by abrasives being dragged across tooth surfaces (Grine et al., 2012; Teaford, 2007a; Ungar & Lucas, 2010). At this point, it is still unclear how macro- and microwear correlate, although it seems that wider scratches – associated with a more abrasive diet – lead to more macrowear (Schmidt, 2010).

Using scanning electron microscopy (SEM), microwear studies in modern humans indicate that the shift from hunting and gathering to agriculture also represents a shift to a softer, less diverse diet, although this varies according to population and habitat (Romero et al., 2013). Traditionally, microwear studies have been carried out using SEM. These, however, are costly, time consuming, and have high rates of intra- and inter-observer error, even when using computerized routines (Grine, Ungar, & Teaford, 2002). A solution to these problems lies on an improved technique, dental microwear texture analysis, which combines point clouds produced by confocal microscopy with scale-sensitive fractal analysis (Scott et al., 2005).

Tooth and Bone Chemistry

Tooth and bone chemistry studies are based on the principle that 'we are what we eat'. In other words, the chemical components of ingested food find their way into an organism's tissues in a measurable fashion (DeNiro & Epstein, 1978, 1981; Sponheimer & Dufour, 2009). Just as wear studies, these inform on the diet of an individual during its lifetime. Whilst stable isotope ratio and trace element analyses are the two main branches of this field of study, this section does not discuss the latter as it is not widely used in the current research climate, due to the biological signals being obscured by diagenesis, and to the need for an in-depth knowledge of the food webs that sustain the organism of interest (Sandford, 1993; Sillen, Sealy, & van der Merwe, 1989; Sponheimer & Lee-Thorp, 2006).

The most common elements used in diet assessment are carbon and nitrogen. Carbon isotope studies distinguish between the types of plants (C_3 or C_4) that are at base of the food chain, while nitrogen isotope analyses track the organism position in the food chain (DeNiro & Epstein, 1978, 1981; Kellner & Schoeninger, 2007; Schoeninger & DeNiro, 1984). Within archaeology, stable isotope studies are more often performed on bone collagen, and bioapatite. However, if available, such studies are applicable to many other issues containing carbon, nitrogen, and hydrogen, including hair, nails, skin, and muscle (Schoeninger, 1995; Tipple, Chau, Chesson, Fernandez, & Ehleringer, 2013).

Stable isotope studies are an extremely powerful tool. Significant contributions so far include the C_4 specialization of *Paranthropus sp*. (Cerling et al., 2013), and the high trophic level of Neanderthals and Late Pleistocene modern humans in Europe (Hedges & Reynard, 2007; Lee-Thorp & Sponheimer, 2006). However, these analyses also have significant limitations. They are expensive and destructive. In addition, both bone apatite and collagen are susceptible to diagenesis – changes in composition after formation (Hoppe, Koch, & Furutani, 2003; King, Tayles, & Gordon, 2011). And finally, low sample sizes compromise the reliability of the interpretation, which is problematic in archaeological contexts where preservation is often incomplete (Pearson & Grove, 2013). Approaches to counteract these problems combine less destructive sampling techniques (Copeland et al., 2008), and better modelling of results (Erhardt & Bedrick, 2012; Fernandes, Millard, Brabec, Nadeau, & Grootes, 2014; Phillips, 2012).

Pathology

Certain pathological elements can also yield information on diet. For instance, caries are the result of fermentation of food sugars by bacteria on the teeth, and can indicate consumption of sugars, like starch. An increase in the incidence of carious lesions is usually linked with the development of agriculture (Humphrey et al., 2014). In contrast, dental calculus might indicate high levels of protein consumption, which increases the alkalinity of the oral cavity and leads to plaque mineralization through microbial activity (Lieverse, 1999; Scott & Poulson, 2012).

2.3.2 CONTEXTUAL EVIDENCE

Dietary assessments are best understood in their palaeoecological context. This approach focusses on plant and animal remains found in the archaeological context of an individual under study, and predicts food availability and consumption, for instance. In addition, these methods are often the most successful in sites where preservation is poor (Reed & Rector, 2007; Ungar & Sponheimer, 2013).

The two main types of plant remains that are most frequently retrieved from archaeological contexts are macro and microfossils. Macrofossils include fossils of

seeds, leaves, wood and other plant parts; the study of such remains has limited application because plants do not preserve well and thus the finds recovered may not be representative of the plant community at the time of death of the individual (Copeland, 2009; Humphrey et al., 2014; Peters, 2007). Microfossils are microscopic remnants of plants with distinctive morphologies that (1) preserve well and (2) are often diagnostic of particular plant taxa and, in some cases, of plant organs. Microfossils encompass phytoliths, starch grains, and pollen grains. The first two are commonly used to reconstruct diet because they are present in consumed plant products, such as fruits and leaves. Specifically, phytoliths are siliceous bodies formed in the leaves, husks, stems, fruits, and some underground organs, while starch grains are energy-storage bodies that occur in greatest numbers in seeds and underground organs. Such microfossils have been recovered from a variety of materials, such as dental calculus, tools, sediment samples, and coprolites (Henry, 2012; Henry, Brooks, & Piperno, 2011). An important find that used the study of microfossils is the consumption of grasses (nutritionally rich, but hard to exploit) by Neanderthals (Henry et al., 2011).

In addition to plant remains, the zooarchaeological record contributes in three unique ways to dietary assessment (Blumenschine & Pobiner, 2007). First, butchery and gnawing marks on bone confirm meat consumption in hominins (McPherron et al., 2010). Second, co-occurrence of this evidence informs on carcass acquisition mode and on the carnivore species with which hominins interacted, either directly or indirectly (Pante, Blumenschine, Capaldo, & Scott, 2012). Third, the zooarchaeological record can test theoretical constraints, such as predator encounter risk, or decision-making on fall-back foods (Alemseged & Bobe, 2009). These three sets of evidence can, together, provide insights into the diets of past people, but also on their behaviour.

2.3.3 MODELLING APPROACHES

After examining the evidence from fossils and their context, one can further glean information by establishing parallelisms. This extends from simple ecological analogies

to complex multi factor models (Ungar & Sponheimer, 2013). Modelling is especially relevant because, as data accumulate, theoretical challenges grow with attempts to compare different types of evidence with different sampling scales, taphonomic histories, and interpretive frameworks (Alemseged & Bobe, 2009; Sept, 2007). Modelling hinges on two key principles: uniformitarism, wherein natural laws are spatially and temporally invariant; and parsimony, which establishes that between competing hypotheses, the best is the one with the fewest assumptions. Two areas that make ample use of models are the study of ecological analogues and genomics.

Ecological Analogues

No extant living animal is an avatar of an extinct hominin. Likewise, no animal model perfectly mirrors a human. Nevertheless, if tailored to the circumstances, using ecological analogues can be very useful. For instance, modern hunter gatherers are commonly used as models for the assessment of prehistoric diets. In fact, the study of forager diets has helped researchers understand just how diverse are human diets, and how adjustable to the environmental context (Johnson, 2013; Kelly, 2013).

Comparative Genomics

The study of genetic adaptations to diet is in its early days. So far, well known genetic adaptations to diet include lactase persistence and starch digestion (Luca, Perry, & Di Rienzo, 2010). On lactase persistence, evidence shows that in most mammals the ability to digest lactose contained in milk declines after weaning due to a decrease in the levels of the lactase enzyme; however, in some human populations, this enzyme persists into adulthood, allowing people to continue milk consumption. Notably, lactase persistence in adults is more prevalent in populations with a long history of pastoralism and of milk production, such as European, East African, Middle Eastern, and South Asian populations. Moreover, this feature has evolved by convergence in some of these populations and co-occurs with the onset of dairying (Bersaglieri et al., 2004; Enattah et al., 2008; Ingram, Mulcare, Itan, Thomas, & Swallow, 2009; Luca et al., 2010; Tishkoff, Reed, et al., 2007). Starch digestion is linked to the gene AMY1. This gene varies extensively in the number of copies across individuals and populations,

and this variation correlates with the amount of starch contained in the diet (Perry et al., 2007). Another example is found in the NAT2 gene, which is involved in the regulation of toxic compounds generated by cooking or other food treatments (Patin et al., 2006).

This section reviewed the methods used in diet assessment in the past. It became clear that the two techniques whose signal is not hindered by phylogeny are dental microwear and stable isotope analyses because they directly target the diet of the individual. As such, they are the best candidates to document variability in the past, and are the techniques used in this work. This chapter will now expand on the theoretical background of these two techniques.

2.4 DENTAL MICROWEAR TEXTURE ANALYSIS

Dental microwear texture analyses were previously discussed in this chapter in the dental macro and microwear section. Together with stable isotope analysis, they are one of the few techniques that offer a quantitative approach to diet assessment. As stated before, dental microwear *texture* analyses offer a solution to the high levels of intra- and inter-observer errors found in traditional microwear analysis (Grine et al., 2002). However, in order to correctly apply and interpret it, one must understand the theories behind the technique, namely the physics of surface roughness measurement (Scott et al., 2006).

2.4.1 DENTAL MICROWEAR TEXTURE TECHNIQUE

As mentioned above, microwear constitutes direct evidence of the food an organism ingested during its lifetime (Grine et al., 2012; Scott et al., 2006; Teaford, 2007a; Ungar, 2012; Ungar & Lucas, 2010; Ungar & Sponheimer, 2013). Furthermore, dental microwear texture analysis (DMTA) renders results comparable across studies and provides more complete characterization of dental microwear surfaces than a simple pit versus scratch dichotomy (Scott et al., 2005). Interestingly, an *in vitro* DMTA

experiment recently confirmed that different food items and abrasive loads produce microwear features that can be characterised through this method (Hua, Brandt, Meullenet, Zhou, & Ungar, 2015).

In broad terms, the current technique comprises four steps: (1) dental moulds are taken using dental silicone; (2) replicas of the teeth are made using an epoxy; (3) a Phase II facet on each tooth is scanned using a confocal microscope; and (4) scale sensitive fractal analyses are run using the Sfrax® and Toothfrax® software (Scott et al., 2006). I will go through each step of the technique to explain its principles.

First, dental moulding (1) and casting (2) is done using traditional techniques, which are highly reliable in recording microwear features (Fiorenza, Benazzi, & Kullmer, 2009; Galbany et al., 2006). These, however, were developed with archaeological material in mind. In contrast, live subjects are more inconvenienced by this procedure and have bacterial film covering microwear features that must be removed prior to moulding (Teaford & Oyen, 1989b). This will be discussed in more detail in the methodology section.

Following this, confocal microscopy (3) is used to examine the high-resolution casts because it allows collection of 3D data without the disadvantages of traditional SEM microscopy. The latter is costly, time intensive, and produces different results depending on microscope characteristics, such as collector type and position, types of electrons used, voltage, and working distance, among others (Gordon, 1988; Scott et al., 2006). Furthermore, for DMTA analysis, only Phase II facets on molar teeth are scanned for control over sampling area, as these are the ones affected during the biomechanics of mastication (see below). Mastication has two intervals: Phase I, which cuts through food, and Phase II, which grinds it (Kay & Hiiemae, 1974). Each of these phases has associated tooth movements that produce different wear facets on teeth. While Phase I facets are found along the surfaces leading away from the crests, Phase II facets are found in the basins of the teeth (Kay, 1977; Kay & Hiiemae, 1974; Krueger, Scott, Kay, & Ungar, 2008). Furthermore, Krueger et al., (2008) found differences between the two

types of wear facets, also observing that Phase II facets were better at distinguishing between diets, although the cause behind this is unclear. It is possible that Phase II facets result from food being pressed between flat opposing surfaces, or perhaps these facets are less restricted by dental anatomy. To control for the badly understood variation and for ease of comparison across studies, DMTA usually uses facet 9, 10n, or x on lower molars, which are illustrated in Figure 2.2.

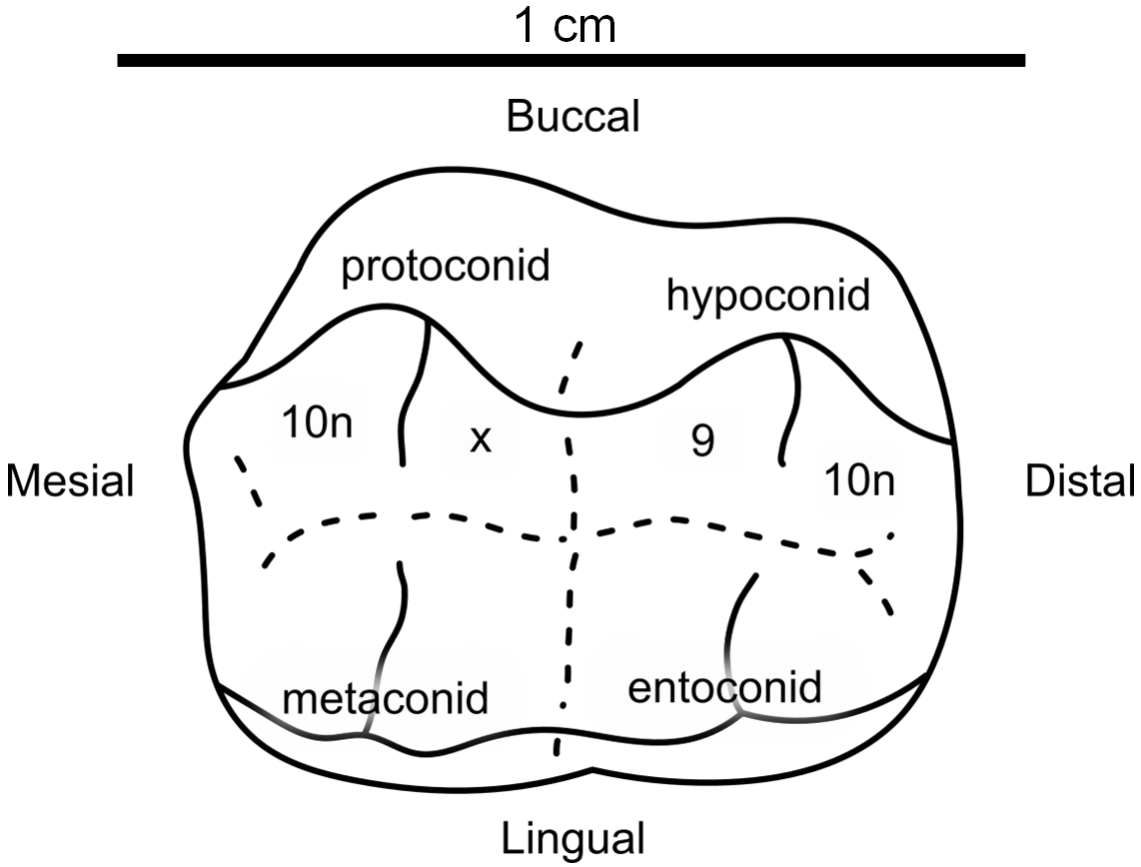


Figure 2.2 Occlusal view of a *Macaca mulatta* right lower second molar, indicating the position of Phase II facets (9, 10n and x). The same facet number is used for matching surfaces on upper and lower teeth that are produced by contact between each other. Facet 9 is situated on the lateral surface of the protocone (upper molars) and on the medial slope of the hypoconid (lower molars). Facet 10n is located on the lateral slope of the hypocone and on the medial facing slope of the hypoconulid. Facet x is located on the anterolateral slope of the protocone and on the posteromedially facing slope of the protoconid. Adapted from Kay (1977) with scale based on Trotter, Hixon, & MacDonald (1977).

Finally, scale sensitive fractal analyses (4) are based on the principle that surface roughness changes with scale of observation. The Sfrax® and Toothfrax® software produce several parameters that characterise a surface in terms of complexity (*Asfc*), scale of maximum complexity (*Smc*), anisotropy (*epLsar*), heterogeneity (*HAsfc*), and

textural fill volume (*Tfv*) (El Zaatari, 2010; Goodall, Darras, & Purnell, 2015; Krueger et al., 2008; Scott et al., 2006). Figure 2.3 and Table 2.1 define and illustrate these features.

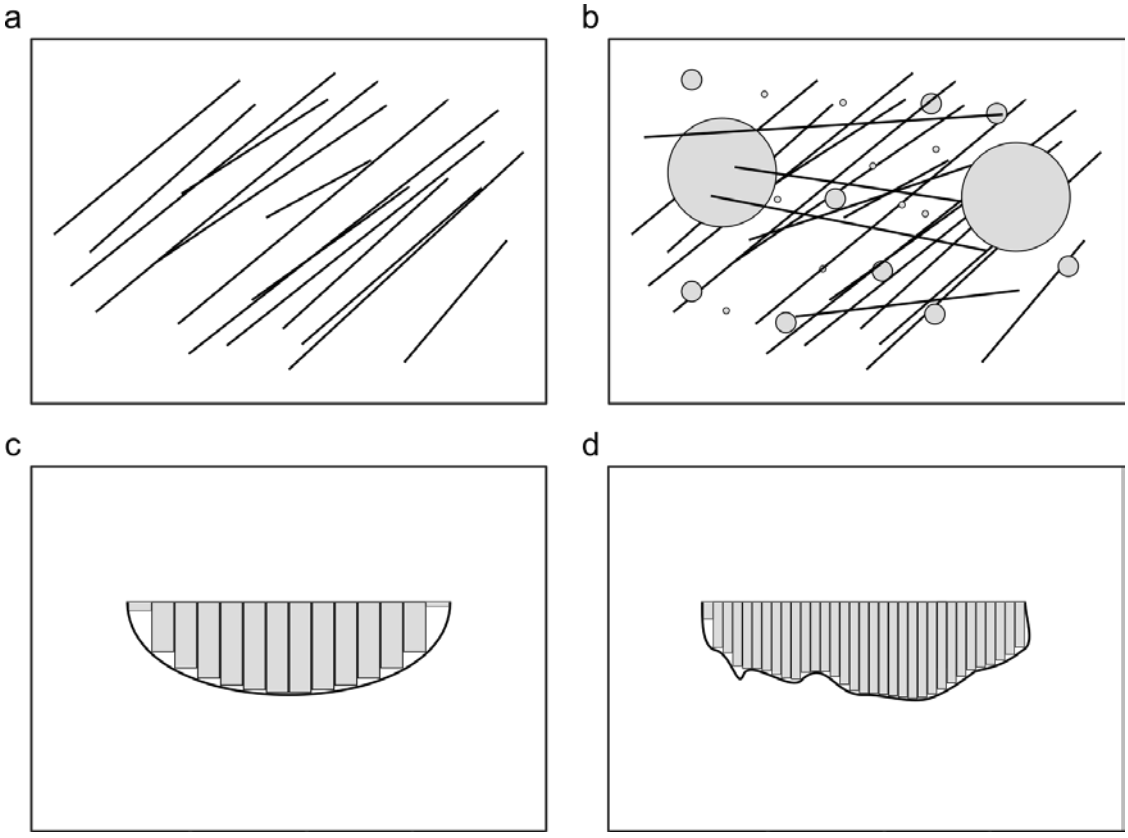


Figure 2.3 Schematic example of microwear features: (a) represents a microwear *surface* with high anisotropic texture (*epLsar*); (b) represents a microwear *surface* with complex texture (*Asfc*); while (c) and (d) are cross sections of a microwear surface that have been filled with square prisms; texture fill volume (*Tfv*) is the result of subtracting (c), the overall shape of the surface, from (d), the texture shape of the surface. Adapted from Scott et al. (2006).

Despite offering a solution to the high intra-observer error of earlier techniques, and thus being an exciting new method, DMTA still suffers from the same limitations that plague microwear studies. First, only some foods leave traces on the tooth surface, and second, microwear surfaces have high turnover rates – also known as the 'Last Supper Effect' (Teaford, 2007a; Ungar & Lucas, 2010).

Grit, phytoliths, or acid are some of the few elements that can leave marks on the teeth. Yet, even though phytoliths are commonly used to explain wear patterns (Reinhard & Danielson, 2005), they may not be strong enough to leave traces in the tooth enamel (Sanson et al., 2007). Moreover, indirect effects of food on teeth might cause most of the observed microwear patterns – for instance, through the introduction of abrasives into

food during processing stages (Grine et al., 2012; Teaford, 2007a). The inverse is also possible, where food processing removes most food elements that could cause striations, which is likely the case among western societies (Teaford, 2007a). However, the 'Last Supper Effect' need not be a handicap. Dental microwear only records the effects of the most recently eaten foods on the teeth, so it gives precise information on the diet of an individual at a given point of his or her life (Grine et al., 2012; Teaford, 2007a).

Considering these limitations and the relative novelty of DMTA, it is crucial to understand how microwear patterns vary between modern human populations with different diets. Interestingly, the need to study animals consuming known foods has been identified for some time (Teaford & Oyen, 1989b). Despite this, *in vivo* microwear studies in animals are still rare (Teaford & Glander, 1991; Teaford & Oyen, 1989a). Instead, microwear work (SEM and DMTA) in modern populations (animal and human) mostly uses museum remains (DeSantis et al., 2013; El Zaatari, 2010; Romero, Galbany, Pérez-Pérez, & De Juan, 2007; Scott, Teaford, & Ungar, 2012). Notwithstanding, such work has produced very thought-provoking results. For instance, one study found that different groups of hunter gatherers differed in the abrasiveness of their diet (El Zaatari, 2010), while a second study found that pastoralists had lower microwear texture values than agriculturalists, interpreted as resulting from a high meat consumption in the first group (Schmidt, Beach, McKinley, & Eng, 2016). Nonetheless, there is still no reliable evidence on how dental microwear differs among living human populations. The only exception, Romero et al. (2013) used SEM technology to compare the buccal microwear of live pygmy Baka hunter gatherers to archaeological samples of pygmies, Bantu pastoralists, and farmers. The results showed that the Baka hunter gatherers consumed low amounts of abrasive foods in relation to the other groups studied, and that this pattern is consistent over time.

Table 2.1

Scale sensitive fractal analysis parameters used to describe a microwear surface. Adapted from Goodall, Darras, & Purnell (2015), Krueger et al. (2008), and Scott et al. (2005).

Parameter Name	Description	Example	Diet
Area scale fractal complexity (*Asfc*)	Measure of change in roughness with scale	A complex surface may have pits and scratches of different sizes overlaying one another	A complex surface indicates a hard diet
Scale of maximum complexity (*Smc*)	Scale at which the surface is most complex	Coarse features (large pits, no fine scratches) may dominate a surface with high *Smc*	High *Smc* indicates a hard diet
Exact proportion Length scale anisotropy of relief (*epLsar*)	Variation in lengths of transect lines measured at a given scale with orientations sampled at 5° intervals across a surface	An anisotropic surface may have a surface relief with similar orientation (i.e. more parallel scratches)	A tough diet will be more anisotropic
Heterogeneity of Area scale fractal complexity (*HAsfc*)	Variation of *Asfc* across a surface	High heterogeneity may indicate differing patterns of scratching and pitting across the surface	A mixed diet will be more heterogenic
Textural fill volume (*Tfv*)	Function of two components: (1) the shape of the surface, and (2) the texture of the surface	Deep features may dominate a surface with high *Tfv*	High *Tfv* indicates a hard diet

2.5 STABLE ISOTOPE RATIOS ANALYSIS

Stable isotope ratio analyses were previously discussed in this work in relation to tooth and bone chemistry studies. These studies have become one of the most frequently used methods in archaeological science because they offer a semi-quantitative approach to diet assessment, and because they often provide the only means of detecting dietary differences across individuals. Nevertheless, as for dental microwear texture analysis, to correctly apply and interpret stable isotope ratio analysis, one must understand the chemical basis of the technique, and of the isotope fractionation in both the environment and the body (Makarewicz & Sealy, 2015; O'Brien, 2015). The latter is particularly relevant to this research because it deals with the trajectory of chemical elements in different isotopic body pools – hair, nail, and breath – in order to build a model applicable to archaeological populations whose most commonly studied tissues are collagen and bioapatite.

2.5.1 STABLE ISOTOPE RATIO TECHNIQUE

Isotopes are chemical elements that have the same number of protons and electrons (and so share the same chemical properties), but vary in the number of neutrons (and therefore in mass), resulting in different reaction rates. This process is termed fractionation and results in different concentrations of each isotope in the products and substrates of the reaction. Thus, as different isotopes of a given element circulate through the ecosystem, different stable isotope ratios will characterize the different components of the system, i.e. the transfer of an element from the ocean to the atmosphere, or from plant to animal tissue, is associated with sequential changes to the natural abundance of isotope ratios (Hoefs, 2004; Sharp, 2007). There are two types of fractionation: thermodynamic and kinetic. The first occurs in systems that are at equilibrium, where temperature determines the speed of reaction (e.g. bone mineral synthesis). Conversely, kinetic isotope fractionation occurs in incomplete or unidirectional reactions, i.e. not at equilibrium, where reaction speed is determined by enzymatic control (e.g. bone collagen synthesis) (Hoefs, 2004; Sharp, 2007).

Isotope fractionation is most noticeable in light elements, like hydrogen (H), carbon (C), nitrogen (N) and oxygen (O), where mass differences between isotopes are relatively large. Heavier elements, like strontium, will not register such difference in chemical reaction speed between isotopes. In addition, isotopes with higher atomic mass are termed heavy isotopes, while the ones with lower atomic mass are light isotopes, which is not to be confused with the light and heavy elements mentioned above. For instance, ^{12}C and ^{13}C are, respectively, the light and heavy isotope of the light element carbon, and the relative difference between the overall masses leads to a faster reaction time for ^{12}C than for and ^{13}C (Hoefs, 2004; Sharp, 2007). The naturally occurring isotopes and respective abundance of the two elements studied in this thesis – C and N – are found in Table 2.2.

Table 2.2
Carbon and nitrogen isotopes, abundance, and international standards (Schoeller, 1999, p. 668).

Element	Stable Isotope	Abundance (%)	International Standard
Carbon	^{12}C	98.889	Pedee Belemnite Limestone (PDB)
	^{13}C	1.111	
Nitrogen	^{14}N	99.634	Atmospheric N_2 (AIR)
	^{15}N	0.366	

For results to be comparable across studies, isotope ratios are reported as delta (δ) values, and are calculated as the difference between the isotope ratio of the sample and that of an internationally recognized standard. Furthermore, since differences in natural abundance of isotopes are very small, delta values are expressed in units *per mil* (‰). This is formalised in Equation 2.1 where X represents the heavy isotope and R represents the ratio of heavy to light isotope. Standards used for carbon and nitrogen are found in Table 2.2. Also, since PDB, the standard for carbon, has a higher abundance of ^{13}C than most biological tissues, $\delta^{13}C$ values in these materials are negative (Hoefs, 2004; Sharp, 2007). Finally, differences between analysed samples are commonly expressed as $\Delta_{\text{sample 1- sample 2}}$.

$$\delta X = \frac{R_{sample} - R_{standard}}{R_{standard}} \times 1000 \quad (2.1)$$

Stable isotope ratios are measured through isotope ratio mass spectrometry (IRMS). This analysis requires gaseous samples, which is achieved through an inlet system used in tandem with the IRMS in a continuous flow system (CF-IRMS). The inlet system used for solid samples (e.g. keratin and collagen) is an element analyser (EA), while the system used for complex mixtures (e.g. breath) is a gas chromatograph (GC) (see Figure 2.4) (Hoefs, 2004; Muccio & Jackson, 2009; O'Brien, 2015).

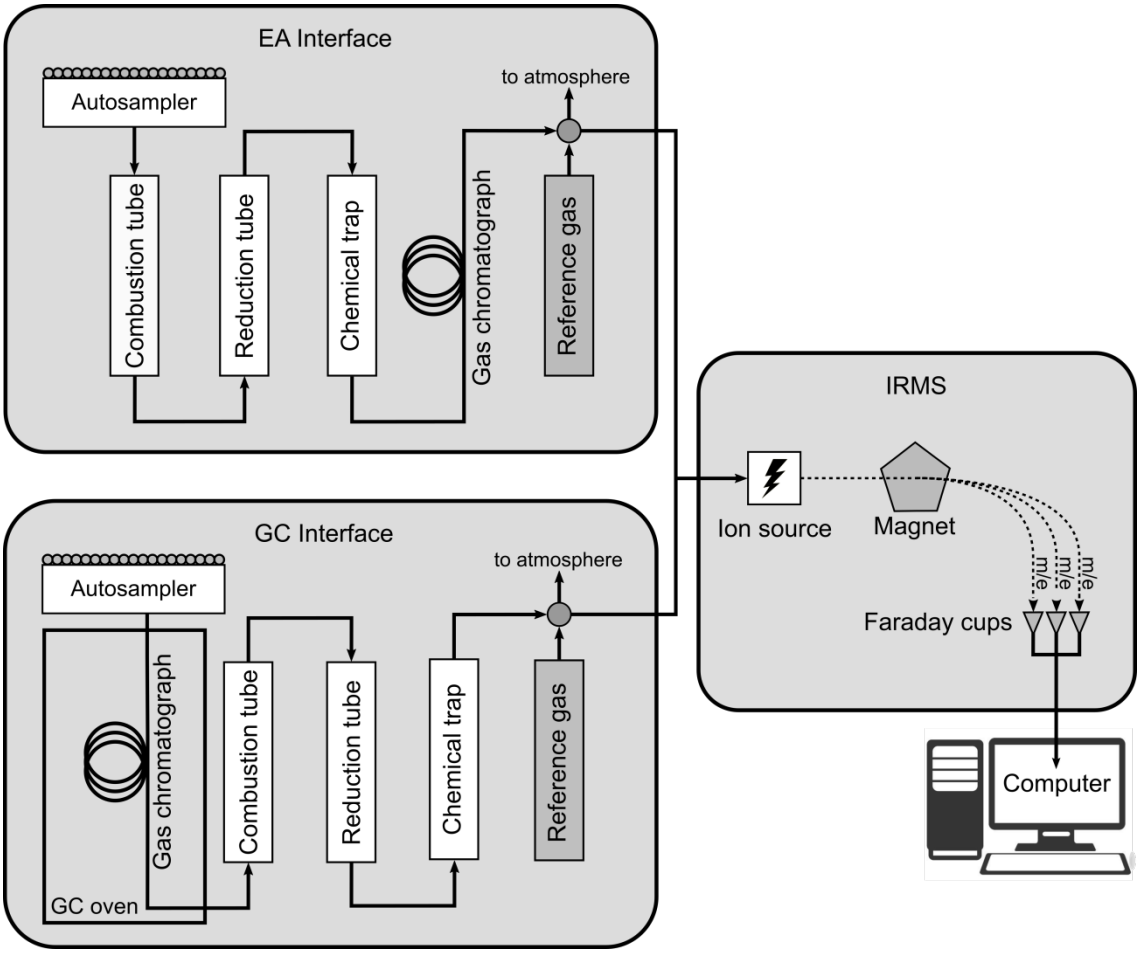


Figure 2.4 Schematic of an element analyser isotope ratio mass spectrometer (EA-IRMS), m/e = mass/energy. Adapted from Muccio & Jackson (2009).

In an EA, a sample encased in a tin capsule is dropped in a combustion tube at 1000 °C in an oxygen pulse. The tin capsule burns exothermically, raising temperatures to *circa* 1800 °C. This step produces carbon dioxide (CO_2), nitric oxide and nitrogen dioxide (NO_x), water (H_2O) and sulphur dioxide (SO_2), which is then fed to a reduction tube at 600 °C where NO_x is reduced to N_2 and excess O_2 is removed. A chemical trap then removes water produced by the combustion. The resulting analyte – N_2 and CO_2 – is

carried in a helium stream, separated by a gas chromatograph, and fed separately to the IRMS. An open split valve allows 2% of the sample gases to flow into the mass spectrometer while the rest is released into the atmosphere. The same valve allows the introduction of international reference gases in alternation with sample gases (Hoefs, 2004; Muccio & Jackson, 2009; O'Brien, 2015).

On the other hand, a GC works in a similar way, except that samples are separated prior to any other step. In this technique, the contents of samples and in-house standards are eluted from the sample vials via a double holed needle into a column gas chromatograph, held at 60°C. The next steps are the same as those of EA: samples are combusted into a combination of gases (e.g. CO_2, NO_x, and H_2O); NO_x is reduced to N_2, and O_2 removed; H_2O is eliminated; and the resulting gases are passed on unto the IRMS through a valve that also allows the injection of international reference gases into the system (Hoefs, 2004).

Subsequently, gases produced by either an EA or a GC enter the IRMS at the ion source, where they are converted into positively charged ions. These are then focused into a beam and pass into a flight tube where a magnetic field deflects their path to different degrees based on their mass to charge ratio (m/e in Figure 2.4). Heavier molecules deflect more than lighter ones with the same charge and so strike different ion detectors (Faraday cups) placed at the end of the flight tube. Produced voltage intensity is proportional to the quantity of ions collected and is recorded through a computer. For nitrogen, detected ions have atomic masses of 28, 29 and 30 ($^{14}N_2^{+1}$, $^{14}N^{15}N^{+1}$, $^{15}N_2^{+1}$) while for carbon, ions have atomic masses of 44, 45 and 46 ($^{12}C^{16}O_2^{+1}$, $^{13}C^{16}O_2^{+1}$, $^{12}C^{16}O^{18}O^{+1}$) (Hoefs, 2004; Muccio & Jackson, 2009; O'Brien, 2015).

After understanding the theories that guide stable isotope ratio analysis, and to interpret the results produced by the mass spectrometer, one must understand how carbon and nitrogen isotopes behave in both the environment and the body.

2.5.2 *ISOTOPIC FRACTIONATION IN THE ENVIRONMENT*

Carbon

The carbon isotope patterning of the biosphere is mainly determined by kinetic fractionation steps associated with carbon fixation of atmospheric CO_2 in plants (see Figure 2.5). However, carbon isotope fractionation is also influenced by any environmental factors that change the source CO_2 or the internal versus external concentration of this gas in the leaf. In general, environmental factors can modify carbon isotope fractionation but internal physiology constrains the response (Cernusak et al., 2013; Farquhar, Ehleringer, & Hubick, 1989; Smith & Epstein, 1971).

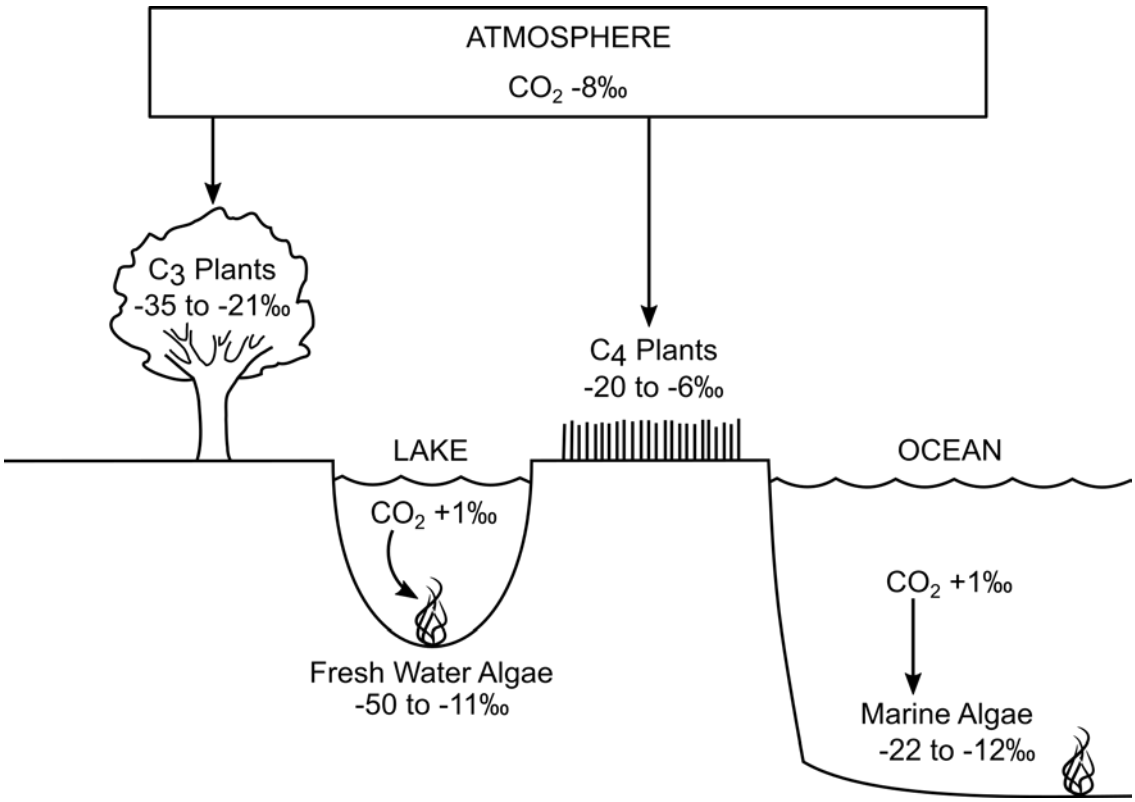


Figure 2.5 The distribution of $\delta^{13}C$ in the ecosystem. Arrows indicate carbon fluxes. Numbers for pools indicate $\delta^{13}C$ values. Adapted from Pate (1994).

At the baseline, atmospheric CO_2 has a $\delta^{13}C$ value around -8‰. This CO_2 is then taken up into plants by diffusing into the internal air space through stomata in the leaves. Stomata are organs that control both gas exchange and water loss in the plant. Overall, plants have lower $\delta^{13}C$ values than atmospheric CO_2 because the enzyme (Rubisco) that binds and fixates CO_2 discriminates against the heavier ^{13}C isotope. The extent of this

fractionation, however, differs according to plant carbon uptake mechanisms, of which there are three: the C_3 pathway characterizes almost all species native to temperate environments, such as trees or herbaceous plants; the C_4 or Hatch-Slack is a modification of the C_3 pathway and is typical of plants native to dry environments, such as tropical and salt grasses; and the CAM (Crassulacean Acid Metabolism) is a modification of the C_4 pathway that characterises succulents, agaves, and bromeliads (e.g. pineapple). C_4 plants discriminate less against ^{13}C than C_3 plants because the first have evolved a CO_2 concentrating mechanism that reduces water loss but also reduces the extent to which Rubisco discriminates ^{13}C. As a result, C_3 plants $\delta^{13}C$ values range between −35‰ and −21‰, C_4 plants between −20‰ and −6‰, and CAM between −33‰ and −14‰ (Bender, Rouhani, Vines, & Black, 1973; Farquhar et al., 1989; O'Leary, 1988; Sealy, 2001; Smith & Epstein, 1971). Considering that few CAM plants are important food sources, they will not be discussed further.

As mentioned before, environmental factors also modify carbon isotope fractionation in terrestrial plants. These environmental factors include all those that alter the rate of photosynthesis or stomatal conductance or, as put above, any that changes the source CO_2 or the internal versus external concentration of this gas in the leaf. A complete review of these factors (atmospheric CO_2 concentration, water, nutrient, and light availability, temperature, altitude, and atmospheric pressure) and their effects can be found in Cernusak et al. (2013). A further intriguing environmental factor is the 'canopy effect', wherein unusually ^{13}C depleted carbon isotope ratios in closed forests are explained by the recycling of depleted carbon released by decaying organic material on the forest floor and by variations at the leaf level in response to decreased sunlight (Bonafini, Pellegrini, Ditchfield, & Pollard, 2013; Schoeninger, 2010). Evidently, the amount and diversity of environmental factors that influence fractionation help explain the high variation in $\delta^{13}C$ values observed across all photosynthetic pathways.

Carbon isotope patterning in the oceans is more complex and less well understood. To start, aquatic plants use both atmospheric and dissolved CO_2 for photosynthesis. The

latter is derived from different sources, but bicarbonate, dissolved inorganic carbon, is one of the main sources. In fact, bicarbonate sequesters most of the world's active cycling carbon. Most of the CO_2 present in water is converted to bicarbonate through an equilibrium isotope fractionation that favours ^{13}C. As a result, bicarbonate has a $\delta^{13}C$ value around +1‰ – +9‰ higher than atmospheric CO_2 (DeNiro & Epstein, 1978; Mook, Bommerson, & Staverman, 1974; Smith & Epstein, 1971). Regarding the photosynthetic pathway, most aquatic plants are C_3 although a few C_4 exist, with consequent fractionation processes (Ueno, Samejima, Muto, & Miyachi, 1988). This produces $\delta^{13}C$ for marine plants values between −22‰ and −12‰ (Smith & Epstein, 1971). As in their terrestrial counterparts, $\delta^{13}C$ values in aquatic plants also vary according to environmental and physiological parameters. Other than the above-mentioned CO_2 source, an example is water flow rate: considering the diffusion of CO_2 in water is slower than in air, CO_2 will be more readily available in fast than in slow flowing water, resulting in lower $\delta^{13}C$ values for the first (Osmond, Vallane, Haslam, Uotila, & Roksandic, 1981). Notwithstanding, these effects average as one moves up the food chain, and so, fish and marine predators differ from animals in a terrestrial C_3 food web – but overlap with C_4 $\delta^{13}C$ values (Schoeninger & DeNiro, 1984). Freshwater aquatic plants deserve a final comment as they are even less studied than marine plants. These plants have very variable $\delta^{13}C$ values, ranging from -50‰ to -11‰ and overlapping with those of terrestrial plants (Keeley & Sandquist, 1992).

Following this, the $\delta^{13}C$ values of animals reflect the plant diet that is at the base of the food chain. C_3 consumers have lower δ ^{13}C values than those that eat C_4 plants (DeNiro & Epstein, 1978). Furthermore, trophic shifts are often modest and hard to estimate (McCutchan, Lewis, Kendall, & McGrath, 2003). A subsequent section on the isotopic fractionation in the human body explores the physiological reasons behind this small trophic shift, as well as the differences between tissues of the same individual.

Nitrogen

To understand the isotope patterning of nitrogen in the biosphere, one must understand how nitrogen fractionates across the food chain (see Figure 2.6). In broad

terms, nitrogen patterning in plants reflects the fractionation associated with nitrogen fixation, but also the different nitrogen sources, the net nitrogen loss, and the mixing of nitrogen pools within the plant. Furthermore, as for carbon, environmental factors also affect $\delta^{15}N$ values by altering the source nitrogen and the extent of fractionation. Finally, nitrogen patterning in animals reflects trophic position (DeNiro & Epstein, 1981; Robinson, 2001).

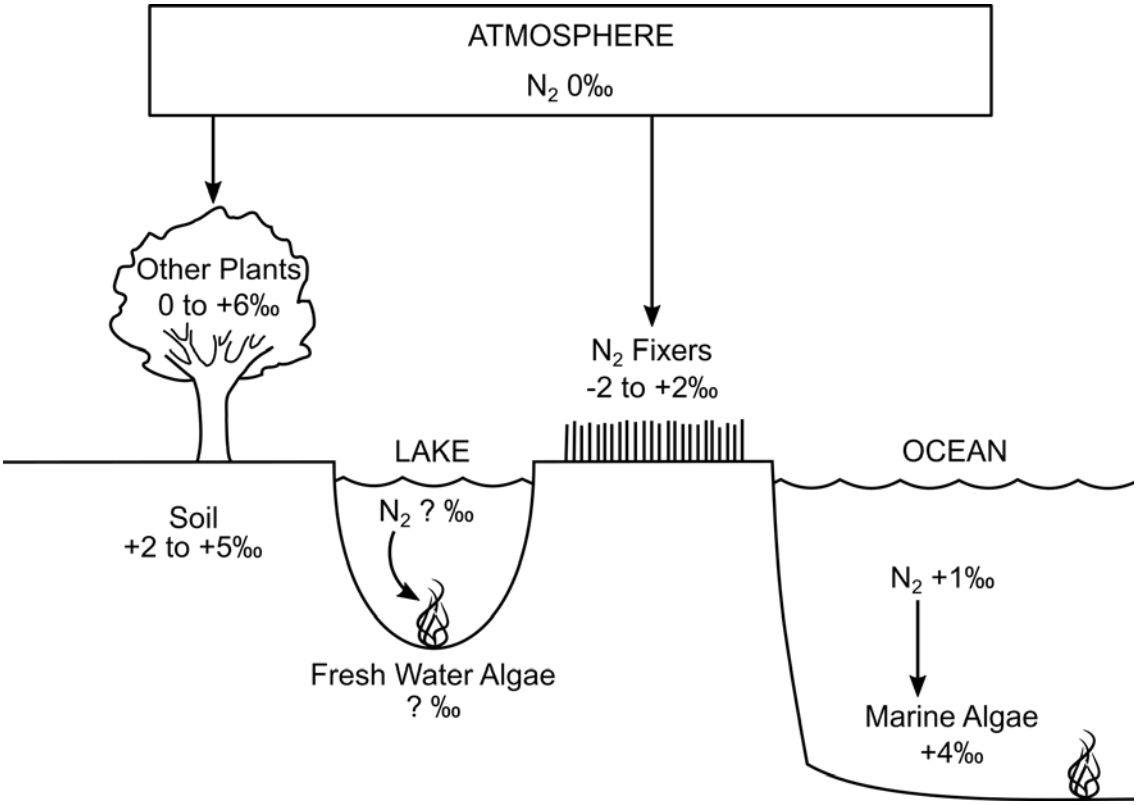


Figure 2.6 The distribution of $\delta^{15}N$ in the ecosystem. Arrows indicate nitrogen fluxes. Numbers for pools indicate $\delta^{15}N$ values. Adapted from Pate (1994).

Most terrestrial plants take up nitrogen directly from the soil but some fix nitrogen from the air spaces in the soil, i.e. from the atmosphere. The proportion between assimilated nitrogen sources depends on its availability and on plant type (Robinson, 2001). On the one hand, the isotopic value of N_2 present in the atmosphere is 0‰ – as this is the standard used in $\delta^{15}N$ calculations – and its transfer to the biosphere depends on bacteria, found in the roots of legumes, or on fungi, found in the biological crust of arid soils (mycorrhyzae). This process has little associated fractionation and $\delta^{15}N$ values of such plants range from −2 to +2‰. On the other hand, ammonium

(NH_4^+) and nitrate (NO_3^-) available in soils originates from the bacterial decomposition of organic matter and from rain. Decomposition processes result in ^{14}N depletion, and so, the isotopic value of soil is often around +2 to +5‰, while that of non-leguminous plants is around 0 to +6‰ (DeNiro & Epstein, 1981; Pate, 1994; Sharp, 2007). Further variation in $\delta^{15}N$ values of plants is a result of (1) nitrogen loss by plants through efflux from roots or volatilization from leaves, and (2) nitrogen movement between distinct pools within the plant. As stated above, the $\delta^{15}N$ of a plant reflects the $\delta^{15}N$ of the nitrogen source(s), nitrogen isotope fractionation, nitrogen gain and loss, and nitrogen pool mixing (Robinson, 2001).

Sources of environmental variation in $\delta^{15}N$ values in terrestrial plants include factors affecting nitrogen composition of soil and fractionation in plants. An example of the first is salinity. Saline soils have higher $\delta^{15}N$ values than non-saline soils, likely due to the higher ^{15}N content of soil nitrate and ammonium in saline environments (Heaton, 1987; Pate, 1994). An example of nitrogen fractionation in plants is nutrient availability. Rich soils have higher $\delta^{15}N$ values because when nitrogen is available, fractionation is high, but when nitrogen is limiting, fractionation is low (McKee, Feller, Popp, & Wanek, 2002). Besides salinity and nutrient availability, other important environmental factors encompass water availability, soil depth, and type and frequency of mycorrhyzae (Robinson, 2001).

In aquatic environments, N_2 dissolved in the ocean has a $\delta^{15}N$ value of about +1‰ and is converted into ^{15}N enriched NH_4^+ and NO_3^-. Furthermore, nitrogen fixation by bacteria and blue green algae provides an additional source of nitrogen, which leads to aquatic plants enriched in ^{15}N in relation to terrestrial plants, and with $\delta^{15}N$ values around +4‰. These values, however, overlap with those of non N_2 fixing terrestrial plants (Sharp, 2007; Wada, Kadonaga, & Matsuo, 1975).

Following up the food chain, $\delta\ ^{15}N$ values of animals relate to their trophic position. Because nitrogen excretion in animals discriminates against ^{15}N, the retention of the heavier isotope results in a trophic level enrichment of about 3-5‰, although it can be higher. Consequently, fish will often have higher $\delta^{15}N$ values than terrestrial animals

because food chains tend to be longer in the marine than in the terrestrial environment (DeNiro & Epstein, 1981; O'Connell, Kneale, Tasevska, & Kuhnle, 2012). However, this signal may be confounded by a variety of factors, such as water or caloric stress associated with desert areas (Hartman, 2011; Reitsema, 2013; Tuross, 2017). As for carbon, the particulars of how nitrogen fractions across different tissues of the body are discussed in the next section.

2.5.3 ISOTOPIC FRACTIONATION IN THE HUMAN BODY

Different tissues within an organism often have different isotopic values. This variation is the result of organism's physiological disparities, such as distinct turnover rates of isotopic pools within the body, and divergent routing from diet into the body. In consequence, different tissues may represent different time periods of an individual's life and have a signature of different parts of its diet (O'Connell et al., 2001). This confuses the interpretation of isotopic results. For instance, Hülsemann et al. (2015) used an offset to transform nail to hair isotopic values, but O'Connell et al. (2001) did not find a significant difference between the two tissues, indicating they could be pooled together without that step.

The first point to address on isotopic fractionation in the human body is elemental turnover. Organisms are in a constant process of remodelling wherein components are lost through waste, sweat, breath, sloughed cells, etc., while new components are incorporated from the diet. Tissues vary in their rate of elemental turnover and consequently represent dietary information from varying time frames. For instance, tooth enamel does not remodel, and so, it represents the isotopic signal of the diet at the time in which it was formed (Loftus & Sealy, 2012). On the other hand, bone collagen has a turnover period of more than ten years, although turnover rates vary depending on age, sex, health and the type of bone analysed (Hedges, Clement, Thomas, & O'Connell, 2007; O'Connell & Hedges, 1999b; Stenhouse & Baxter, 1979). Furthermore, as in the case of enamel, hair and nail record the isotopic signal at the time they were formed, but because they grow continuously, that record is replaced

throughout the individual's life: 1 centimetre (cm) of nail represents three months' growth (Zaias, 1980), while 1 cm of hair represents one month growth (Saitoh, Uzuka, Sakamoto, & Kobori, 1969). However, the turnover ratios associated with the amino acid pools supplying hair synthesis ensue that it may take several months for individuals' hair to reach an isotopic steady state after a diet change (Ayliffe et al., 2004; Hülsemann, Flenker, Koehler, & Schaenzer, 2009). Conversely, breath will signal a diet change after only three hours, simply associated with digestion time (Ayliffe et al., 2004). Finally, and for all body pools mentioned, changes in $\delta^{15}N$ take place faster than in $\delta^{13}C$, due to differences in the way carbon and nitrogen are processed in the body. In detail, de- and transamination move nitrogen from one amino acid to another, leading to a faster equilibrium of amino groups, while the carbon skeletons of the amino acids are more stable (Hülsemann et al., 2009; O'Connell & Hedges, 1999a).

The second point on this theme is diet to tissue routing. This matter is straightforward in nitrogen considering that 98% of the nitrogen in the body is found in proteins and its constituent amino acids (Schoeller, 1999). As a result, $\delta^{15}N$ represents the protein part of the diet, in an almost perfect protein to protein routing. On the other hand, carbon is present in all major macronutrients – proteins, carbohydrates, and lipids – and so, the mechanisms involved are more complicated. This is explored further below.

Carbon

Two models have been proposed to explain the flow of carbon into the body: the scrambling and the protein routing models. The first suggests that all macronutrients are equally integrated into tissues so that $\delta^{13}C$ reflects total dietary carbon, while the second suggests protein to protein routing, so that $\delta^{13}C$ of tissue *protein* reflects dietary protein (Ambrose & Norr, 1993; Chisholm, Nelson, & Schwarcz, 1982; Hedges & van Klinken, 2000; Tieszen & Fagre, 1993; Vogel & van der Merwe, 1977). The most likely hypothesis is that a combination of the models applies to different degrees in different circumstances, resulting in varying discrimination factors between tissues. To help clarify this matter, Figure 2.7 illustrates the flux of carbon across the body. In this flow model, largely based on Hedges and van Klinken (2000), the influx of carbon into the

body is diet, which is subject to several processes of mechanical and chemical digestion, and is eventually absorbed in the digestive tract. Diet has three main components: protein, carbohydrates, and lipids. The out flux from the system is mostly composed of respired CO_2, although some is excreted through waste. The body has several body tissue or pool components, of which several can be measured isotopically: hair and nail, bone collagen, breath, and bioapatite and tooth enamel. Importantly, the first three tissues are made of protein while the other three are not. Other key components of the model are the pool of amino acids and energy metabolism, which encompasses the anaerobic glycolysis in the cell's cytoplasm and the aerobic Krebs cycle in the cell's mitochondria.

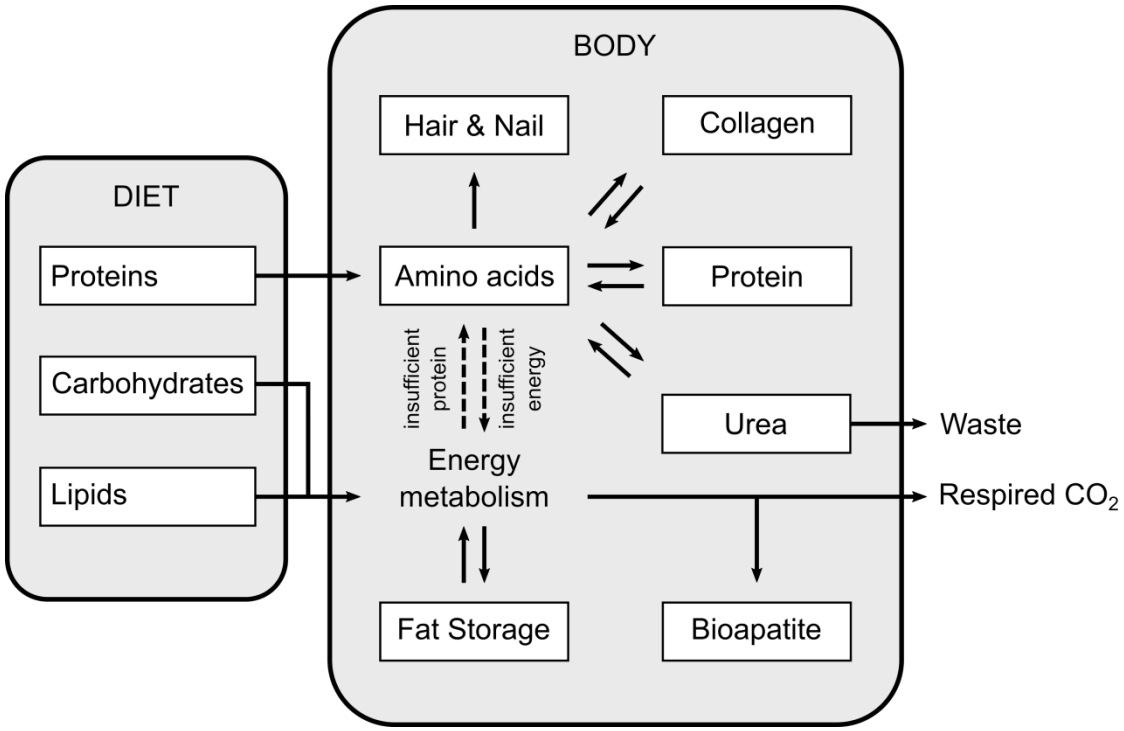


Figure 2.7 Model of the flux of carbon from diet to the body. Arrows indicate the direction of the flux and are likely associated with fractionation. Adapted from Hedges & van Klinken (2000).

Next, one must address the movement of the dietary components, also termed macronutrients, within the system. Namely, ingested protein is denaturated in the stomach and hydrolysed to amino acids and polypeptides, which are then absorbed and transported to the liver. Most of these amino acids are used in protein production within each cell, while amino acids in excess may be used as a source of energy, particularly if energy consumption from carbohydrates and lipids is insufficient.

Producing energy from amino acids, however, is highly inefficient. Even though proteins are as energy dense as carbohydrates, most of that energy is lost as heat during deamination. This process removes the nitrogen (in the form of the amino group (NH_2)) from the amino acid, transforming them into keto acids, which are then transformed into acetyl sugars usable by the Krebs cycle. In the case of humans, ammonia (NH_3) produced by deamination is transformed into urea ($CO(NH_2)_2$) by the liver, which is then excreted (Schoeller, 1999; WHO/FAO/UNU, 2007). In a similar way to protein, carbohydrates and lipids are hydrolysed into simple sugars and fatty acids, respectively, absorbed in the intestine, and then transported to the cells of the body, where they are used to generate energy through glycolysis and the Krebs cycle mentioned above. Notably, carbohydrates are more efficient for energy production, while fat is more suitable for energy storage. In fact, excess carbohydrates enter the fatty acid pathway and are stored as fat. The reasons behind this are twofold: (1) the hydrophobic nature of lipids makes them a more compact form of energy storage than hydrophilic carbohydrates, which easily diffuse in the cell milieu, becoming easily available for energy production, and (2) despite the higher energy content of lipids, transport into the mitochondria and oxidation therein for energy yield are inefficient processes. Interestingly, when the body is breaking down fat at a higher rate than usual, the Krebs cycle becomes flooded with the acetyl sugars being produced by fat oxidation (partly due to the higher energy content of fat versus carbohydrates), and so some of these acetyl groups are reversibly transformed into keto acids (in an inverse reaction to that that follows protein deamination). Finally, this energy metabolism (from carbohydrates, lipids, and protein) produces CO_2, which is released into the blood stream and eventually into the atmosphere through respiration; almost 90% of the ingested carbon is eventually excreted this way (FAO/WHO, 1998; FAO, 2010; Hedges & van Klinken, 2000).

Besides the overall behaviour of macronutrients within the body, different amino acids move differently through the metabolism, with consequences to protein routing. To be more precise, there are 20 common amino acids and several minor amino acids. Common amino acids are divided into indispensable and dispensable: indispensable

amino acids represent only 12.2% of the total but cannot be synthesized by the organism, while dispensable amino acids can be synthesised by the organism from glycolysis or Krebs cycle metabolites, although this has significant energy costs. This has two consequences: (1) indispensable amino acids are always derived directly from dietary protein, and (2) dispensable amino acids will only be produced by the body if protein consumption is insufficient. In addition, at least five dispensable amino acids have particularities that result in their preferential protein routing: tyrosine and hydroxylisine can only be synthesised from indispensable amino acid precursors, while glycine, proline and hydroxylproline require four enzymes to be synthesized *de novo* (versus one enzyme for the rest of dispensable amino acids), with consequent high energy costs (Ambrose, Butler, Hanson, Hunter-Anderson, & Krueger, 1997; Hedges & van Klinken, 2000; Schoeller, 1999). Notwithstanding, there are other confounding factors on protein metabolism: (1) urea recycling in most animals results in the synthesis of amino acids from the nitrogen of dietary and own tissue protein, and from the carbon skeletons of other dietary components (Houpt, 1963), and (2) microbial activity in herbivores metabolises dietary components to non protein nitrogen and so, protein synthesis partly 'mixes' carbon and nitrogen from all dietary components (Fogel & Tuross, 2003; Houpt & Houpt, 1971). In conclusion, in most animals, carbon in protein is biased towards dietary protein, although not only representative of it. Due to amino acid composition, collagen is theorized to have around 65% protein routing of carbon, which has been confirmed in feeding experiments (Ambrose & Norr, 1993; Froehle, Kellner, & Schoeninger, 2010; Jim, Jones, Ambrose, & Evershed, 2006), whereas hair protein (keratin) is theorized to have about 25% protein routing, again confirmed in an experimental setting (Hedges et al., 2009).

Carbon isotopic fractionation occurs during amino acid metabolism; however, the effects tend to be modest and it is still unclear where or why they take place (Ambrose & Norr, 1993; Tieszen & Fagre, 1993). Collagen-diet spacing varies between 3.7 to 6‰ and is largely explained by its high glycine content (31.9% of total amino acids), an amino acid enriched in ^{13}C relative to other amino acids. In fact, collagen is enriched in ^{13}C in relation to muscle protein, for instance, by about 2‰ (Bocherens & Drucker,

2003; DeNiro & Epstein, 1978; Hare, Fogel, Stafford Jr., Mitchell, & Hoering, 1991; Hedges & van Klinken, 2000; Jim et al., 2006; O'Connell et al., 2001).

In contrast, collagen is only 1-2‰ enriched in relation to hair or nail keratin because keratin is rich in cysteine and serine (28.1% of total amino acids in hair and 21.9% in nail), which are metabolically close to glycine and have similar enrichment in ^{13}C (O'Connell et al., 2001). Consequently, hair and nail are also enriched in ^{13}C relative to dietary protein by 1-2‰ (DeNiro & Epstein, 1978; O'Connell et al., 2001). Finally, modest differences are found between hair and nail $\delta^{13}C$ (Hülsemann et al., 2015; O'Connell et al., 2001).

As mentioned before, circulatory CO_2 is a by-product of the energy metabolism that is released into the atmosphere in breath. In humans, this respired CO_2 is ^{13}C-depleted in relation to diet by 1-2‰ (Hedges et al., 2009). This is not due to a fractionation step between blood and breath CO_2, although its existence cannot be excluded, but due to the oxidation of ^{13}C-depleted lipids. Lipid synthesis discriminates against ^{13}C and so, when the body oxidizes lipids to obtain energy, it feeds ^{13}C-depleted CO_2 into the blood stream (DeNiro & Epstein, 1977). Additionally, breath $\delta^{13}C$ varies throughout the day according to the proportions of protein, carbohydrates, and fat being oxidised for energy (Hedges et al., 2009; Passey et al., 2005; Schoeller et al., 1984).

Although most of the dissolved carbon in blood is released in breath, some will find its way into the bioapatite in bones and enamel in teeth. This process has an isotopic fractionation that is dependent on body size, and that averages around 12-13‰ for primates (Crowley et al., 2010; Passey et al., 2005). In summary, breath, bioapatite, and enamel $\delta^{13}C$ all track the whole dietary carbon, and not only protein. Experimental support for this assertion is found in controlled fed animal studies (Jim, Ambrose, & Evershed, 2004) and in human studies where sugar consumption and energy balance was successfully tracked through the $\delta^{13}C$ of breath and blood components (Fakhouri et al., 2014; Nash, Kristal, Hopkins, Boyer, & O'Brien, 2014; Whigham et al., 2014).

A final aspect of carbon fractionation in the body concerns collagen-bioapatite spacing. Added effects of bioapatite precipitation, and lipid metabolism result in a spacing of about 7‰ for herbivores and 4‰ for carnivores, with omnivores falling somewhere in the middle (O'Connell & Hedges, 2017; Passey et al., 2005). Therefore, protein metabolism likely plays a minimal role in this spacing, considering that collagen-bioapatite spacing differences between carnivores and herbivores result from differences in bioapatite-diet spacing and not from collagen-diet spacing (O'Connell & Hedges, 2017).

Summing up, particulars of the body physiology translate in $\delta^{13}C$ differences between tissues, which must be taken into account in analyses that compare $\delta^{13}C$ results across an organism's tissues. Notwithstanding, these tissue spacings cancel each other out to some extent, resulting in small differences between consumers and diet when analysed whole. In other words, trophic shifts in carbon are modest, especially when compared with the differences between C_3 and C_4 plants at the base of the trophic chain, and hence $\delta^{13}C$ values primarily track that plant diet (DeNiro & Epstein, 1978; McCutchan et al., 2003).

Nitrogen

Compared to carbon, nitrogen metabolism is simpler because 98% of the nitrogen in the body is found in proteins, thus $\delta^{15}N$ values always represent the protein part of the diet (Schoeller, 1999). For this reason, stable nitrogen isotopes cannot be determined for some of the tissues discussed above: bioapatite, tooth enamel, and breath. In order to understand nitrogen metabolism, Figure 2.8 models the flux of nitrogen through the body. The main influx of nitrogen into the body is the protein part of the diet, while the main out flux is waste. In fact, 80-90% of the ingested nitrogen is excreted as urea (Hedges & van Klinken, 2000). Hair, nail, and collagen can be analysed isotopically, and other important components of the model are the pool of amino acids and urea. From the previous section, it is clear how protein is processed by the body: the ingested protein is broken down, absorbed, and transported to the liver, then the liver

uses most of the amino acids, and those in excess are catabolised and eventually excreted in the form of urea (Schoeller, 1999).

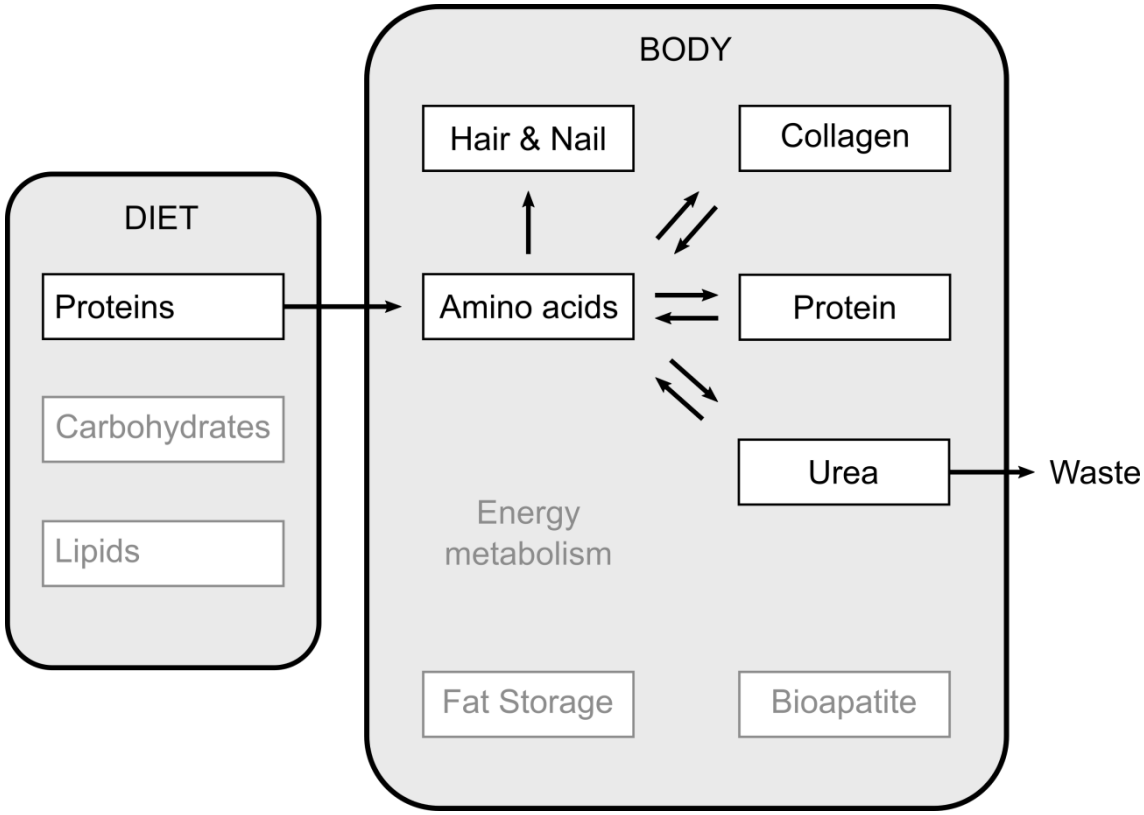


Figure 2.8 Model of the flux of nitrogen from diet to the body. Arrows indicate the direction of the flux and are likely associated with some fractionation; light grey elements indicate important body components that are not part of the nitrogen metabolism. Adapted from Hedges & van Klinken (2000).

Metabolic processing of protein causes an increase in $\delta^{15}N$ values through the trophic chain, which indicates that animals discriminate against ^{15}N. However, a permanent discrimination against ^{15}N would mean an endless ^{15}N increase in adult tissues; wherein the older the tissue, the higher its $\delta^{15}N$, but this is not observed. Instead, the trophic shift might be a result of ^{15}N accumulation during growth, when animals are in positive nitrogen balance (Poupin, Mariotti, Huneau, Hermier, & Fouillet, 2014). These authors also identify several metabolic processes likely involved in nitrogen fractionation: protein synthesis, amino acid intracellular metabolism, intestinal absorption, urea production, and urea recycling. In other words, it is not yet clear where and how in the body does this fractionation occur (O'Connell et al., 2012; Schoeller, 1999). Nevertheless, this trophic enrichment is independent of the type of nitrogen excretion (i.e. urea, ammonia or uric acid), since mammals, fish, birds,

reptiles, and insects all have similar enrichments (Caut, Angulo, & Courchamp, 2009; O'Connell et al., 2012).

In humans, the trophic enrichment ($\Delta^{15}N_{diet\text{-}body}$) is usually reported as varying between 2.5 to 3.5‰, depending on the tissue analysed (Bocherens & Drucker, 2003; McCutchan et al., 2003; O'Connell et al., 2012). However, O'Connell et al. (2012) found that $\Delta^{15}N_{diet\text{-}collagen}$ could be as high as 6‰, which suggests traditional palaeodietary studies may overestimate the level of animal protein consumed in the past. Furthermore, hair keratin is depleted in $\delta^{15}N$ by 0.86‰ in relation to collagen, which translates in a $\Delta^{15}N_{diet\text{-}hair}$ of around 5.0-5.3‰ (O'Connell et al., 2001, 2012), although lower values have been reported (Hedges et al., 2009), while nail is about 0.6‰ more enriched in ^{15}N than hair (Hülsemann et al., 2015; O'Connell et al., 2001). These divergences in $\delta^{15}N$ cannot be fully explained by differences in amino acid composition, as in carbon, and hence it is possible that some metabolic process is involved (O'Connell et al., 2001).

This trophic enrichment in ^{15}N allows researchers to detect different aspects of the food web, such as meat and fish consumption, and weaning. Ocean dwellers are usually part of longer food chains, and so have higher $\delta^{15}N$ values, which are then passed on to the food chain. An example is found in Yup'ik Eskimos, whose diet has a high marine content (Nash et al., 2009). Weaning detection works on the same idea; wherein a breastfed infant is effectively one trophic level higher than the mother, while a weaned one is not (Tsutaya & Yoneda, 2015). However, high $\delta^{15}N$ values do not necessarily indicate longer food chains. For instance, in hot and arid environments, animals often have high $\delta^{15}N$ values because the plants at the base of the food chain have high $\delta^{15}N$ values because of low water availability (Hartman, 2011). One must not forget that variation at the base of the food chain is always passed on up the food chain.

Trophic enrichment is also affected by other environmental and physiological factors. A complete review can be found in Caut et al. (2009). As a rule, most of these factors work by interfering with the nitrogen balance (i.e. by altering the rate of nitrogen excretion relative to nitrogen intake) and encompass diet quality, growth, and starvation. Usually, higher $\delta^{15}N$ values reflect high protein intake (McCue & Pollock,

2008; O'Connell & Hedges, 1999a). However, one must also consider that trophic fractionation is low on high protein intake because excreted dietary nitrogen is high, resulting in a lowering of $\delta^{15}N$ values for these diets (Robbins, Felicetti, & Sponheimer, 2005). Next, growth is linked with low $\delta^{15}N$ values, possibly due to an increased utilization of nitrogen (Fuller et al., 2004b). This contradicts the idea expressed by Poupin et al. (2014), that trophic enrichment is the result of ^{15}N accumulation during growth. In contrast, during nutritional stress, $\delta^{15}N$ values increase, possibly because the body resorts to catabolising its own tissues, which are already enriched in ^{15}N in relation to diet (Fuller et al., 2004a; McCue & Pollock, 2008; Reitsema, 2013).

In addition to these confounding variables, there are other limitations to the interpretation of nitrogen isotopic values. First, they cannot distinguish between different types of animal protein, such as meat, eggs, or dairy (O'Connell & Hedges, 1999a). Second, it is still impossible to know in what quantity a staple needs to be consumed in order to alter the consumer isotopic values – for example, fish intakes of up to 20% may not be apparent in $\delta^{15}N$ values (Hedges & Reynard, 2007).

It becomes clear through these last sections that carbon and nitrogen metabolism is far from simple, and that in samping the easily available body pools for isotopic analyses, one creates the need for additional steps before translating any findings to the archaeological record.

2.6 SUMMARY OF CHAPTER 2

To recapitulate, diet is a key component to understand population ecology, ultimately acting as an evolutionary driver, and it is particularly crucial in an East African context. In this very diverse environment, diet could explain how populations structure themselves in relation to habitats and how they respond to environmental change – through, for instance, population expansion followed by contraction into refugia. Also, there are many methods to assess diet in the past, but dental microwear texture and stable isotope ratio analyses are the most compelling because they offer a

quantitative approach that can be applied across a wide range of contexts. However, it is still unclear how the results of these methods vary between modern populations with known variable diets (O'Connell & Hedges, 1999a; Sillen et al., 1989). Therefore, this thesis aims to build a model based on data from tropical African populations that may eventually be used as a framework for interpreting patterns in archaeological and fossil data.

CHAPTER 3

STUDY POPULATIONS

This chapter encompasses relevant ethnographic information about the studied populations. Therefore, it includes the reasoning behind the choice of populations studied; descriptions of the ethnic groups, with a focus on the communities studied; a characterization of the local environment; an account of their history and origins; a few notes on social structure; and information on the subsistence strategies. However, this chapter focuses on information relevant to the diet of these populations, and does not aim to be a comprehensive review of the ethnography of the groups included in the thesis.

3.1. STUDY POPULATIONS CHOICE AND LOCATION

As mentioned in the Introduction, the groups sampled as part of this work were six: the El Molo, fishers on the south-eastern margin of Lake Turkana (Kenya); the Turkana, pastoralists on the south-western margin of Lake Turkana (Kenya); the Luhya, agriculturalists in Webuye (Kenya), but also fishers in Lake Victoria; the Luo, fishers on the northern margin of Lake Victoria (Kenya); and the Baka, hunter gatherers of the south-eastern rainforests of Cameroon. Therefore, these six communities belonged to five ethnic groups with four subsistence strategies in different geographical locations. The location of fieldwork sites in Kenya is shown on Figure 3.1, whereas the fieldwork site in Cameroon is shown in Figure 3.2.

The targeting of tropical African populations with traditional diets emerged from a number of reasons. First and foremost, tropical African populations were chosen because this work is part of the IN-AFRICA Project, which looks at changes in prehistoric African populations in Kenya. Then, the absence of any substantial body of referential work on the diversity of isotopic signatures and tooth wear in East African people led to the choice of tribes (Hülsemann et al., 2015; Romero et al., 2013).Within this choice, one must consider that globalization has led to a convergence into carbohydrate-rich diets across the world, albeit not necessarily to a convergence of isotopic values (Valenzuela, Chesson, Bowen, Cerling, & Ehleringer, 2012). This dietary globalization is more prevalent in the more well-connected urban areas, where, for the same reason, access to isotopic samples by researchers is also easier and, hence, overrepresented in datasets. However, considering that most key events in human evolution took place in Africa, and that the populations that lived through these events did not follow carbohydrate-rich Western diets, it makes sense to frame such events in data collected from African populations with more traditional diets.

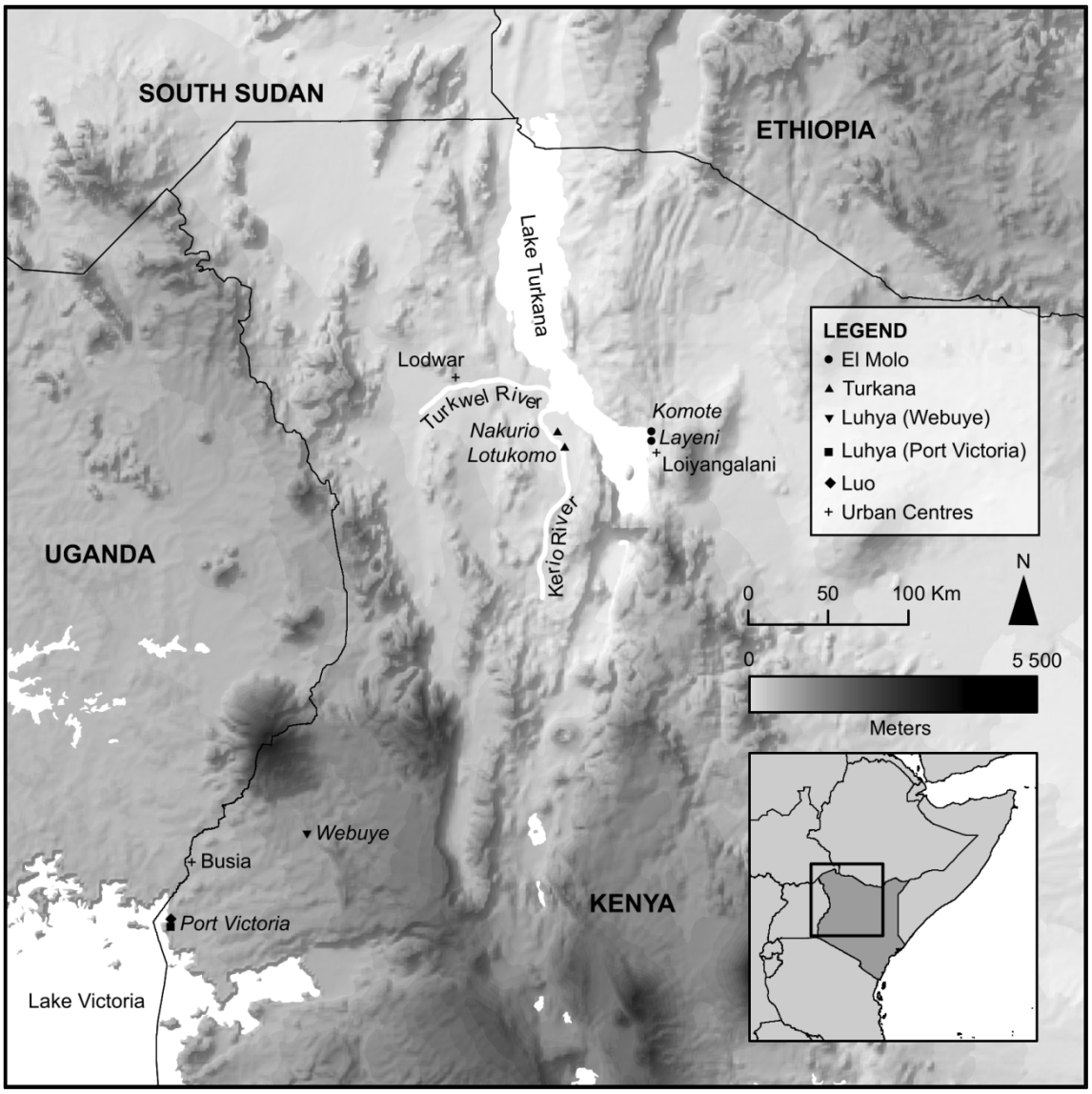


Figure 3.1 Map illustrating the location of fieldwork sites in Kenya, as well as other locations mentioned in text. The names of the fieldwork sites are in *italic*. Samples from the El Molo tribe were collected in Komote and Layeni in the south-eastern margin of Lake Turkana (Marsabit County); samples from the Turkana tribe were collected in Nakurio and Lotukomo on the south-western margin of Lake Turkana (Turkana County); samples from the Luhya tribe were collected in Webuye (Bungoma County) and in Port Bunyala (Busia County); samples from the Luo tribe were collected in Port Bunyala.

In addition, an ideal sample would represent a high diversity of subsistence patterns, namely agriculturalists, fishers, pastoralists, and hunter gatherers. Appropriate samples were sought in Kenya, as this is the country where the IN-AFRICA project is based – for this reason, all, but the hunter gatherer population, were from Kenya. Unfortunately, food producing communities are absorbing the few hunter gatherer communities that remain in Kenya, through the marginalisation of the latter and consequent loss of land (Ehret, 1998; Tishkoff et al., 2009). Therefore, Dr. Ramírez-

Rozzi, from Centre National de la Recherche Scientifique (CNRS), collected hair and breath samples from the hunter gatherer Baka of Cameroon.

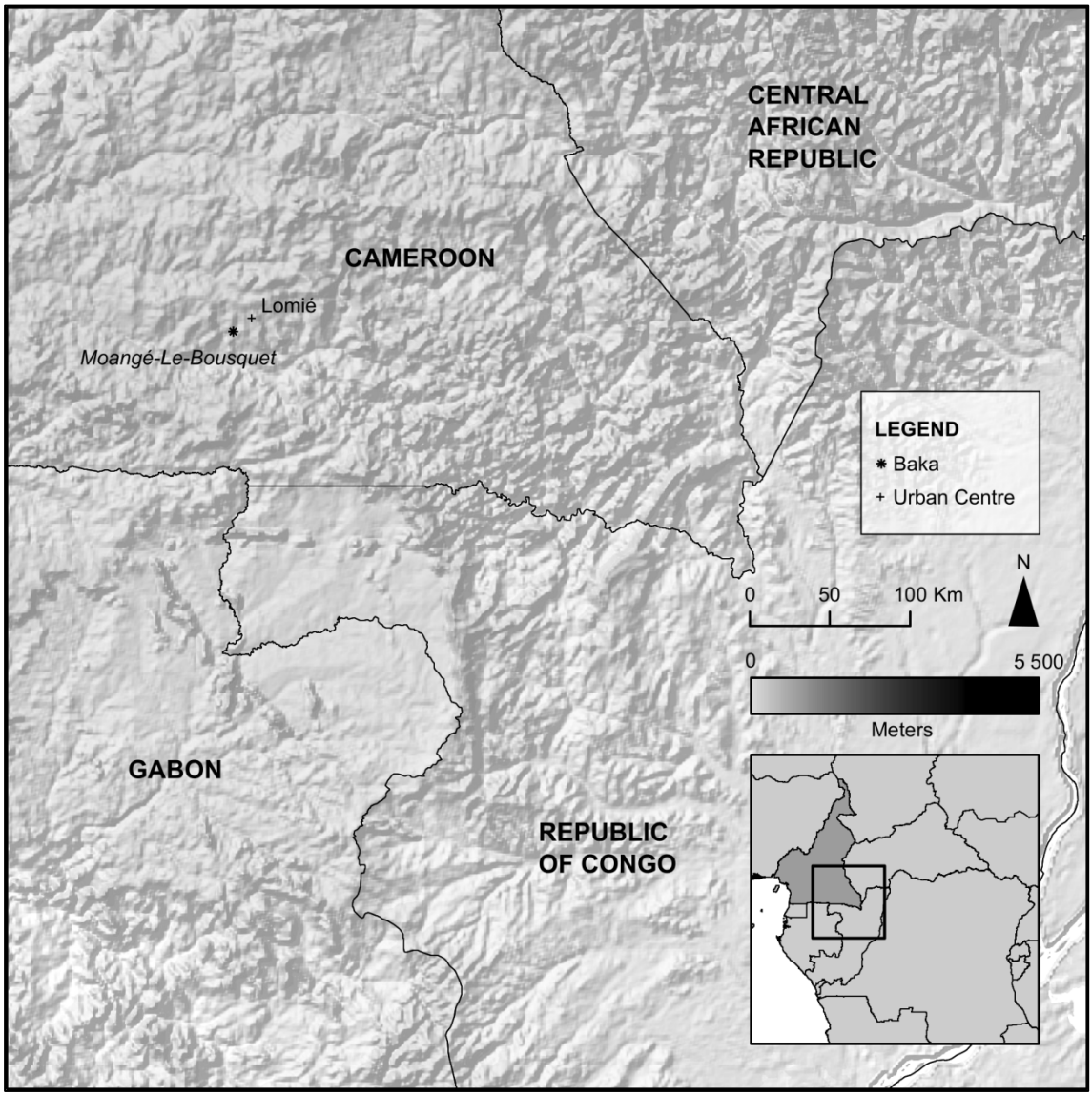


Figure 3.2 Map illustrating the location of Dr. Ramírez-Rozzi fieldwork site in Cameroon, as well as Lomié, the urban centre mentioned in text. The name of the fieldwork site is in *italic*. Samples from the Baka tribe were collected in village Moangé-Le- Bosquet in the Dja Biosphere Reserve (Lomié District, southeastern Cameroon).

Finally, fishers are overrepresented in this sample; it includes three sampled populations of fishers and only one of each of the other subsistence patterns. The reasons behind this are fourfold. First, it is likely that the human remains found so far within the IN-AFRICA project belonged to fisher forager populations (Mirazón Lahr et al., 2016; Mirazón Lahr, Foley, & Mbua, 2009; Mirazón Lahr & Foley, 2014), thus the study of present fishing population diets might provide a glimpse into these past

fishing population diets. Second, more investigation is needed on the isotopic values of fish, particularly fresh water ones, since these vary more than the isotopic values of terrestrial foods (Chapter 2, p. 39). Third, the previous point is concomitant with a lack of research into the role of freshwater resources in human evolution and diversification (Erlandson, 2001). And fourth, the choice of different fisher populations makes it possible to test varied influences on diet, such as ethnic affiliation and location.

Having established the principles that guided population choice, the following sections address each ethnic group in turn, and discuss all information that might be relevant to this study.

3.2. EL MOLO

The El Molo (El-Molo, or Elmolo) are a small fisher-forager community living in Elmolo Bay on the south-eastern shore of Lake Turkana (Marsabit County, Eastern Province, Kenya) (Figure 3.1). They are a very small community of circa six hundred people distributed across two settlements five kilometres (km) apart, Layeni (02°49'36''N, 36°41'50''E) and Komote (02°51'19''N, 36°41'33''E) (Jeffery, 2018; Kiura, 2005). Data collection took place on both villages. Interestingly, the term 'El Molo' is a derogatory one used by surrounding pastoral communities to denote the low status and poverty of any Lake Turkana fishing communities, particularly in reference to their lack of livestock. The inhabitants of Elmolo Bay call themselves Gurapau, or 'people of the lake' (Heine, 1980; Jeffery, 2018; Scherrer, 1978). However, this denomination is not used here, as the literature seldom does so.

3.2.1. ECOLOGY

This area is characterised by arid and semi-arid climates. Loiyangalani, the closest town about 10 km south, sits at 376 metres (m) above sea level, and registers an average annual temperature of 28 degrees Celsius (°C) and less than 200 millimetres (mm) annual rainfall. This rainfall is seasonal with the rainy season lasting from

March-April until the end of May while the driest season spans from October until December. In addition, due to its strong alkalinity, the water of Lake Turkana is inadequate for drinking or irrigation. This hampers agriculture and indeed, most of the other inhabitants in the area are pastoralists, a practice that further advances desertification due to the continuous overgrazing (Vallerini & Ceccanti, 2011).

3.2.2. HISTORY

Equally important is the complex and obscure history of the contemporary El Molo, which can be put together through turn of the twentieth century explorers' accounts in the region, oral histories collected in the 1970's, and linguistic and genetic analyses. In brief, early explorers' accounts of the El Molo probably encompass different communities, since this term can refer to any fishing communities that have become impoverished through loss of livestock. Descriptions of one of these, however, closely match the contemporary El Molo. Thus, it is likely that the fisher-forager culture of this community has deep roots in time, but also that they have accommodated impoverished pastoralists from neighbouring communities with consequent cultural and biological exchange (Jeffery, 2018; Scherrer, 1978; von Höhnel, 1894). The El Molo's deeper origins are more difficult to disentangle due to the lack of written records. Nevertheless, oral history, linguistics and preliminary genetic analyses indicate that the El Molo, who presently speak Samburu, previously spoke a Cushitic language (part of the Western Omo-Tana branch of the Lowland East Cushitic languages), a member of the Afro-Asiatic language family (Boattini et al., 2013; Heine, 1980; Jeffery, 2018; Tishkoff et al., 2009). The Western Omo-Tana language group probably expanded from Ethiopia around 3 Ka and occupied this area until the early nineteenth century when they contracted back, possibly pushed by Samburu speakers (Hayward, 1984; Sobania, 1978). However, both linguistic and genetic studies point to a complex history of cultural and biological assimilation of groups into local fishing communities such as the El Molo (Hayward, 1984; Miyawaki, 2008; Sobania, 1978; Tishkoff et al., 2009). A full account of this fascinating tale is found elsewhere (Jeffery, 2018).

3.2.3. CULTURE

Following the history of the El Molo, one must discuss the cultural structure of this population, particularly those aspects that affect their subsistence. This community is organised into seven patrilineal clans: Marle (or Oriarpurla), Origalgite, Origalgito, Orisayo, Orikara, Orisole, and Origaya. Each clan has political and social importance; an example of the first is the role of the clan headmen as mediators in conflicts within the community, while an example of the latter is the prohibition of marriage within the same clan. Furthermore, each clan is linked to a totemic animal and has particular food taboos, responsibilities, and supernatural powers. The food taboos are of particular interest to this project: the Marle do not eat crocodile, turtle, Nile perch abdomen, and certain birds; the Origalgite do not eat camels, or the tail end of mud fish; the Origalgito do not eat donkey; the Orisayo do not drink milk or eat fish eyes; the Orikara do not eat tiger fish, catfish, or hippo calves; the Orisole do not eat tiger fish; and, finally, the Origaya have no registered food taboos (Heine, 1980; Jeffery, 2018; Scherrer, 1978). Further to the clan structure, El Molo men also separate into Korchir and Marsara, where membership alternates with each male born into a family – i.e. a Korchir man's first son is Marsara, his second son Korchir, and so on. The single purpose of this division seems to be the formation of hippo hunting groups. Even though no hippo hunts have taken place since the 1980's, mostly due to the disappearance of the species in the area, a lot of the collective imagination of the El Molo still lies with this practice (Jeffery, 2018; Scherrer, 1978). Finally, food sharing is paramount in this community; fishermen have a social obligation to share their catch, particularly with those unable to provide for themselves, such as the elderly, widows, or single mothers (Kiura, 2005; Scherrer, 1978). This practice strengthens community ties, while avoiding the worse effects of starvation.

3.2.4. SUBSISTENCE

Turning to the specificities of the El Molo diet, fish is their main staple and they consume it fresh every day, either cooked in a stew or fried in oil. Surplus fish is sun-dried and sold in markets for profit. The El Molo themselves only consume this dried

fish when fishing returns are low. The most commonly caught types of fish are catfish (*Bagrus spp.*), Nile perch (*Lates nilotius*), different types of tilapia (Cichlidae family: *Oreochromis niloticus*, *Sarotherodon galilaeus*, *Tilapia zillii*), mudfish (*Clarias spp.*) and tigerfish (*Hydrocynus forskahlii*) (Kiura, 2005, p. 259; Scherrer, 1978, p. 39). Interestingly, Kiura (2005, p. 181) states that Nile perch is usually processed for sale and not eaten. Next, and as mentioned above, the El Molo no longer consume hippo (*Hippopotamus amphibius*), despite its central role to the culture. Crocodile (*Crocodylus niloticus*) also used to be part of their diet, but nowadays the El Molo are not allowed to hunt them, unless one becomes particularly aggressive towards people or property (tribal informant, personal communication). Meanwhile, turtles (*Trionyx triunguis*), like turtle eggs and crocodile eggs, are considerably easier to catch and are consumed opportunistically, although they make up a small part of the total diet (Kiura, 2005; Scherrer, 1978). In addition, the El Molo have also begun to adopt some aspects of the Samburu culture (see Figure 3.3), namely by keeping small livestock, like goats, and sheep. These are few in number and are kept outside the villages due to the lack of pasture. Besides, the El Molo argue they only keep livestock as an insurance against adversity or to be used as currency (Kiura, 2005, p. 97). Conversely, birds, such as pelicans, Egyptian geese, flamingos, cormorants and the domesticated chicken, are readily accepted as food. Nevertheless, some clans will not eat birds (see the taboos mentioned above), and wild birds are not commonly caught, while chickens, and chicken eggs, are reported as being eaten by children only (Kiura, 2005, p. 93). Lastly, other foodstuffs included in their diet are those acquired through commerce or relief food programmes; these include yellow maize, maize and wheat flour, beans, rice, cooking oil, tea, and sugar (Kiura, 2005, p. 90; tribal informant, personal communication).

Most El Molo suffer the consequences of fluoride contamination. Up until the installation of a fresh water pipeline five years ago, the El Molo drank directly the alkaline and saline water of Lake Turkana (Vallerini & Ceccanti, 2011). This water has a high fluoride content due to the combined effects of volcanic rock weathering and evaporative concentration (Gaciri & Davies, 1993). In small quantities, fluoride is key in

the prevention of caries, but in large quantities, it causes serious tooth and bone deformities. Severe cases of dental fluorosis (Figure 3.3), are characterised by teeth stained yellow or brown and mechanical breakdown of the tooth surface (Aoba & Fejerskov, 2002), whereas skeletal fluorosis is characterized by bone deformation, sometimes with crippling results (Krishnamachari, 1986). Reports of such extreme cases of dental and skeletal fluorosis amongst the El Molo date from at least the 1960's and extend to the present day (Kiura, 2005, p. 82; Scherrer, 1978, p. 84). Finally, Kiura (2005) conducted isotopic studies on El Molo hair, the results of which will be considered vis-à-vis those obtained in the present work.

Figure 3.3 El Molo woman with traditional Samburu necklace and signs of enamel fluorosis on teeth (Marsabit County, Kenya). Photo courtesy of Joe Jeffery.

3.3. TURKANA

Unlike the El Molo, the Turkana are a large and dispersed community. They make up the majority of the Turkana County population, and, as of 2009, had a population size of over nine hundred thousand (2009 Kenya Census). Traditionally, the Turkana

practised nomadic pastoralism. In fact, considering that pastoralism is defined by a high protein and fat intake but low energy intake, the Turkana are an extreme example of an African pastoralist population since they have the highest proportion of animal products in their diet (Galvin, 1985; Galvin & Little, 1999). However, over the past few decades, this community has begun to settle, particularly in central Turkana County. As a consequence, they have diversified their livelihood to include some form of agriculture and wage labour, and especially products from famine relief programmes (Campbell, Leslie, Little, Brainard, & Deluca, 1999; McCabe, 1990). The Turkana people sampled for this study live in two villages located close to IN-AFRICA's fieldwork sites, the villages of Nakurio (02°52'23''N, 36°08'30''E) and Lotukomo (02°46'57''N, 36°12'17''E) in the Nakurio Sub-division of Kerio Division, Turkana Central District (Turkana County, Rift Valley Province, Kenya) (Figure 3.1).

3.3.1. ECOLOGY

Turkana County is characterized by a hot and dry climate and an arid and semi-arid desert vegetation. For the most part, conditions become wetter and greener as one moves towards the escarpments in the west and away from the arid plains in the east. 70-80 km northwest of the study sites, Lodwar is the closest urban centre; it sits 496 m above sea level, and it registers an average annual temperature of 29°C and a mean annual rainfall around 200 mm (Climate-data.org, 2017b). Although some of this rain may fall between October and December, the majority occurs from March until July (Opiyo, Nyangito, Wasonga, & Omondi, 2014). Also important is the Kerio River that cuts through this region and that, despite its intermittency, supports a belt of riverine forest, an important feature for the maintenance of small gardens in the area (Olang, 1984).

3.3.2. HISTORY

Following the description of the Turkana surroundings, comes the question of how and when did they arrive here. Despite the lack of written records for the Lake Turkana region until the late nineteenth century, in-depth linguistic and genetic

analysis revealed the origins of the Turkana in a higher detail than that of the El Molo. Thus, the Turkana are part of the 'Karamajong Cluster', a group of populations that are culturally similar, occupy adjacent territories in East Africa, and speak closely related Teso-Turkana or Ateker languages. These languages belong to the Eastern Nilotic language group, in turn part of the Nilo-Saharan language family (as opposed to the Afro-Asiatic family to which the El Molo language belongs) (Boattini et al., 2013; Dyson-Hudson, 1999; Tishkoff et al., 2009; Vossen, 1982). Linguistic reconstructions propose that a proto-Eastern Nilotic language group established itself in the east banks of the Nile by 3 Ka, in modern day northern Uganda and southern South Sudan. This group then split into east and south branches, with the eastern branch occupying northern Turkana County and giving rise to the contemporary Lotuko of South Sudan, and the Maasai, Samburu, and Camus of Kenya, while the southern route occupied north eastern Uganda and formed the proto Teso-Turkana (Ehret, 2002; von Vossen, 1982). The latter then continued to split and to expand into what is their present territory. Around three hundred years ago, one of these splits gave origin to the modern day Turkana through the isolation of a Jie population, another member of the Karamajong Cluster that nowadays still lives in northeast Uganda (Dyson-Hudson, 1999; Vossen, 1982). The Turkana then spread in two waves into what is today the Turkana County of Kenya; the first in the eighteenth century took them from the Ugandan escarpment into the plains of the Rift Valley, whereas the second wave in the mid nineteenth century took them partially north but mostly to the south and south east, to the Turkwell and Kerio rivers and beyond. Throughout this expansion process, the Turkana overcame and assimilated many of their neighbouring populations (Boattini et al., 2013; Dyson-Hudson, 1999; Lamphear, 1988). The reasons behind the Turkana's success over their predecessors are still unclear. A possibility lies with Turkana drought-resistant zebu cattle, which adjusted better to the climate than the pre-existent Sanga cattle, but also in the choice of Turkana people to herd cattle in the mountains, where it was left untouched by the epizootics that occasionally sweep the region (Dyson-Hudson, 1999; Lamphear, 1988). Also important, pastoralism in this

region significantly predates the arrival of the Turkana, with the earliest evidence dating back to 4.5 Ka (Lane, 2013).

3.3.3. CULTURE

As became clear from the previous sections, the Turkana region is an unforgiving environment. These harsh conditions lead to a social organization centred on resource exploitation. Accordingly, the simplest grouping level is the *awi*, which consists of a man, his wives, their dependent children, and their livestock. Indeed, an *awi* constitutes not only a residential group but also a livestock management unit. Due to patchy resource distribution, the *awi* can break up into smaller units, although the head of the *awi* will still make decisions regarding movements and livestock. The next level of social organization is the *adakar*, which refers to a 'neighbourhood' or geographically close group of camps. Ties of kinship or friendship maintain these familial camps and members share water sources and grazing fields. The *adakar* plays an important role in information sharing, social interaction, and security by numbers. However, since each *awi* moves independently, the *adakar* spontaneously forms and dissolves, with no formal membership maintained. Nonetheless, an *adakar* usually develops from families with overlapping home ranges or *ere*, an area inherited by the *awi*'s head from his father where he has rights over pastures and water sources. In addition, the Turkana divide themselves into 19 subpopulations or *ekitela*. These sections are territorial and an individual's membership is determined at birth and not by residence. For instance, the people sampled in this study belonged to the Ngisonyoka section, which spans the area between the Turkwell and Kerio River. Other social divisions within the Turkana bind members across territories. Individuals may call upon these affiliations if they find themselves forced to move beyond the usual home range, an event that used to be common when pastoralism nomadism predominated. Examples are the *emachar*, a patrilineal descent group but also the brand used to mark the animals; initiation ceremonies that bond men of the same age set; and, finally, moieties that alternate by generation along the patrilineal line. On the whole, Turkana social organization is very

flexible at all levels, adjusting to the demands of water and food of both humans and herds (Dyson-Hudson & Dyson-Hudson, 1999).

3.3.4. SUBSISTENCE

The flexibility of Turkana social organization is a direct consequence of subsistence demands, and thus, Turkana subsistence is itself highly flexible and variable. For instance, the Turkana keep the most species diverse herds of all East African pastoralists, which include cattle, goats, sheep, donkeys, and camels (Brainard, 1991). Furthermore, the exact composition of the herd depends on the specificities of the environment. On a continuum from high to low water requirements, cattle sit at the first end, followed by sheep, then goats, and finally camels. Thus, communities with cattle migrate east towards the dry plains during the rainy season, reserving the grassy mountains for the dry season, whereas communities with mostly camels rely on the plains and avoid moving long distances at a time, although they will migrate continuously in pursuit of pasture availability (Brainard, 1991). Accordingly, water and vegetation availability are the key factors dictating Turkana movements, albeit other factors also play a role, such as predation and disease risk, as well as raiding likelihood (Brainard, 1991). Taking the above into account, it follows that the Turkana diet will vary according to season, location, and herd composition.

The most detailed study of the Turkana diet was carried among the Ngisonyoka between 1981 and 1991 (Galvin, 1985; Galvin & Little, 1999). This study found that secondary animal products, like milk and blood, constitute the majority of a Turkana traditional diet. In fact, they will only eat meat during ceremonies or when an animal dies of natural causes. In more detail, during the wet season, milk makes up more than 90% of consumed calories while in the dry season this number drops to around 30%. During this time, people consume more blood, diversify food staples, or simply eat less. For this reason, the Turkana's traditional food intake sits considerably below the energy requirements determined by the World Health Organization (WHO) (Brainard, 1991; Galvin, 1985; Galvin & Little, 1999).

Subsequently, one must address the food diversification that supplements the Turkana diet. This practice is particularly important during the dry season, but becomes a crucial source of nutrients as communities settle. This sedentarisation is more prevalent in central Turkana, and results from a process of transition into the cash economy of modern Kenya, and loss of livestock, which, in turn, results from drought, livestock disease, raiding, poor management of herds, or simple bad luck (Campbell et al., 1999). In general, cereals, like sorghum and maize, are the main staples added to the diet. In Galvin's (1985) study, cereals contributed about 24% of calories during the dry season, but none during the wet season. These, as well as other cultivated goods, like beans, sugar, and tea, are mostly obtained through purchase, barter, gifts, or government aid (Galvin & Little, 1999). Nevertheless, the Turkana also engage in 'opportunistic' sorghum cultivation along the shores of the Kerio River, a practice that probably predates European colonization (Campbell et al., 1999; von Höhnel, 1894). On the contrary, the consumption of fish works only as a backup resource wherein communities along the lakeshore (like the El Molo) provide refuge in times of shortage. Finally, wild plants and hunted game are consumed opportunistically. Of the two, wild plant consumption is more prevalent, and the most commonly eaten are *engol*, the doum palm fruit (*Hyphaene ventricosa*); *edome* fruit (*Cordia sinensis*); *eedung* fruit (*Boscia coriacea*); and *edapal* fruit (*Dobra glabra*) (Brainard, 1991; Galvin & Little, 1999). In brief, the Turkana mostly rely on domestic animal products but diversify the diet upon necessity.

Figure 3.4 Turkana woman in Lotukomo (Turkana County, Kenya). Photo courtesy of Joe Jeffery.

3.4. LUHYA

The Luhya (Luyia, Abaluyia, Baluyia, Luluyha, Luhia, Kiluyia, or Wakavirondo) are a predominantly agriculturalist community that numbers over five million people (2009 Kenya Census), making it the second largest ethnic group in Kenya, after the Kikuyu (Ehret, 1998; Ford, 1955, p. 13; MacArthur, 2013). This is in sharp contrast to the six hundred El Molo, or even to the nine hundred thousand Turkana. The Luhya live mainly in the Western Province of Kenya, north of Lake Victoria, although some live in eastern Uganda (Kanyoro, 1983; Wagner, 1949; Were, 1967). Notably, despite the fact that the Western Province comprises four different Counties (Kakamega, Vihiga, Bungoma, and Busia), its total area (~7 400 km^2) is significantly smaller to that of Marsabit (~67 000 km^2) or Turkana Counties (~71 500 km^2). If one balances this information against the differences in population numbers, it follows that the Luhya live in an area far more densely populated than the El Molo or the Turkana. An important point accompanies this simple observation; the Luhya have better connections to the 'outer world' than either of the previous two populations. They

have better infrastructure, and overall access to health, education, and commerce (*First Bungoma County Integrated Development Plan 2013-2017*, 2013; Wagner, 1949). This has important consequences regarding their subsistence.

The Luhya people sampled for this study resided in Webuye (0°35'57''N, 34°46'47''E), Bungoma County, and Port Bunyala (0°5'47''N, 33°58'42''E), Busia County (Figure 3.1, p. 55).

3.4.1. ECOLOGY

For clarity, this part addresses only Webuye in Bungoma County, while the following Luo section considers Port Bunyala in Busia County. Thus, Bungoma County sits within Lake Victoria Basin and has a humid climate in the highlands with sub-humid conditions in the rest of the County (see Figure 3.5) (*First Bungoma County Integrated Development Plan 2013-2017*, 2013). Webuye is 1532 m above sea level, and the mean annual temperature is around 20°C while the mean annual rainfall is over 1600 mm. Although precipitation falls continuously throughout the year, the month with most rainfall is May (Climate-data.org, 2017d). These conditions make Bungoma County ideal for agriculture (*First Bungoma County Integrated Development Plan 2013-2017*, 2013).

3.4.2. HISTORY

The Luhya are an intriguing ethnic group because "they share no myth of a founding father, no single narrative of historical descent and migration, no standard set of cultural practices, no common language" (MacArthur, 2013, p. 353). Instead, they came together under the term Luhya in order to defend their territory in the aftermath of the Kakamega gold rush of the 1930's. At the time, the different Bantu communities in the area chose to assuage their internal competition and to adopt the same denomination in order to stand against foreign miners, British bureaucrats, and encroaching European farmers (MacArthur, 2013). This, however, does not mean that the Luhya do not share any cultural ties. After all, their united front had to be based on common ground. Rather, it means that they are a "hybrid community" (Were, 1967, p. 60),

wherein different clans originate from different migration waves into the area followed by a constant flux of people between groups (MacArthur, 2013; Were, 1967).

In broad terms, the Luhya belong to the Mashariki Bantu branch of the of the Benue-Congo languages, which is part of the Niger-Congo language family (in parallel to the Afro-Asiatic El Molo and the Nilo-Saharan Turkana) (Ehret, 1998). However, some Luhya clans claim an origin in Egypt (Misri), which might result from the assimilation of Nilo-Saharan Kalenjin or Luo people (Ehret, 2002; Were, 1967). Thus, archaeological, genetic, and linguistic evidence indicate that the Mashariki originated in West Africa, arrived in the western margins of the African Great Lakes around 3 Ka, and went through a series of expansions and splits (Boattini et al., 2013; Clist, 1987; Ehret, 1998). One of these expansions into the southwest of Lake Victoria around 2.5 Ka gave rise to a pre-Luhya group. The timing of their arrival into the north of Lake Victoria, however, is controversial: Ehret (1998) claims that by 1.7 Ka the Luhya already occupied the area while Were (1967) defends a much later arrival in the seventeenth century. It is possible that they arrived early in time but only became the predominant population in the area later on (Ehret, 2002). Nonetheless, it is clear that the Bantu Luhya mixed extensively with Nilo-Saharan people, both with the Southern Nilotic Kalenjin that inhabited the area before their arrival, and with the Western Nilotic Luo that arrived later, in the sixteenth and seventeenth centuries (Castrì, Garagnani, Useli, Pettener, & Luiselli, 2008; Ehret, 2002; Tishkoff et al., 2009).

3.4.3. CULTURE

As for previous sections, social organization follows historical background. In this context, the largest social division of the Luhya is the tribe, of which there are eighteen, each with a distinct dialect. In Bungoma County, most people belong to the Bukusu tribe, while in Busia County people belong to the Khayo, Marachi, and Nyala tribes, although in Port Bunyala the Nyala tribe is more prevalent (Kanyoro, 1983; Were, 1967). Then, the tribe breaks into clans. Each clan forms a patrilineal, exogamous, and territorial unit. In other words, each clan traces their paternal line to a common

ancestor, refrains from intermarrying, and inhabits a common land (Wagner, 1949, p. 53). Interestingly, the word for clan, *oluhia*, is likely the origin of the term Luhya itself. It derives from the verb *okhuyia*, "to burn", or from the noun *oluhia*, "fire-place on a meadow", in reference to a common hearth, a common family (MacArthur, 2013, p. 358; Wagner, 1949, p. 55). Furthermore, each clan has specific cultural traits, such as myths, ritual rules, and food taboos. The latter prohibit consumption of certain plants, the meat of certain animals, or the milk of cows of a certain colour (Wagner, 1949, p. 74). Also, as in the El Molo, clan leaders' have no defined rights or duties; aside from the general precept of maintaining peace within the community (Wagner, 1949, p. 77). Instead, as for the Turkana, life centres on the family, the smallest social unit; these usually consist of a man, his wife, and their unmarried children. The family inhabits an isolated homestead built in the middle of the family gardens, and is economically self-sufficient. Outside the family unit, there is little organised cooperation in economic production and no marked occupational differentiation (Ford, 1955, p. 13; Wagner, 1949, pp. 40–41). However, Luhya's social structure and characteristics are currently going through dramatic changes due to the development and urbanization of the area, the conversion to Christianity, and the migration of people out of Luhya land in search of work (*First Bungoma County Integrated Development Plan 2013-2017*, 2013; MacArthur, 2012; Wagner, 1949).

3.4.4. SUBSISTENCE

In contrast to the El Molo and the Turkana, there has been no in-depth study of Luhya's subsistence. Nonetheless, due to the recent development of the area and the large population, one can find relatively complete reports on food security in the area (*First Bungoma County Integrated Development Plan 2013-2017*, 2013). These clearly show that the Luhya follow more Westernised diets than those of the two previous populations.

Despite this, the Luhya are traditionally agriculturalists (Ehret, 1998; Ford, 1955; Wagner, 1949) and one may review the historical Luhya diet to identify trends still

present today. Hence, the Mashariki Bantu group brought with them into East Africa a form of food production centred on yams, known as the West African planting tradition (Ehret, 1998, pp. 5, 13). Afterwards, by 2 Ka and likely under the influence of the Sudanic agricultural tradition, they diversified agricultural practices to include more seed crops, such as grains and beans (Ehret, 1998, 2002). In the fifteenth to the sixteenth centuries, the Luhya, together with the Kalenjin, developed a banana-based highland agriculture in the slopes of Mt. Elgon, north of Lake Victoria, that depended on rainfall, rather than irrigation, a pattern kept until recent times (Ehret, 2002, p. 396; Ford, 1955). In the twentieth century, records show that the Luhya grew maize, finger millet, sorghum, and bananas in the arable fields around the homestead, cultivated rice in swamp areas, and kept cattle on communal grazing grounds (Ford, 1955, p. 13; Wagner, 1949).

Finally, recent accounts of food production and consumption do not distinguish between the Luhya and other inhabitants of a region. Thus, this Luhya section covers information for Bungoma County, while the Luo section includes information for Busia County. This emphasizes the fact that not all Luhya share the same subsistence pattern. Nevertheless, most people in Bungoma County engage in some form of agricultural activity, while formal employment accounts for less than 20% of the County's labour force (*First Bungoma County Integrated Development Plan 2013-2017*, 2013). Today, main crops include maize, beans, finger millet, sweet potatoes, bananas, and other vegetables. Cash crops, those sold for profit, are sugar cane, cotton, palm oil, coffee, sunflower, and tobacco. Notably, *ugali*, a thick porridge of maize meal, is a basic staple, like in many other areas of Kenya (Wagah, Bader, Deligia, & Dop, 2005). In fact, in Webuye, people claimed that "a day without *ugali* is a bad day; one feels weak and cannot work" (tribal informant, personal communication). In addition, animal husbandry is an integrated part of land use; families usually keep cattle and chicken, but also sheep, goats, donkeys, pigs, and bees. On the contrary, the only communities engaged in subsistence fishing are those living close to water bodies (*First Bungoma County Integrated Development Plan 2013-2017*, 2013). Overall, the Luhya have a more diverse diet than the previous populations, with an emphasis in agricultural produce.

Figure 3.5 Luhya man (Joseph, my field assistant in Webuye) next to Broderick or Nabuyole Falls (Bungoma County, Kenya).

3.5. LUO

Like the Luhya, the Southern Luo (Joluo or Lwo) form a very large community. Numbering over four million, they are the third largest ethnic group in Kenya (2009 Kenya Census). They inhabit the north-eastern margins of Lake Victoria, mostly in the Counties of Nyanza Province, with a few pockets located in Busia County, Western Province. Outside Kenya, the Southern Luo are also present in northern Uganda and in Northern Tanzania (DuPré, 1968; Ogot, 1967; Were, 1967).

The Luo were traditionally pastoralists and historical texts report more emphasis on cattle keeping and fishing than on agriculture when compared to the Luhya (DuPré, 1968; Ford, 1955). Furthermore, the community sampled in this study belonged to Port Bunyala (0°5'47''N, 33°58'42''E), Busia County (Figure 3.1, p. 55), where commercial fishing is an important economic activity (Figure 3.6) (Ford, 1955). However, similarly to the Luhya and the Turkana, the lifestyle of the Luo is going through dramatic changes, due to recent development in the area (DuPré, 1968).

3.5.1. ECOLOGY

Busia, capital of the County, is the closest city about 50 km north of the field site and 1222 m above sea level. It has a tropical climate with a mean annual temperature of 22°C and mean annual rainfall over 1600 mm (Climate-data.org, 2017a). Close to the lake, however, conditions are drier, and rainfall normally varies between 700 and 1300 mm per year. There are two rainy seasons, the long rains from middle March to middle May, and the short rains from October to November (DuPré, 1968; Ford, 1955). These conditions make the area suitable for agriculture.

3.5.2. HISTORY

Contrary to the Luhya, the Luo of Kenya claim an historically clear ethnic affiliation (DuPré, 1968; MacArthur, 2013). Yet different lines of evidence regarding their origin are incongruent. Linguistically speaking, they belong to the Western Nilotic language group of the Nilo-Saharan language family, a sister branch to the Eastern Nilotic group of which the Turkana are part. However, genetically, they cluster with the Luhya in the Niger-Congo family (Boattini et al., 2013; Castrì et al., 2008; Tishkoff et al., 2009). In fact, Boattini et al. (2013) states that, in Kenya, mitochondrial DNA (mtDNA) of Nilo-Saharan and Niger-Congo populations form a homogeneous cluster, while Castrì et al. (2008) adds that the Luo are characterized by a high percentage of mtDNA haplogroups typical of Bantu-speaking groups. This evidence testifies to the complex population history of East Africa and to the genetic flux maintained between groups. Despite this, this section describes Luo origins according to their Western Nilotic

linguistic affiliation, considering that this aspect has far more important repercussions regarding their identity, culture, and subsistence today.

Based on archaeological and linguistic evidence, Ehret (2002) suggests that around 3 Ka the Western Nilotic people formed a population named Jii (not to be confused with the Jie from which the Turkana originated) that spread from the north into the Sudd region, an area of flooding plains in the west banks of the Nile river in southern South Sudan (similarly to the Eastern Nilotic people in the east banks of the Nile described in the Turkana section). This population continued to inhabit these plains for a long time, while splitting into their modern-day cultural descendants: the Dinka, the Nuer, and the Luo-speaking group. Subsequent, population movements out of this area probably only took place much later, in the fifteenth century, when the Luo started moving south towards north Uganda. In the following centuries, this population continued to expand and to move south, mostly towards the northwest of the Great Lakes region, but eventually reaching the east of Lake Victoria in the seventeenth century, where they gradually assimilated various Luhya-speaking communities (Ehret, 1998, 2002; Ogot, 1967). Notably, Luo migration was a multi-phased event, and there are up to five identified migration waves into east Lake Victoria alone, with different lineages moving independently and forming different communities (Ogot, 1967). As a result, Luo-speaking populations today occupy an area ranging from South Sudan and Ethiopia, to Uganda, Congo, Kenya, and Tanzania. They are usually divided into two groups: one in the north, formed by the Shilluk, the Anuak, the Acholi, the Langu, and the Palwo, and another formed by the Alur, the Padhola, and the Southern Luo (Ogot, 1967). This last population is the one studied here.

3.5.3. CULTURE

Regarding cultural and social structure, the present-day Southern Luo divide into twenty seven subgroups, each in turn composed of several clans and sub-clans (Ogot, 1967). The relationships among those are complex: first, a *piny* refers to both the territory and the members of a sub-group; next, a patrilineal, land holding lineage or

clan is termed *dhoot,* and the land it occupies is a *gweng*; within the *dhoot*, political, legal, economic, and ritual issues are managed either by a *ruoth*, a chief, or by *wuon dala,* a group of elders or homestead heads that represent the major segments of the lineage; finally, a *piny* and a *gweng* coincide if the patrilineal lineage also forms a subgroup. At a smaller scale, most Luo homesteads or *dala* are composed of a father, his wives, his sons and their wives, and his sons' sons. Like for previous populations, each *dala* forms an economically independent unit. But unlike previous populations, the Luo farm, herd, but also fish if close to water bodies. Thus, each *dala* is responsible for both a garden and a herd (DuPré, 1968; Ogot, 1967). In similar manner to the Luhya, this structure is currently suffering significant change. Population increase and the introduction of a cash economy forced young Luo to seek work in urban centres, outside the *dala*, while epizootics and raiding prohibitions led to a greater emphasis on agriculture, and to the decline of herds as prime wealth (DuPré, 1968).

3.5.4. SUBSISTENCE

The characteristics of Luo subsistence follow the same pattern of that of the Luhya. First, a thorough study of the Luo's diet is impractical due to the development of Busia County, and the sheer numbers of this population. Instead, one finds reports on food security in the area (Andika, Mwangangi, & Kitaka, 2011). Hence, and as mentioned above, this section includes information for Bunyala District, within Busia County, and not for the Luo in general. However, to highlight plausible differences between the Luhya and Luo living in Port Bunyala, one may review the historical Luo diet.

As members of the Nilo-Saharan family, the Southern Luo brought with them the Sudanic agro-pastoral tradition. Initially, this tradition is thought to have been characterised by the farming of sorghum and pearl millet and the keeping of cattle. Subsequently, it probably expanded the range of crops and animals to include gourds, calabashes, sesame, cowpeas, cotton, Bambara groundnuts (*Vigna subterranea*), sheep, and goats. As these people expanded, they also absorbed other populations that depended on fishing and hippo hunting and incorporated those foods into their diets

(Ehret, 1998, p. 6, 2002, p. 124). By comparison, recent descriptions of diet among the Luo report on the already mentioned increased reliance on agriculture, on an abandonment of cotton as cash crop in favour of maize and rice, and on a high reliance on fish and groundnuts in the flats fringing the lake (DuPré, 1968; Ford, 1955). These researchers also describe a reliance on market goods going back to before European arrival. These markets traded mostly on surplus agricultural products, or other commodities such as pottery, and iron working.

Turning to the food security report from the Bunyala District, the bulk of the local agriculture production consists of rice, maize, sorghum, and beans. The first is a cash crop and is mostly sold to areas outside the district, while the last three form the main staple in the region. Food availability, however, fluctuates throughout the year: from April onwards, after the long rains, fresh vegetables are easily found, with beans becoming available from June onwards; this abundance dwindles throughout the summer and is renovated if the short rains in October and November are plentiful; if they are not, people turn to root crops, like cassava and sweet potatoes. Regarding livestock, chicken and cattle are the most common, although goats and sheep are also found. Despite the presence of dairying animals among the livestock, the local breeds do not produce high milk yields. Fish, as mentioned several times, is a main staple close to the lake, although most is exported to processing plants and major markets like Nairobi and Mombasa. The most commonly caught fish are, in order, Nile perch (*Lates nilotius*), tilapia (*Oreochromis niloticus*, *Sarotherodon galilaeus*, *Haplochromis sp.*), *dagaa* (*Rastrineobola argentea*), and a few mudfish (*Clarias sp.*), and lungfish (*Protopterus sp.*). Fish are usually consumed fresh but may be preserved, especially for transport, by chilling, sun drying, smoking, and deep frying. Finally, key items that households purchase include maize, cassava, sweet potatoes, beans, and sorghum. These usually come from Uganda, with imports increasing in the dry season and decreasing in the wet season. On the whole, the Luo rely heavily on agricultural produce, but also on fish in areas close to Lake Victoria, as in this study (Andika et al., 2011).

Figure 3.6 Luo man tending his net in Port Bunyala (Busia County, Kenya).

3.6. BAKA

The last population in this study, the Baka (Bangombe, Bibayak, or Babinga), are a hunter gatherer, nomadic population that lives in the Central African rainforest. To be more precise, they inhabit south eastern Cameroon, northern Republic of Congo, northern Gabon, and south western Central African Republic. In terms of population size, they fall between the El Molo and the Turkana, numbering between thirty and forty thousand (Bahuchet, 2014).

The Baka are one of several ethnic groups in Central Africa collectively known as Pygmies, in their case, part of the Western Pygmy group. The Pygmies are characterised by their short stature, semi nomadism, hunting and gathering, use of temporary camps with domed huts, and regular exchange with neighbouring farmers (Hewlett, 1996). Furthermore, genetic studies indicate that these groups all share a common ancestry, preceding the other non-Pygmy groups living in the same area (Tishkoff et al., 2009; Verdu, 2014). And yet, recent studies have highlighted the variability among Pygmy populations, namely in hunting techniques, housing type, language, and social organization (Bahuchet, 2014). This fact, added to the derogatory undertone of the term 'Pygmy', led to the proposal of a new denomination as hunter gatherers of the Congo Basin. This is still far from widely adopted, though.

The community sampled for this study by Dr. Fernando Ramírez-Rozzi was located in Moangé-Le-Bosquet (3°4'34''N, 13°31'47''E), Lomié District, Cameroon (Figure 3.2, p. 56).

3.6.1. ECOLOGY

The Baka of Cameroon live in the tropical forest, characterised by closed evergreen forest. The closest urban centre is Lomié, 10 km to the northeast and 653 m above sea level, which registers a mean annual temperature of 23°C and over 1600 mm mean annual rainfall (Climate-data.org, 2017c). The tropical forest has two rainy seasons, the first from March to June, and the second between September and November. These conditions make the area suitable for agriculture. Interestingly, all Pygmy groups today depend partially on agricultural produce. Furthermore, even though it is theoretically possible to survive exclusively on hunting and gathering in some parts of the Cameroon rainforest (Sato, 2014), there is no such hunter gatherer group today nor any historical record of it (Bailey et al., 1989).

3.6.2. HISTORY

Relevant information on the origins and history of the Baka gleans mostly from genetic studies, but also from linguistic studies. Regarding linguistics, all forest foragers adopted the languages of their current or previous farmer neighbours, which renders linguistic reconstruction of population history difficult, not to say impossible. In the case of the Baka, they speak an Ubangian language, which is one of the Adamawa-Gur languages, part of the Niger-Congo language family, of which the Luhya are also part (Bahuchet, 2014; Hewlett, 1996). Nevertheless, the Baka share 20% of their vocabulary with the contiguous Pygmy group Aka, despite the fact that each group speaks non-intelligible languages (the Aka speak a Bantu language) and that there is minimal interaction between the two groups. This shared vocabulary mostly concerns rainforest activities, such as hunting, and yam and honey gathering, which led researchers to speculate on an original forest forager language and on a common origin for the Aka and the Baka (Bahuchet, 1993; Hewlett, 1996).

By comparison, genetic studies provide a more complete account of events. As described in the literature review, the Baka, together with other modern African hunter gatherers, are probably the descendants of an early modern human population, which split between 130 and 70 Ka (Tishkoff et al., 2009; Verdu, 2014). Following this, Pygmies probably diverged from other hunter-gatherers, but it was likely before 65 Ka, whereas western and eastern Pygmy populations share a common origin at 25 Ka (Batini et al., 2011; Patin et al., 2009; Tishkoff, Gonder, et al., 2007). In addition, the expansion of food producing populations into the area is not only associated with western Pygmy fragmentation, but also with admixture patterns between the Pygmy populations and the new arrivals. This process is particularly biased towards a non-Pygmy male gene flow into the Pygmy population (Batini et al., 2011; Patin et al., 2009; Tishkoff et al., 2009; Verdu, 2014). Taken together, this evidence depicts the Baka as having a deep history in time and an ancient dependency on the forest, but also a record of complex contact with their non-Pygmy neighbours.

3.6.3. CULTURE

When compared to the previous populations, the Baka likely have the most fluid social structure of all the groups. Overall, the Baka are organised into clans with little hierarchy between them and members of different clans often live together in the same camp (Bahuchet, 2014; Joiris, 1996). At a higher scale, the regional band consists of a set of permanent camps, with each camp situated four to eight kilometres away from a farmer village. The relationship between camps within a band is established through matrimonial ties and visits are frequent, with visitors easily integrated into the social structure of the host camp. Furthermore, since families often share a ritual membership (*yeli*) linked to particular spirits, familial ties between the regional band lead to co-guardianship of the tutelary spirits of the dead. As a result, each regional band shares economic, social, and religious structures (Hewlett, 1996; Joiris, 1996).

In turn, each permanent camp consists of thirty to one hundred individuals, where the nucleus of the camp corresponds to the descendants of one lineage, while the wives and husbands of the group belong to another lineage. These lineages often bear the name of animals, plants, or objects to which dietary or behavioural prohibition is related (Hewlett, 1996; Joiris, 1996). In addition, there is no designated leader within each camp. Instead, the camp communicates, organises activities, and resolves problems through *mosambo*, a public speaking protocol that is heard twice a day (Lewis, 2014). In parallel with the absence of a leader, the Baka practise demand-sharing, resulting in an egalitarian society, similarly to the El Molo. Within this system, the Baka form food sharing clusters of 3-4 households, within the wider residential camps. In addition, most material items are communal, and need determines who can claim it, especially if they are consumable; refusal is impolite, even offensive. In contrast, knowledge, such as herbal remedies, can be exchanged, negotiated, and traded (Dyble et al., 2016; Lewis, 2014).

Two final points are important regarding Baka's social organization. First, they form patrilineal, patrilocal, and exogamous groups, characteristics that allow men of the same family to hunt together. And second, these communities maintain complex, albeit

problematic, relationships with neighbouring farmers and often trade labour, as well as meat and other forest products, for agricultural produce. Farmers and Pygmies maintain independent social functions, but cultural and (unidirectional) biological admixture is common (Bahuchet, 2014; Hewlett, 1996; Joiris, 1996).

3.6.4. SUBSISTENCE

As for the Turkana and the El Molo, researchers have conducted thorough studies of the Baka diet, despite the difficulty in following and registering Baka diet in the forest (Sato et al., 2012; Yamauchi et al., 2000). Before addressing these studies, one must pay attention to overall features of Baka subsistence. Of particular relevance, the area in south eastern Cameroon, where the Baka community of this study lives, is connected by regular public transport due to its closeness to the more populated areas of West Africa (Hewlett, 1996). In consequence, the Baka began to settle in the 1950's and to change the subsistence focus from hunting and gathering to farming (Yamauchi et al., 2000). Despite this, the Baka still spend four to five months of the year in the forest, mostly during the dry seasons (Bahuchet, 2014).

Regarding hunting and gathering activities, the Baka practise sexual division of labour: men collect honey and hunt, the latter mostly using trap lines, but also spears in the case of medium-sized game; women, on the other hand, cook, work in agricultural fields for village women, and collect labour intensive, but dependable foods, such as wild yam, edible leaves, fruits, insects, mushrooms, small fish, and crustaceans (Hewlett, 1996; Lewis, 2014). Furthermore, Baka's plant crops include cassava, plantain banana, taro, okra, and maize, with cacao introduced recently as a cash crop (Yamauchi et al., 2000).

Overall, starch and cereals constitute the majority (70-85%) of the energy intake of the Baka, while the remainder is divided between fish and meat (8-15%), and vegetables, fruits and other minor products (7-15%). The high proportion of starchy foods in Baka diet is probably the result of a dependency on staples, such as plantain banana and wild yam (Sato et al., 2012; Yamauchi et al., 2000).

Figure 3.7 Baka hunter in Moangé-Le-Bosquet (Lomié District, Cameroon). Photo courtesy of Dr. Fernando Ramírez-Rozzi.

3.7. SUMMARY OF CHAPTER 3

From the sections above, it is clear that the populations studied in this thesis establish complex interactions with the environment that considerably affect their subsistence. Furthermore, most, if not all populations are currently going through significant changes to their traditional way of life, due to the increased connectedness with the world at large. Table 3.1 aims to bring together information on these populations, so it may better contextualise the following chapters. Namely, this table includes information on population size, on the ecology of the area, on linguistic and genetic

affiliation, on the broad classification of subsistence patterns, on the presence of seasonality in that subsistence, on the existence of food taboos, and on the possibility of recent changes to diet, as a result of cultural admixture. All of these are potentially important factors in the characterization of the diet of these people.

Table 3.1
Summary table on main characteristics of study ethnic groups: ecology, linguistic and genetic affiliation, livelihood, seasonality, food taboos, and recent changes to diet.

	El Molo	Turkana	Luhya	Luo	Baka
Population size	~ 600	~ 900 000	>5 000 000	>4 000 000	>30 000
Ecology	semi-arid	semi-arid	sub-humid	sub-humid	humid
Temperature	28°C	29°C	20°C	22°C	23°C
Rainfall	<200 mm	~200 mm	>1600 mm	700-1300 mm	>1600 mm
Affiliation					
Linguistic	Afro Asiatic	Nilo Saharan	Niger Congo	Nilo Saharan	Niger Congo
Genetic	Afro Asiatic	Nilo Saharan	Niger Congo	Admixed	Khoisan pygmy
Traditional livelihood	fishers	pastoralists	agriculturalists	fishers and agriculturalists	hunter gatherers
Seasonality	no	yes	some	some	yes
Food taboos	yes	no	yes	no	yes
Recent changes to livelihood	no	yes	yes	yes	yes

CHAPTER 4

MATERIALS AND METHODS

This chapter contains all relevant information concerning data collection and processing within this thesis. With this goal, it includes an account of fieldwork, information on sampled participants and foods, and a description of the different methods used to carry out the diet questionnaires, dental microwear analysis, and stable isotope ratio analysis. Finally, this chapter addresses the details of statistical analysis and the ethical and safety implications of working in Kenya with living human subjects.

4.1. SAMPLED SUBJECTS

After the detailed discussion of population choice in the previous chapter, the present chapter addresses the criteria of subject choice and data collection. As a rule, a sample size of thirty individuals was set as the minimum for each group. This has been shown to allow the reliable estimation of population means in isotope studies, although no similar guideline exists for microwear studies (Pearson & Grove, 2013).

Fieldwork data collection among the El Molo took place in late April and early May 2015, during the wet season; among the Turkana during the month of August 2015, just after the wet season and concurrent with the IN-AFRICA archaeological fieldwork season; in Webuye, during the first week of September, and in Port Victoria during the rest of that month, with both regions going through the dry season; finally, the Baka were sampled by Dr. Ramírez-Rozzi in late October and early November of that same year, during the short wet season. Throughout data collection in Kenya, field assistants were employed as translators, but also as aides in the dental moulding process.

Regarding subject criteria, participants had to be in general good health, and not suffer from any of the diseases (e.g. diabetes) known to lead to isotopic fractionation (Reitsema, 2013). Similarly, dental moulding was only done on individuals with good dental health (e.g. few cavities) and with no tooth sensitivity, i.e. individuals who showed no discomfort when having pressured air blown into teeth. Furthermore, only lower teeth were moulded, aiming for right versus left balance, although an individual's preference regarding side was respected. Some individuals who could not have dental moulds taken still participated in the rest of the study. This study aimed at a balanced female versus male ratio amongst those sampled, and to include only adult participants. This, however, was not always possible, since adults in general and men in particular were often away from the settlements during the day. Hence, the female versus male ratio is biased towards females, driven mostly by the Baka sample, and age ranges reveal the inclusion of a few teenagers (see Table 4.1). Finally, this study included pregnant and lactating women, even though these conditions may result in isotope shifts (Fuller et al., 2004b; Reitsema, 2013). Excluding these women would

further decrease the number of eligible participants and thus, it was opted to include them but to register the information for later analysis.

If an individual fulfilled the criteria above, I explained the study to them, and invited them to participate and to sign the consent form (as described in the ethics section in this chapter). Once that procedure was complete, data collection began.

A completed example of a data collection form is provided in Appendix A, Figure A.1 to A.4. First, I assigned each person an identification number consisting of an ethnic group code, followed by a number (e.g. TK26 is the 26th Turkana participant). Then, I registered other identification data, such as clan, age, and sex, and, if female, whether pregnant or lactating, and for how long. Next, I conducted the diet questionnaire and collected hair. Initially, nail was to be collected only if participants' hair was too short for collection. However, because hair is used in witchcraft, several people were reluctant to give it, which led me to include more nail samples than initially planned (see Table 4.1). To address this, if people were willing to provide hair, I would also ask for a nail sample, in order to study differences between the two tissues. This was also difficult, since most people had very short fingernails. In at least two occasions, people provided toenails instead. In addition, most study participants had their breath collected. Breath was sampled around 11 am each morning, and when that was impossible, at least three hours after the individual's last meal. This set up was necessary because breath isotope values vary throughout the day, particularly following a meal (Hedges et al., 2009; Whigham et al., 2014). As a result, breath collection also proved challenging. People sometimes signed up as participants in the morning and had breath collected, but then never turned up for the rest of the study. Other times, people who signed up after 11 am and had all other data collected, never showed up for breath collection the next day. As a final procedure, if viable, I took moulds of participants' teeth.

These issues surrounding breath and hair sampling added to some people's inability to have moulds taken, resulted in several incomplete data sets (e.g. an individual with

diet questionnaire, nail, and dental moulds, but no breath). This led me to strive for thirty complete data sets per population.

In contrast, Dr. Ramírez-Rozzi did not collect diet questionnaires, nail, or dental moulds from the Baka. First, the researcher felt diet questionnaires were unsuitable for the Baka population since they often forget what they eat during forest trips. Second, all Baka were willing to supply hair, although Dr. Ramírez-Rozzi also reported some level of reluctance. And third, dental microwear had already been studied for this population. To address this gap in data for the Baka, I will address isotopic results in relation to published data on Baka diet and microwear (Romero et al., 2013; Yamauchi et al., 2000). Nevertheless, one must keep in mind that these results will not be fully comparable, considering that the methodologies used in the published Baka diet and microwear study differed from those employed in the present work.

Table 4.1 summarises information on demographics and data collected for all populations while Appendix A (Table A.1 to A.6) includes the same information per individual.

Table 4.1
Demographics and collected data across sampled populations.

Parameter	Populations						Total
	El Molo	Turkana	Luhya (Webuye)	Luhya (Port Vict.)	Luo (Port Vict.)	Baka	
N	31	37	32	31	29	33	193
Age (year)							
mean (SD)	25.3 (8.6)	28.2 (8.4)	35.2 (11.5)	37.4 (10.4)	31.0 (8.9)	33.2 (12.6)	31.7 (10.9)
range	[15, 62]	[15, 54]	[18, 71]	[22, 58]	[19, 54]	[13, 60]	[13, 71]
Gender							
male (%)	14 (45)	18 (49)	16 (50)	14 (45)	14 (48)	9 (27)	85 (44)
female (%)	17 (55)	19 (51)	16 (50)	17 (55)	15 (52)	24 (73)	108 (56)
pregnant	2	1	0	1	1	1	6
breastfeeding	6	9	2	4	4	9	34
Data Collected							
diet survey	31	37	32	31	29	—	160
hair	27	26	19	19	19	33	143
nail	9	20	15	18	21	—	83
breath	30	33	32	30	29	32	186
moulds	31	30	30	30	29	—	150
right (%)	19 (61)	11 (37)	16 (53)	19 (63)	14 (48)	—	79 (53)
left (%)	12 (39)	19 (63)	14 (47)	11 (37)	15 (52)	—	71 (47)

Notes. *N* = number of individuals; SD = standard deviation.

4.2. SAMPLED FOODS

In addition to the data collected from individuals, food items were also sampled in the field. Table 4.2 lists sampled foods from all study sites.

Table 4.2
Plant and fish samples from study sites. Scientific names followed by '?' indicate uncertainty in denomination.

Location (Ethnic Group)	Category	Sample	Scientific name
Layeni (El Molo)	C_4 plant	Sorghum	*Sorghum bicolor*
		Maize	*Zea mays*
	Fish	Tilapia	*Oreochromis niloticus*
		Golefish	*Hydrocynus forskahlii* ?
		Flatfish	*Barbus turkanae* ?
		Catfish	*Bagrus sp.*
		Mudfish	*Clarias sp.*
Nakurio (Turkana)	C_3 plant	Doum palm fruit	*Hyphaene ventricosa*
	C_4 plant	Sorghum	*Sorghum bicolor*
Webuye (Luhya)	C_3 plant	Beans (legume)	*Phaseolus vulgaris*
		Cow peas (legume)	*Vigna unguiculata*
	C_4 plant	Maize	*Zea mays*
Port Victoria (Luhya and Luo)	C_3 plant	Soya beans (legume)	*Glycine max*
		Beans (legume)	*Phaseolus vulgaris*
		Cassava	*Manihot esculenta*
		Rice	*Oryza sativa*
		Cow peas (legume)	*Vigna unguiculata*
		Green peas (legume)	*Pisum sativum*
	C_4 plant	Sorghum	*Sorghum bicolor*
		Finger millett	*Eleusine coracana*
		Yellow maize	*Zea mays*
	Fish	*Dagaa* (bone and whole)	*Rastrineobola argentea*
		Ofulu (bone and whole)	*Haplochromis sp.*
		Nile perch	*Lates nilotius*
		Helicopter fish	*Protopterus sp.* ?
		Esebu	*Clarias sp.* ?
		Tilapia	*Oreochromis niloticus*

Even though analysing food staples was not the main goal of this study, it is important to understand the food web that sustains an organism of interest. The latter is

particularly true for lake foods, since isotopic patterning in these environments is highly variable (Ambrose et al., 1997; Keeley & Sandquist, 1992; Raymond & Bauer, 2001). Nevertheless, it was opted to exclude animal products from analysis since these items cannot be brought from outside the European Union into the United Kingdom without previous authorization ("Bringing food, animals or plants into the UK," 2014). As a result, samples were limited to plants and fish, which, as far as possible, represented all items of this category mentioned during diet interviews. In addition, only the bones of fish were sampled, except for *Dagaa* and *Ofulu,* which were dried in the sun, and hence were easy to transport. In general, participants supplied most of the samples upon request, except for the plant samples from Port Victoria, which were purchased at the local market. Finally, for some of the fish, it was not possible to determine the scientific names with certainty, since local denominations were not found in the literature (see Table 4.2).

4.3. DIET ASSESSMENT METHODOLOGY

To characterise the diet of the studied populations, dietary histories and semi-quantitative food questionnaires were collected (Figure 4.1), based on Gibson (1990, 1993). These methods are the most commonly used to assess food intake in groups, mainly because they are less time intensive when compared to other methods, such as estimated or weighed food records, which require a trained researcher to accompany the subjects and record all consumed food for a given time. However, since the methods used here rely on the subject accuracy, they are more prone to error in assessment of food intake, particularly due to incorrect estimation of portion size (Gibson, 1990, p. 89). This is a matter of concern that will be discussed in full in later chapters.

4.3.1. FIELDWORK PROTOCOL

Figure A.3 and A.4 in Appendix A illustrate a completed diet questionnaire. The dietary history section (Figure A.3) consisted of a twenty four hour recall of actual

intake, where participants listed all food consumed the previous day, and collection of general information on the overall eating pattern of the subject, both at mealtimes and between meals (Gibson, 1993). In addition, participants were asked for details on any foods they particularly avoided or favoured in an effort to detect food taboos and preferences.

Following this, the semi-quantitative food frequency questionnaire (Figure A.4) consisted of three sections: a list of foods, a portion section, and a frequency of use section. Regarding the list of foods, this was initially designed based on ethnographic data. However, it was continuously updated while in the field, particularly by adding food items mentioned by subjects during the dietary history. Furthermore, and as far as possible, the food list remained consistent across populations. To illustrate, Turkana participants were still asked if they consumed fish or other lake products, even though they are known to rarely do so. This was aimed at reducing interviewer bias by not assuming what a given population was eating. After ascertaining if an individual ate a particular food item, a choice of three portion sizes was available, small (S), medium (M), and large (L). These were based on the United States Department of Agriculture National Nutrient Database for Standard Reference (USDA), as reported by Gibson (1993, pp. 16-17). Next, there were five categories for the frequency of food use, daily (D), weekly (W), monthly (M), yearly (Y), and rarely or never (N). Besides choosing the most appropriate category according to the subject's response, it was also recorded the number of times each food item was consumed. Furthermore, participants were questioned on non-dietary behaviour that might potentially affect results, like smoking or drinking alcohol. To conclude the interview, participants were asked if there were any other items they recalled eating or consuming, which had not been addressed up until that point. If the answer was positive, details were taken on these new items, inclusively by adding them to the food frequency questionnaire.

Figure 4.1 Conducting a diet questionnaire among the Turkana in the IN-AFRICA camp, Turkana, Kenya. Photo courtesy of Dr. Frances Rivera.

4.3.2. ANALYSIS

Due to the difficulty of transforming data from dietary histories into quantitative information, the diet of the studied populations was mostly characterised through the semi-quantitative food frequency questionnaire, with dietary histories providing mostly corroborating or supplementary information.

For the semi-quantitative food frequency questionnaire, food intake in grams per day (g/day) was calculated by multiplying the frequency by the amount of each reported food. The frequency categories had previously been converted to a daily basis; by dividing food consumed weekly by 7, monthly by 30, and yearly by 365.

In addition, the intake of energy in kilocalories per day (kcal/day) and nutrients (g/day) was calculated using food composition tables. In detail, the nutrients studied were protein, fat, and carbohydrates. Notably, not all foods reported by participants were found on a single food composition table, and thus, several tables had to be used, with the necessary attention to their equivalence. This is a common procedure, since many

of the tables themselves glean data from different analyses and reports. As a rule, the food composition table used was the International Minilist (IML), which was developed specifically for foods consumed in rural Kenya (Murphy, Gewa, Grillenberger, & Neumann, 2004; Murphy, Weinberg-Andersson, Neumann, Mulligan, & Calloway, 1991). Moreover, the two other tables used were the Food and Agriculture Organization of the United Nations Database (FAO, 2016) and the United States Department of Agriculture National Nutrient Database for Standard Reference (USDA, 2015). Finally, the nutrient contents of blood could not be found in any of the above tables, and thus this information was obtained from Galvin (1985), the only study so far, to the best of my knowledge, to have analysed blood for its potential as food. Following this, energy and nutrient content of each consumed food was calculated by multiplying the food intake (described in the above paragraph) with the energy or nutrient content in 100g of that food, and then dividing the result by 100 to obtain energy and nutrient intake per day (Equation 4.1).

$$\underset{(kcal\ or\ g/day)}{Energy\ or\ nutrient} = \left(\underset{(g/day)}{Food\ intake} \times \frac{\underset{(kcal\ or\ g)}{Energy\ or\ nutrient\ content}}{100g\ of\ food} \right) \Big/ 100 \qquad (4.1)$$

Following this, all foods were divided into groups: animal, fish, C_3 plants, C_4 plants, and other, in a similar procedure to that found in Patel et al. (2014). Accordingly, food, energy, or nutrient intake for a certain group was calculated by adding all relevant elements of that group. For instance, animal protein (g/day) was derived from the sum of protein from meat, dairy products, and eggs. It must be stressed that this categorization is imperfect. Or rather, if one considers that, in simple terms, carbon isotopes track C_3 versus C_4 plant consumption, while nitrogen isotopes track trophic position, it would be reasonable to have two different classification systems, one for each isotope. However, this is unfeasible, since it is often difficult to discern the details of the food web. For instance, it is impossible to know, within this research design, if cattle were feeding on C_3 or C_4 plants.

The matter of food composition is discussed in more detail in Chapter 5, which reports on the results of the semi-quantitative food questionnaires.

4.4. DENTAL MICROWEAR TEXTURE METHODOLOGY

In brief, lower molars were moulded using a polyvinylsiloxane silicone putty-wash system while in the field, casted using an epoxy at the Garrod Laboratory, scanned using a confocal microscope, and analysed using scale sensitive fractal analysis, both at the Ungar Laboratory, Fayetteville, United States of America. The moulding and casting process reproduces features with resolutions to a fraction of a micron and maintains the resolution for many years (Galbany, Martínez, & Pérez-Pérez, 2004; Teaford & Oyen, 1989b). The exact protocol used for dental moulding in living participants was that developed by Sarah Livengood at the Ungar Laboratory (Livengood, 2015, personal communication), while the protocol for dental casting followed standard methods (Galbany et al., 2006, 2004). Importantly, whilst dental moulding in archaeological material simply requires the use of an alcohol-soaked cotton to clean the teeth of interest, followed directly by moulding of the teeth with an appropriate material (Galbany et al., 2006, 2004), dental moulding in living subjects requires removal of bacterial film prior to moulding. This biofilm is the result of the activity of mouth bacteria, which obscures dental microwear, and that must be mechanically or chemically removed (De Nardo, Chiappe, Gómez, Romanelli, & Slots, 2012; Teaford & Oyen, 1989b). Even though some dental moulding has been conducted on living primates (Dennis, Ungar, Teaford, & Glander, 2004; Teaford & Glander, 1991; Teaford & Oyen, 1989b) and on living humans (Romero et al., 2013; Teaford & Tylenda, 1991) for the purpose of dental microwear studies, Sarah Livengood developed a method specifically targeting living people, and biofilm removal. Thus, this method was the one chosen for this study.

4.4.1. FIELDWORK PROTOCOL

Dental moulding only took place if the subject (1) agreed to the procedure, (2) had no complaints of tooth ache or sensitivity, (3) had good general dental health, and (4) reported no discomfort when having air directed on teeth for five seconds minimum.

To begin, general dental health was evaluated using a plane mouth mirror and registered on the data collection form (Appendix A, Figure A.1, and A.2). Afterwards, the air test consisted of having air from a clean, mixing tip connected to an air compressor, directed on teeth.

If the subject was cleared for participation, dental moulding began (Figure 4.2). First, participants ate one cup of popcorn on a side of their mouth of their choosing. Next, they rinsed the mouth with half a cup of water, brushed the lower teeth using a pre-pasted, unflavoured toothbrush (MediInn® Disposable Toothbrushes), for two minutes, and again rinsed the mouth with half a cup of water. Afterwards, the surface of the teeth was cleaned using an oral irrigation device (Panasonic® EW-DJ10 Travel Oral Irrigator), steadily increasing the pressure until the maximum and for about 1 to 2 minutes, i.e. until the 165 millilitres (ml) water tank was empty. Following this, an initial impression was taken by mixing Affinis® putty (Coltene Whaledent) into a dental tray, placing a polythene sheet over it as spacer, and inserting the tray into the subject's mouth. This step served only to create space for the teeth within the putty, working towards a better impression. At this point, a clean cotton roll was inserted under the tongue to absorb any excess saliva and a tongue depressor was placed over the tongue to keep it from moving. The tongue depressor was held by the research assistant from this moment on, and participants were instructed to keep their mouths open until the end of the procedure. Subsequently, teeth were dried by wiping them with a cotton roll and by having air from the air compressor directed at the teeth for at least two minutes. Then, Affinis® light body wash was directly applied unto the teeth using a dispenser gun adapted with a mixing tip and an oral tip. Next, more impression material was quickly applied to the tray, and this was inserted unto mouth and held, without applying force, for five minutes. Once fully set, the tray was

removed, and the impression checked for defects. If there were any, such as air bubbles, the procedure was repeated. Finally, moulds were disinfected using Aseptoprint® spray, and stored in a zip bag labelled with the subject's code and date.

Considering the complexity of this procedure, several safety measurements were taken. First, both the research assistant and I used vinyl gloves, which were replaced between each participant. Furthermore, mixing tips and oral tips used in the dispenser gun and in the air compressor were also discarded after each single use. Moreover, participants were advised to not swallow the water from the irrigator, lest they choke. Instead, the water dripped down their chins, and they wore a rain coat to protect their clothes. Equally important, both the air compressor and the irrigator were demonstrated on participant's fingers in advance, so they could become familiar with the feeling prior to its use. Last, but not least, participants were continuously monitored for any signs of discomfort (e.g. eyes watering, unable to hold mouth open, emotional cues from facial expressions), and repeatedly asked if they were faring well. If at any point, the subject showed uneasiness, the procedure was interrupted.

Figure 4.2 Making dental moulds in Lotukomo, Turkana, Kenya. Photos: material preparation (left); waiting mould to set (top right); finished mould (bottom right). Photos courtesy of Joe Jeffery.

4.4.2. LABORATORY PROTOCOL AND ANALYSIS

Cast preparation used Epotek® 301 resin and hardener (Epoxy Technologies, Inc.). First, moulds were removed from the dental tray, cleaned using a standard detergent, and dried. Next, a wall of Affinis® putty was built around the moulds to prevent epoxy from seeping out of the mould before setting. Following this, moulds were weighed, paired in weight ranges, and secured to plastic boxes using Blu Tack® and packing peanuts around the moulds. These measures kept weight balance and decreased movement in the centrifugal step that followed. Afterward, epoxy was mixed in a 4:1 (A+B) weight ratio, stirred using a wooden spatula for two minutes, and centrifuged for one minute at 3000 rpm, in order to remove air bubbles. Then, quickly but carefully, the epoxy was poured into moulds using a plastic Pasteur pipette until the occlusal surface was covered. At this point, the complete apparatus (plastic box, moulds, packing peanuts, and epoxy) was reweighed and any weight imbalances corrected for, using Blu Tack®. Immediately after, the plastic boxes with the moulds were centrifuged once again for three minutes at 2500 rpm, to force the epoxy into every feature of the microwear surface, and to remove any air bubbles trapped at the silicone-epoxy boundary. After this, the moulds were checked for spillage. If there was none, the casts were left to cure for 48 hours. In contrast, if there was spillage, the leak was corrected using Affinis® putty, and more epoxy added, followed by a new centrifugal step. Once a cast was fully cured, it was removed from the mould by gently loosening the sides until the cast popped out. Finally, the casts were put under a light microscope to check for any casting errors, such as remaining air bubbles. If any were found, the mould was re-casted, considering that the accuracy of microwear features is maintained until the fourth consecutive replica from the original mould (Galbany et al., 2006).

As the final step, and as mentioned at the beginning of this section, the casts were scanned and analysed at the Ungar Laboratory. Each individual was represented by one lower molar (first, second or third) with the best microwear preservation, on which a crushing facet (i.e. a Phase II facet: 9, 10n, or x) was scanned using a Sensofar® Plμ Neox white-light scanning confocal microscope fitted with a 100x objective. This yielded a cloud of points for each surface, with a lateral sampling interval of 0.17

micrometre (μm), a vertical resolution of 0.005 μm, and a field of view of 138×102 μm. Data were collected from four adjoining fields, which were automatically stitched together to form a common data cloud, with an area around 242×181 μm (Figure 4.3). The data for each scanned individual were then imported into the SolarMap® software, where they were levelled, and cleaned of any small identifiable defects, such as dust particles. The resulting *.sur files were then analysed with the Sfrax® and Toothfrax® software, using scale sensitive fractal analysis. Following, five variables were calculated: surface complexity (*Asfc*), scale of maximum complexity (*Smc*), anisotropy (*epLsar*), heterogeneity (*HAsfc*), and texture fill volume (*Tfv*) (see Chapter 2, Table 2.1, p. 32).

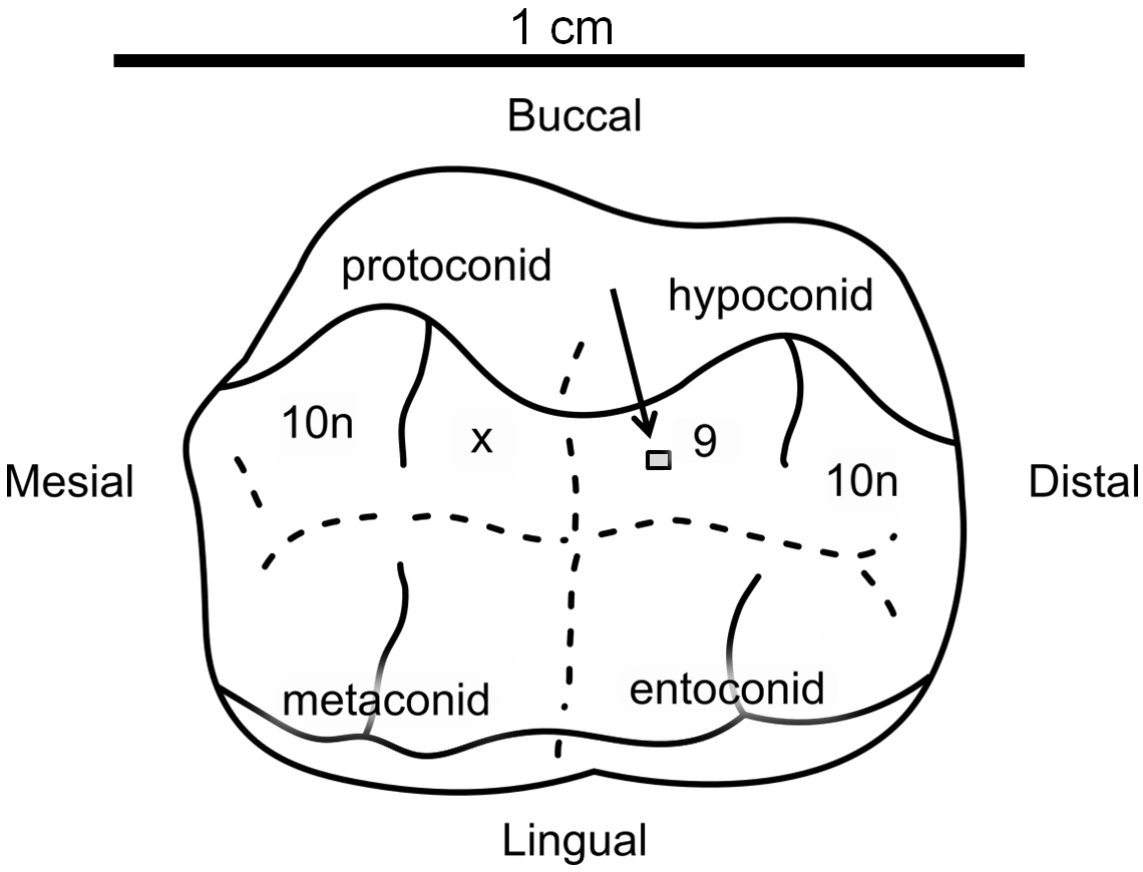


Figure 4.3 Occlusal view of a *Macaca mulatta* right lower M_2, indicating the position of Phase II facets (9, 10n and x) and the size (to scale) of the area scanned by the confocal microscope. Adapted from Kay (1977) with scale based on Trotter et al. (1977).

4.5. HAIR AND NAIL STABLE ISOTOPE METHODOLOGY

The protocols used for hair and nail collection and analysis were those used in the Garrod Laboratory at the McDonald Institute of Archaeology, University of Cambridge, and based on O'Connell & Hedges (1999a) and O'Connell et al. (2001).

4.5.1. FIELDWORK PROTOCOL

A hair sample of 20-30 hairs of at least 2 cm length was cut at the nape of the head, as close to the scalp as possible, and then stored in a plastic bag labelled with the subject's code and the date of sampling (Figure 4.4) (O'Connell & Hedges, 1999a). Unlike in previous studies, the sample was not secured in micropore tape, since the very curly hair tended to stick to the tape in its entirety and not only at the cut end. This was not an issue since this research did not focus on studying diet variation across time, an avenue of enquiry also rendered impractical by the often short hair. Meanwhile, nail samples were collected by simply clipping the distal edge of a fingernail and storing it in a plastic bag, also appropriately labelled (O'Connell et al., 2001).

Figure 4.4 Collecting hair among the El Molo in Layeni, Marsabit, Kenya. Photo courtesy of Joe Jeffery.

4.5.2. *LABORATORY PROTOCOL AND ANALYSIS*

Hair and nail are made of virtually pure keratin and thus minimal sample preparation is required prior to isotopic analysis. Nevertheless, since both are exposed to the environment, they must be cleaned of contaminants (O'Connell & Hedges, 1999a; O'Connell et al., 2001).

Each hair sample was cleaned by two successive immersions in a 2:1 mixture of methanol and chloroform (MeOH:$CHCl_3$) for 30 minutes each, followed by two 30 minutes rinses in distilled water. The solvent was then discarded and hair samples were freeze-dried. Finally, 0.8 ± 0.1 milligrams (mg) of sample were weighed into tin capsules for analysis.

Each nail sample was cleaned by sandblasting, removing the surface layer of nail together with any surface contamination. Next, samples were rinsed in distilled water, and left for 30 minutes in an ultra-sonic bath, followed by one immersion in 2:1 mixture of MeOH:$CHCl_3$ for 30 minutes, also in an ultra-sonic bath, and then by a final rinse in distilled water for 30 minutes, again in an ultra-sonic bath. Following this, samples were freeze dried and 0.8 ± 0.1 mg were weighed into tin capsules.

Finally, isotopic analyses were carried out at the Godwin Laboratory, Department of Earth Sciences. Each sample was analysed in triplicate using a Costech elemental analyser (EA) coupled in continuous-flow mode to a Finnigan Delta V mass spectrometer. Carbon and nitrogen results are reported as δ values in parts permil (‰), in comparison to international standards, VPDB and AIR respectively (see Chapter 2, Table 2.2, p. 34). Based on replicate analysis of international and in-house standards, analytical errors are less than 0.2‰ for both carbon and nitrogen.

4.5.3. *ASSESSMENT OF QUALITY OF RESULTS*

The isotopic results of hair and nail samples were assessed as valid using the carbon-nitrogen (C/N) ratio and replicate measurement errors. The theoretical atomic C/N ratio of keratin is 3.4, and any values that fall outside the range of 3.0 to 3.8 suggest the

sample was contaminated, and thus, that the result should be discarded (O'Connell & Hedges, 1999a; O'Connell et al., 2001). In addition, replicate measurement errors are typically around 0.3‰ for $\delta^{13}C$ and 0.4‰ for $\delta^{15}N$. Accordingly, in each instance where the standard deviation across replicates was higher than these values, the data were inspected in order to identify the source of variation (e.g. one run driving the variation due to a glitch in the analysis). If necessary, more replicates were analysed.

4.6. BREATH STABLE ISOTOPE METHODOLOGY

The protocols used for breath collection and analysis were those used in the Garrod Laboratory at the McDonald Institute of Archaeology, University of Cambridge, and based on Slater et al. (2004).

4.6.1. FIELDWORK PROTOCOL

Ideally, breath collection takes place after an overnight fast, first thing in the morning (Hedges et al., 2009; Slater, Preston, & Weaver, 2004). However, this was logistically impractical, considering I often spent the night at least 30 minutes away from the study site, and that I had to adjust data collection to the everyday demands of the participants. Instead, I aimed to collect breath just before lunch, at 11am, or at least three hours after a meal, as already described.

A breath sample was obtained by having the subject blow through a straw into a 12 ml evacuated Exetainer vial (Labco, High Wycombe, United Kingdom) (Figure 4.5) (Slater et al., 2004). Prior to sampling, all tubes were labelled with the subject's code, the date, and the time of last meal. Next, subjects were requested to inhale deeply, hold it, and place the straw in their mouth. Meanwhile, the vial was uncapped, and the other end of the straw guided to the tube bottom. Subjects would then exhale fully, and when they were almost finished, the vial was pulled and capped immediately. Afterwards, the time of sampling was recorded on the tube and a square of parafilm wrapped around the vial's top, for safer transport. Finally, the procedure was repeated in cases

where sampling success was not certain, such as when participants accidently left the straw inside the tube or when exhalation was weak.

Figure 4.5 Collecting breath among the Turkana in Nakurio, Turkana, Kenya. Photo courtesy of Joe Jeffery.

4.6.2. LABORATORY PROTOCOL AND ANALYSIS

Breath isotopic analyses require no preparation prior to isotopic analysis. However, whilst hair and nail samples are fed to the IRMS through an EA inlet system, breath requires a gas chromatography (GC) inlet system (Chapter 2, Figure 2.4, p. 35). Charles Belanger, at the Iso-Analytical Limited (Crewe, United Kingdom), carried out the breath analyses. In this laboratory, each sample was analysed once using a GC coupled in continuous flow with a Europa Scientific Hydra 20-20 mass spectrometer. In addition, these analyses were conducted against a reference gas (IA-R060) diluted to 3.3% CO_2, which has been calibrated against a secondary reference (NBS-19) that was traceable to the international VPDB standard. Like before, carbon results are reported as δ values in parts per mil. Finally, and based on replicate analysis of in-house standards, analytical errors are less than 0.2‰.

4.6.3. ASSESSMENT OF QUALITY OF RESULTS

In contrast to hair and nail protocol, the C/N ratio in breath cannot be used to assess the quality of breath isotopic results, since nitrogen is absent from these samples. Therefore, it is difficult to evaluate the contamination of breath samples. Nevertheless, samples with insufficient CO_2 (<0.1%) were not analysed and samples with low CO_2 concentration (<3%) were marked, and were treated with caution in later analysis.

4.7. FOOD STABLE ISOTOPE METHODOLOGY

The protocols used for food collection and analyses were those used in the Garrod Laboratory at the McDonald Institute of Archaeology, University of Cambridge. Protocols for fish collagen were based on Richards & Hedges (1999), while plant material requires no pre-treatment, and hence can be analysed directly after grinding. Notably, two of the fish samples (*Dagaa* and *Ofulu)* were whole fish, and not just the bone. Therefore, these samples were analysed twice: once, whole; and another time just the bone (after dissection). The protocol used in both instances was that of bone collagen.

4.7.1. FIELDWORK PROTOCOL

Each sample was prepared for transport by placing it in a 50 ml plastic container, screwing the lid on tightly, and wrapping Parafilm around the lid. In fish samples, for better preservation, silica gel was added to the containers before closing them.

4.7.2. LABORATORY PROTOCOL AND ANALYSIS

Extraction of collagen from fish bone involves four steps: defatting, which cleans fat from the bone; demineralisation, which removes the inorganic component of the bone; gelatinisation, which dissolves collagen and allows removal of contaminants; and freeze-drying, which extracts water (Figure 4.6) (Richards & Hedges, 1999).

During defatting, fish bones were soaked in 200 ml of 2:1 MeOH:$CHCl_3$, placed in an ultra-sonic bath for 30 minutes, and then left standing overnight in the fume hood. The following day, if there was any fat left (i.e. fat accumulated at the bottom of the tube), the samples were soaked in new solution, and again subject to the procedure above. If there was still fat left after three rinses, the procedure was repeated consecutively, i.e. without the overnight standing period, until all fat was eliminated. Once this was achieved, the samples were rinsed twice in 50 ml distilled water, placing them in an ultra-sonic bath for 30 minutes each time.

Afterward, samples were weighted, and then demineralised in 8 ml 0.5 mole (M) hydrochloric acid (HCl) at 4°C. Samples were shaken twice a day and acid was replaced every two to four days. On average, it took one week for samples to demineralise completely. Once demineralised, samples were rinsed three times in distilled water.

Subsequently, samples were gelatinised in acidic water (pH 3.0) at 75°C for forty-eight hours. The solution was then filtered into pre-weighed tubes using an Eezee filter (8 μm), and then freeze-dried. Once freeze-dried, the tubes were weighed in order to estimate collagen yield. Finally, 0.8 ± 0.1 mg of the resulting collagen was weighed into tin capsules for analysis.

In contrast, preparation of plant samples for analysis was far simpler. These were ground until homogeneous using a percussion mortar, and then weighed into tin capsules for analysis. Importantly, $\delta^{13}C$ and $\delta^{15}N$ had to be measured in two different runs of analyses, since the C/N ratio of plants is highly variable. In particular, nitrogen content is lower than carbon content for most plants. Hence, a first analysis used 1 ± 0.1 mg to measure $\delta^{13}C$, while a second analysis measured $\delta^{15}N$ using a weight calculated from the carbon to nitrogen ratio detected in the first analysis.

Finally, and in the same way as for hair and nail, isotopic analyses were carried out at the Godwin Laboratory, Department of Earth Sciences. Each sample was analysed in triplicate using a Costech EA coupled in continuous-flow mode to a Finnigan Delta V

mass spectrometer. Carbon and nitrogen results are reported as δ values in parts per mil, in comparison to international standards, VPDB and AIR respectively (see Table 2.2, p. 34). Furthermore, and based on replicate analysis of international and in-house standards, analytical errors are less than to 0.2‰ for both carbon and nitrogen.

Figure 4.6 Analyzing fish samples in the Garrod Laboratory, University of Cambridge. Photos: fish samples demineralization (top left); fish samples in freeze-dryer (bottom left); collagen weighing (right). Photo courtesy of Josie.

4.7.3. ASSESSMENT OF QUALITY OF RESULTS

In the same way as hair and nail, the isotopic results of collagen samples were assessed as valid using the C/N ratio and replicate measurement errors, but also the collagen yield. The theoretical atomic C/N ratio of collagen is 3.17, and any values that fall outside the range of 2.9 to 3.6 suggest the sample was contaminated, and thus, that the result should be discarded (DeNiro, 1985). In the same way as for hair and nail, replicate measurement errors fall around 0.3‰ for $\delta^{13}C$ and 0.4‰ for $\delta^{15}N$. Accordingly, in each instance where the standard deviation across replicates was higher than these values, the data were inspected in order to identify the source of

variation (e.g. one run driving the variation due to a glitch in the analysis). If necessary, more replicates were analysed. In addition, the collagen yield, in percentage of the start mass of the bone sample, is an indicator of preservation. In general, carbon yields over 13% and nitrogen yields over 5% are judged as well-preserved. Preservation, however, is unlikely to be a concern in the samples examined here, since they are modern. In fact, modern samples have carbon yields from 40 to 50%, and nitrogen yields between 14 and 18% (Ambrose, 1990).

There is no equivalent independent test to evaluate contamination in plants, since their C/N ratios are highly variable, as mentioned in the above section.

4.8. STATISTICAL ANALYSIS

Statistical analyses were performed on the R platform (R Core Team, 2016 version 3.3.1), with all graphics produced within the same program. The only exception is that for power analyses, I used the program G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). For the sake of brevity, the code used for statistical analysis, as well as many of the analysis details, are included in the Appendices of each Chapter.

The normality of the data was checked graphically using histograms or quantile-quantile plots, and formally using Shapiro-Wilk tests. In addition, equality of variances was tested using Levene's test. Furthermore, descriptive statistics of central tendency and variability included the mean and standard deviation, but also medians and interquartile ranges. These last two measures were reported because they are robust to outliers and non-normal data, whereas the first two were reported because Shapiro-Wilk test is very sensitive to outliers, and because other studies (used for comparison) report only these measures, and not medians or interquartile ranges (Field, Miles, & Field, 2012, p. 182).

The statistical tests used to investigate the data are described throughout the chapters, so that the choice of each test is justified within the context of its statistical question.

However, one can say that, as a rule, when the assumptions of the test were not met, robust versions of the tests were used (e.g. when data were non-normal, I used a test with bootstrapping and trimmed means, instead of the standard test). In addition, and whenever possible, this work reported on the effect sizes of the tests. An effect size is a quantitative measure of the magnitude of an observed phenomenon that is less dependent on sample size than traditional tests of statistical significance, and thus, can aid data interpretation (Field et al., 2012, p. 57).

4.9. ETHICS AND HEALTH & SAFETY

A final issue to address regards the ethical implications of working with living participants in remote areas. Before any data collection took place, ethical and safety issues were discussed with my supervisor and with relevant individuals. All procedures taken followed guidelines set by the Safety Manual of the Department of Archaeology and Anthropology (June 2014 version).

Fieldwork data collection in Kenya fell under the IN-AFRICA Project permit granted to Prof. Marta Mirazón Lahr by the Government of the Republic of Kenya (NACOSTI/P/15/2669/4758). An exploration license was also obtained from the National Museums of Kenya that addresses this thesis data collection. Copies of these can be found in Appendix A, in Figure A.5 and A.6.

Risk assessments, insurance, and 'Leave to Work Away' were in place before the beginning of all fieldwork. Ethics approval for fieldwork in Kenya was sought at the Department Level, as per University guidelines (Appendix A, Figure A.7). During fieldwork, and prior to data collection, all procedures were duly explained to participants. This explanation used the support of posters that contained images of every step in the procedure and brief descriptions in English and Swahili (included in Appendix A, Figure A.8). After that, all Kenyan participants signed a consent form; this form was also written in English and Swahili and participants kept the Swahili version, which included the researcher contacts for posterior reference. In cases where

the participant was illiterate, consent forms were read out to them and a thumb print was taken in lieu of a signature. Two examples of signed consent forms are included in Appendix A (Figure A.9), one with a signature and one with a thumb print. Safety measures regarding human tissue and dental moulds collection are addressed above.

Regarding transport into the United Kingdom, no special measurements were taken for the following reasons: first, hair, nail, and breath do not fall under the Human Tissue Act 2004, according to the Human Tissue Authority; second, dental moulds, once properly disinfected, are non-hazardous; and third, food samples did not exceed 2 kilograms (kg) and did not contain any animal products and so were not subject to any special import rules ("Bringing food, animals or plants into the UK," 2014).

Hair and breath samples from Cameroon (Baka) were obtained through collaboration with Dr. Fernando Ramírez-Rozzi at the CNRS. Consent was oral, as commonly used in his research projects (Romero et al., 2013). Dr. Ramírez-Rozzi felt this was the most appropriate approach, considering the Baka cannot read or write and dislike having their fingers dirty with ink. His work falls under the agreement between the Ministére de la Recherche Scientifique et Technique in Cameroon and L'Institut de Recherche pour le Devélopment (IRD) in France. Appendix A contains Dr. Ramírez-Rozzi declaration on the informed consent of the Baka (Figure A.10), and an ethics statement regarding his work in Cameroon signed by the Head of the CNRS (Figure A.11).

Laboratory work was conducted in the Dorothy Garrod Laboratory, Department of Archaeology & Anthropology, University of Cambridge. All risk assessments were in place and adequate training was provided by Catherine Kneale, Senior Research Assistant at The McDonald Institute for Archaeological Research. A final ethics statement was obtained from the Department of Archaeology & Anthropology at Cambridge regarding the analysis of samples collected in Cameroon (Appendix A, Figure A.12).

4.10. SUMMARY OF CHAPTER 4

This chapter has outlined the methodology employed in this thesis, covering the sampling strategy, and the collection and analysis of: dietary information, human tissues and food, and dental moulds and casts. In addition, the assessment of the quality of isotopic results is discussed in the appropriate sections. Following this, a section on statistical analysis describes the programs and statistics used in this work; and a final section discusses the ethical and safety measures taken during the field and laboratory work.

CHAPTER 5

DIET ASSESSMENT RESULTS AND DISCUSSION

This chapter focuses on the characterisation of the diet of Kenyan populations through semi-quantitative food questionnaires. To achieve this, the chapter addresses (1) the composition of foods that make up the populations' diets, (2) the dietary intake of these people, (3) the adequacy of the diet to fulfil nutritional demands, in the light of the dietary recommendations of the World Health Organization (WHO), the Food and Agriculture Organization (FAO) and the United Nations University (UNU), (4) the composition of diets in terms of macronutrients and food types, and (5), the comparability of diet characterised through questionnaires and that reported by the ethnography, as reviewed in Chapter 3.

5.1. FOOD COMPOSITION

To recapitulate from the previous chapter, the nutritional content of the populations' diet was estimated using food composition tables, namely the IML (2004), the FAO (2016), and the USDA (2015). Table 5.1 lists compositional information regarding all food items mentioned by participants during food questionnaires, whereas the responses to all diet questionnaires are found in Appendix B.1.

From this table, two points become apparent. First, some of the food items are recipe dishes, of which some are produced in factories, and some are prepared at home: the former includes bread, spaghetti, soda, and biscuits, while the latter comprise *githeri*, a dish made with beans and maize, *ugali*, a thick porridge of maize flour (already mentioned in Chapter 3, p. 70), *chapati*, a fried wheat flat-bread, and *mdazi*, a fried wheat dough (to which sometimes sugar is added). Such mixed dishes are difficult to assign to one of the food categories mentioned in Chapter 4 (i.e. animal, fish, C_3 plant, C_4 plant, or other). For instance, *githeri* is made of C_3 beans and C_4 maize. However, all the above are made mostly of plant foods, and hence, all fall under either the C_3 or the C_4 category. Ultimately, I opted to assign the category C_4 if any C_4 ingredient was present, and the C_3 category to the remaining. Home prepared dishes pose an additional problem. Unlike products produced in factories, which follow standardized recipes, these dishes have nutritional values that vary both between households and within households. Notwithstanding, by calculating mean intakes across relatively long periods of time, as is the case in semi-quantitative food questionnaires, one dilutes the impact of this variation (Murphy et al., 2004, 1991). Second, Table 5.1 reveals the likely main food staples of these populations based on the number of times individuals declared to consume a given food item. These include milk (n = 145), eggs (n = 132), and fish (n = 142), but also rice (n = 155), beans (n = 153), ugali (n = 158), and githeri (n = 139). In particular, the high prevalence of fish is expected, considering the three fisher populations included in the sample. These results, however, are only indicative, since they do not take into account frequency of consumption.

Table 5.1
List of all food items mentioned by participants during the diet questionnaires and of all relevant information for energy and nutrient analysis.

Food Item	Category [a]	Count [b]	Source Code [c]	Medium Portion (g) [d]	Energy (kcal/100g)	Protein (g/100g)	Fat (g/100g)	Carbohydrate (g/100g)
Milk	Animal	145	1	184	69	3.6	4.1	4.5
Eggs	Animal	132	1	112	155	12.6	10.6	1.1
Poultry	Animal	127	1	84	285	26.9	18.9	0
Cattle	Animal	124	1	112	269	24.9	18	0
Goat	Animal	108	1	85	241	32.9	10.9	0.5
Sheep	Animal	91	1	85	265	16.9	21.4	0
Camel	Animal	44	2	112	267	19.6	20.3	—
Pork	Animal	35	1	56	273	27.6	17.2	0
Donkey	Animal	16	2	112	175	28.14	6.05	0
Blood	Animal	12	3	184	180	24.9	3.6	11.9
Turtle	Animal	10	2	112	138	26.8	2.9	0.9
Other Birds	Animal	8	1	112	285	26.9	18.9	0
Crocodile	Animal	8	2	112	121	19.4	4.8	0
Cheese	Animal	8	1	150	403	24.9	33.1	1.3
Other Eggs	Animal	4	1	64	155	12.6	10.6	1.1
Butter	Animal	4	2	7	105	18.5	2.9	0
Dog [e]	Animal	1	—	112	—	—	—	—
Frogs	Animal	1	2	112	73		0.3	0
Wild cat [e]	Animal	1	—	112	—	—	—	—
Termites	Animal	1	2	25	535	35.14	42.6	2.8
Fish	Fish	142	1	112	105	18.5	2.9	0
Rice	C_3 plant	155	1	175	130	2.4	0.2	28.6
Beans	C_3 plant	153	1	185	127	8.7	0.5	22.8
Cabbage	C_3 plant	123	1	60	22	1	0.4	4.5
Fruit	C_3 plant	120	1	119	92	1	0.5	23.4
USO's (e.g. potatoes)	C_3 plant	115	1	122	93	2	0.1	21.6

Table 5.1 Continued

Food Item	Category [a]	Count [b]	Source Code [c]	Medium Portion (g) [d]	Energy (kcal/100g)	Protein (g/100g)	Fat (g/100g)	Carbohydrate (g/100g)
Cowpeas	C_3 plant	65	1	171	116	7.7	0.5	20.8
Bread (White)	C_3 plant	49	1	44	274	8.8	3	51.9
Tomatoes	C_3 plant	48	1	62	21	0.9	0.3	4.6
Carrots	C_3 plant	46	1	62	45	1.1	0.2	10.5
Spaghetti	C_3 plant	43	1	250	141	4.8	0.7	28.3
Doum Palm Fruit	C_3 plant	28	3	119	390	3.9	0.8	84.1
Chapati & Mdazi	C_3 plant	26	1	125	288	6.3	9.1	49
Pumpkin	C_3 plant	14	2	94	31	1.1	0.1	6.1
Sweet Potatoes	C_3 plant	10	1	114	103	1.7	0.1	24.3
Cassava	C_3 plant	3	1	114	131	1.1	0.3	31.9
Beetroot	C_3 plant	1	2	94	44	1.68	0.18	9.96
Ugali	C_4 plant	158	1	114	125	2.1	1.3	26.2
Githeri	C_4 plant	139	1	134	163	6.3	1.4	31.1
Sorghum	C_4 plant	103	1	83	332	10.3	3	73.7
Sugar Cane	C_4 plant	20	1	3	376	0	0	97.3
White Maize	C_4 plant	1	1	83	108	3.3	1.3	25.1
Sweets &Soda	C_4 plant	118	1	340	41	0	0	10.4
Honey	Other	18	1	20	304	0.3	0	82.4
Biscuits	Other	3	1	62	478	5.1	21.1	67.9

Note. USO = underground storage organs.

[a] It should be stressed that categorization is imperfect (Chapter 4, p. 93); for example, animal should be divided into C_3 and C_4 animal, according to their diet; however, it is impossible to reach that level of detail within the food web.

[b] Count = Number of times a food item was reported; food items are ordered from most reported to less so, within each category.

[c] Source Code: 1 = IML (Murphy et al., 2004, 1991);
2 = FAO (2016);
3 = Galvin (1985).

[d] Portion sizes are based on USDA, as reported by Gibson (1993, pp. 16-17) (Chapter 4, p. 91).

[e] It was not possible to find nutrient information for 'dog' or 'wild cat'. Considering these were only mentioned once, it was opted to exclude them from the analysis.

5.2. DIETARY INTAKE

Prior to comparisons across food categories, or across populations, one must describe and consider the estimation of dietary intake. To achieve this, wet weight, energy, and macronutrient intake for the Kenyan populations studied are summarised in Table 5.2 and Figure 5.1. In brief, these reveal low food intake and high variability across all populations. Furthermore, Figure 5.1 singles out several participants as outliers in their respective population. In Table 5.2, some of the data sets are not normally distributed, according to Shapiro Wilk tests (Appendix B, p. 325-328). For this reason, and as described in the previous chapter, this table reports both standard and robust measures of central tendency and variability: i.e. means and standard deviations, and medians and interquartile ranges.

Before evaluating the adequacy of these diets, one must address possible issues with measurements, and ascertain the reliability of these results, according to food requirements set by international bodies.

Errors in the collection and recording of food consumption data encompass many aspects (Gibson, 1990, p. 86), such as respondent biases, e.g. participants report what they perceive as a 'good diet', interviewer biases, e.g. researchers probe for information to different degrees, respondent memory lapses, e.g. participants unintentionally omit or add consumed foods, incorrect estimation of food portion size, e.g. participant's concept of an average portion deviates from the standard, and coding and computation errors, e.g. dietary information is misreported or inaccurate nutritional values are used to calculate consumption.

Table 5.2
Summary of wet weight, wet energy, and nutrient intake for all populations, and by gender.

Population	*n*	Wet Weight [a] (g/day)		Wet Energy [b] (kcal/day)		Protein (g/day)		Fat (g/day)		Carbohydrate (g/day)	
		mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	31	853 (186)	*867 (227)*	1000 (226)	*1016 (190)*	**60 (13)**	*63 (12)*	24 (7)	*22 (10)*	136 (37)	*134 (34)*
male	14	873 (186)	*931 (228)*	995 (276)	*1004 (327)*	60 (12)	*60 (15)*	24 (7)	*22 (10)*	136 (52)	*136 (69)*
female	17	837 (190)	*844 (154)*	1003 (184)	*1032 (151)*	**60 (14)**	*63 (5)*	25 (8)	*23 (10)*	135 (22)	*134 (24)*
Turkana	37	**612 (432)**	*459 (472)*	**980 (552)**	*892 (652)*	**41 (29)**	*33 (21)*	**16 (12)**	*12 (8)*	152 (83)	*141 (110)*
male	18	797 (538)	*616 (740)*	1114 (590)	*1011(794)*	**51 (36)**	*36 (46)*	**21 (22)**	*14 (13)*	174 (95)	*174 (92)*
female	19	437 (184)	*392 (203)*	**853 (495)**	*797 (370)*	**32 (16)**	*26 (16)*	**11 (6)**	*9 (6)*	131 (65)	*126 (59)*
Luhya (Webuye)	32	783 (220)	*758 (338)*	**875 (310)**	*783 (375)*	41 (15)	*40 (22)*	18 (7)	*18 (10)*	**138 (59)**	*133 (64)*
male	16	747 (226)	*677 (268)*	815 (265)	*741 (324)*	38 (16)	*36 (160)*	17 (8)	*15 (12)*	130 (40)	*129 (64)*
female	16	818 (215)	*790 (352)*	**936(347)**	*882 (324)*	44 (13)	*46 (20)*	20 (5)	*20 (8)*	**147 (74)**	*139 (68)*
Luhya (Port Vict.)	31	1037(307)	*1027 (478)*	1176 (344)	*1215 (464)*	61 (18)	*61 (26)*	22 (9)	*22 (14)*	187 (57)	*185 (79)*
male	14	1154 (297)	*1140 (319)*	1350 (328)	*1318 (316)*	70 (20)	*74 (19)*	27 (8)	*29 (12)*	213 (53)	*215 (60)*
female	17	941 (289)	*940 (285)*	1033 (302)	*992 (370)*	53 (12)	*55 (16)*	**19 (8)**	*17 (10)*	166 (53)	*170 (61)*
Luo (Port Vict.)	29	1131 (273)	*1102 (439)*	1199 (307)	*1172 (408)*	58 (13)	*56 (13)*	20 (5)	*20 (8)*	200 (60)	*197 (93)*
male	14	1228 (260)	*1216 (363)*	1285 (322)	*1184 (438)*	62 (13)	*63 (17)*	22 (5)	*23 (7)*	215 (62)	*188 (82)*
female	15	1041 (261)	*1019 (350)*	1118 (280)	*1137 (412)*	54 (13)	*55 (11)*	19 (5)	*19 (5)*	187 (57)	*204 (85)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range. Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 323-326).

[a] Wet weight is the weight of consumed diet, including macronutrients, but also water, ash, and others.

[b] Wet energy is the total energy of consumed diet, including macronutrients, but also water, ash, and others.

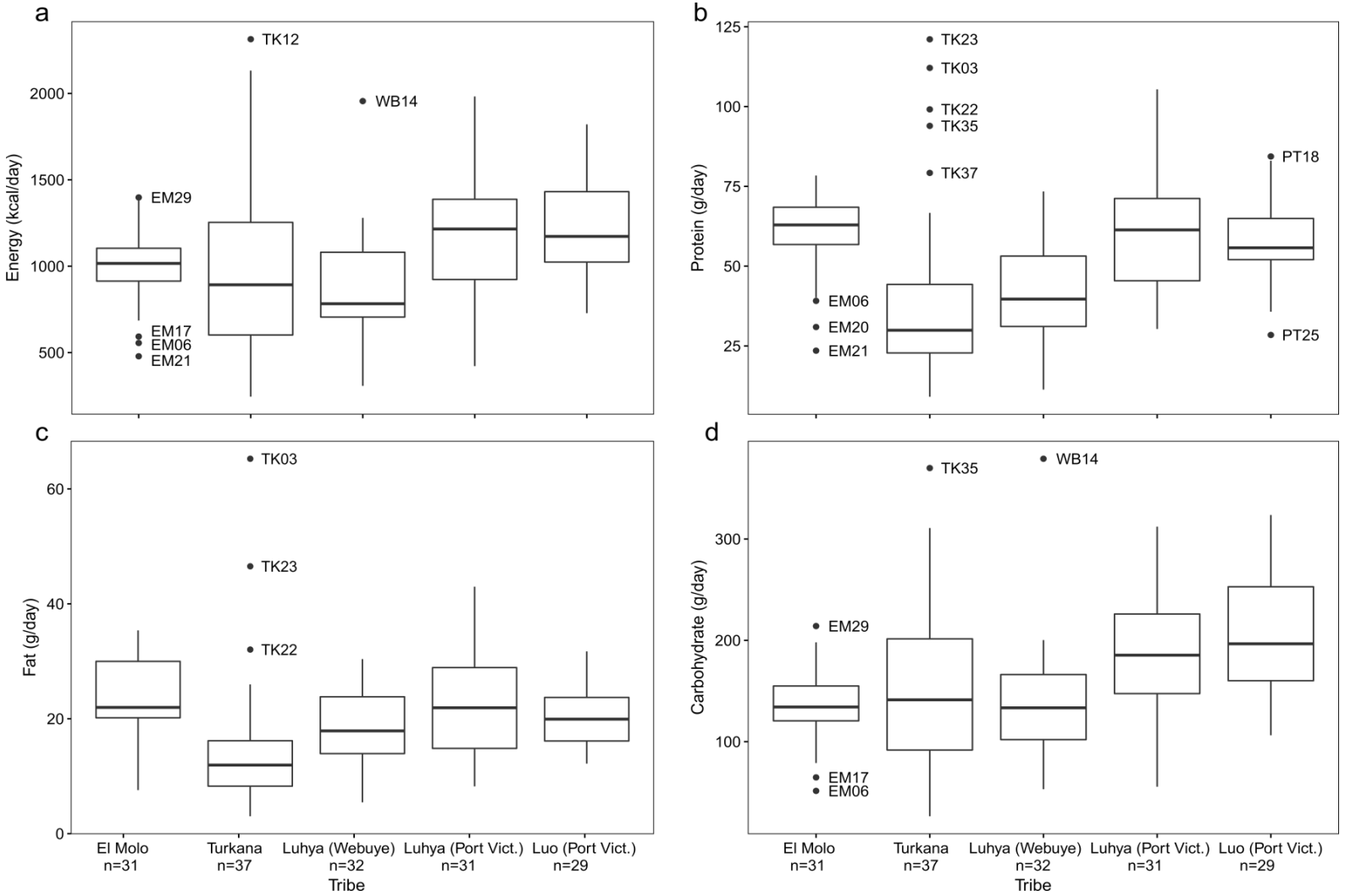


Figure 5.1 Box plots of energy (a) and nutrient (b, c, d) intake for all populations. Outliers are labeled with identification code of individual and number of individuals per population/ box plot are indicated below the x-axis label with 'n='.

Although all the above mentioned errors are likely to affect the results presented here to some degree, several mechanisms were in place to counteract them. For instance, questioning participants on long lists of foods aimed at decreasing respondent memory lapses. In fact, and as mentioned in the previous chapter, incorrect estimation of portion size is probably the largest source of measurement error for recall methods (Gibson, 1990, p. 89). Nonetheless, Hunter et al. (1988) found that the use of portion sizes in semi-quantitative food questionnaires does not introduce a large error in the estimation of food and nutrient intake. In addition, Gibson (1990) goes on to say that recall methods should suffice for estimation of energy and macronutrient consumption, although they may be inadequate to study micronutrient consumption (e.g. indispensable amino acids), which necessarily deal with very small amounts. As a final point, since the same questionnaire was used for all Kenyan populations, the level of error is more or less constant across participants. This means that these results are comparable amongst themselves, even if inter-study errors require caution in the comparison with results from other studies, such as the ones conducted among the Turkana (Galvin, 1985) or the Baka (Sato et al., 2012; Yamauchi et al., 2000).

Despite the cautionary point on comparability, one can only address the high within-population variability observed in this study (Table 5.2, Figure 5.1) by considering it vis-à-vis other research. In the literature, such fluctuations in dietary intake are common. By way of illustration, when using twenty four hour recalls collected from over 100 individuals, Galvin (1985, pp. 190, 196) estimated energy intakes for Turkana men during the wet season that varied from 400 kcal/day to more than 4000! Similar ranges were reported for women and children for all seasons, with the smallest range for all groups and seasons well above 1000 kcal/day. This range decreased for the weighted records results from 28 individuals (i.e. where the researcher recorded, weighed, and measured all food consumed for a specific period), but never dropped below a range of 500 kcal/day. In addition, Galvin describes similar variation to the one reported here for protein intake, the only nutrient studied in her thesis. A final point concerns the outliers identified in Figure 5.1. By reviewing the corresponding questionnaires, it was confirmed that these outliers are not the result of mistyping

errors. In addition, removing these outliers from the data set simply leads to the labelling of other individuals as outliers, based on the amended inter quartile range (Appendix B, p. 322-325). Instead, these outliers represent part of the true variation observed in these populations' diet and were kept in the data set for all further analysis (Field et al., 2012).

Taken together, all the above demonstrates that the variation observed in this study is real, and not exclusively a product of methodological inaccuracies. However, dietary intake may be biased by incorrect estimation of portion sizes. One must now look at whether this diet meets dietary requirements for adults.

5.3. DIETARY ADEQUACY

To determine if energy and nutrient intake meets population needs, one must compare dietary intake values with either (1) published models of diet requirements, or (2) estimates of energy expenditure and nutrient losses from the body. Considering that no analyses of energy expenditure were carried out for this study, I will apply the first approach, by comparing energy and macronutrient values to FAO, WHO, and UNU recommendations (FAO, 2010; WHO/FAO/UNU, 2004, 2007; WHO/FAO, 2003).

In addition, and considering that these organisations often report on macronutrient intake as a percentage of total energy, these percentages were calculated for each macronutrient. To achieve this, a protein digestibility of 80% was applied to protein intake and Atwater factors were used to convert macronutrient intake from g/day into kcal/day (FAO, 2003, pp. 57–58). This was then divided by the energy intake per day from Table 5.2 to obtain the macronutrient:energy percentages reported on Table 5.3.

Table 5.3
Macronutrient intake as percentage of total energy for all populations, and by gender.

Population	*n*	Protein (%)	Fat (%)	Carbohydrate (%)
El Molo	31	19	22	54
male	14	19	22	55
female	17	19	22	54
Turkana	37	13	14	63
male	18	15	16	64
female	19	12	11	62
Luhya (Webuye)	32	15	19	63
male	16	15	18	64
female	16	15	19	63
Luhya (Port Vict.)	31	17	17	64
male	14	17	18	63
female	17	17	16	64
Luo (Port Vict.)	29	15	15	67
male	14	16	15	67
female	15	15	16	67

Whereas the specifics of protein digestibility are discussed in the section below on protein intake, the Atwater factors, and their relevance, are discussed here. To convert weight units to energy units, the Atwater system uses a single factor for each macronutrient, regardless of the food in which it is found. This system is based on the heats of combustion of macronutrients, after correction for losses in digestion, absorption, and urea excretion. These factors are 4.0 kilocalories per gram (kcal/g) for protein, 9.0 kcal/g for fat, and 4.0 kcal/g for carbohydrates (FAO, 2003, p. 23). However, the Atwater system has been the subject of much dispute since it was first developed in 1896. First, this system groups fibre with carbohydrate, and fibre has a lower energy factor of 2.0 kcal/g. Second, macronutrients from different foods yield different amounts of energy per gram. For instance, factors for protein vary from 2.44 kcal/g to 4.36 kcal/g, while factors for fat vary from 8.37 kcal/g to 9.02 kcal/g, and factors for carbohydrate vary from 2.70 kcal/g to 4.16 kcal/g. Third, and last, the Atwater system is based on the heats of combustion of macronutrients, and not on the actual energy available to produce adenosine triphosphate (ATP) in the body. In particular, the

Atwater system encompasses the energy from food that is lost to microbial fermentation and obligatory thermogenesis, and thus, fails to enter the body energy metabolism (FAO, 2003, pp. 24–26). Nevertheless, the Atwater system remains, in many cases, the most adequate system to convert g/day to kcal/day. In detail, FAO (2003, pp. 44-59) informs that when food composition tables do not distinguish between different types of carbohydrates, as is the case here (Table 5.1), the more general conversion factor of 4.0 kcal/g is the most suitable. In addition, using different conversion factors for each food item has a small impact on calculated energy values (difference around 5%), and should only be used when specific analyses of food items have been conducted. And finally, the energy requirements established by the WHO/FAO/UNU report (2004) are based on energy expenditure estimations (plus the energy needs for growth, pregnancy, and lactation). This means that the energy requirements are based on body heat production, which includes the heat of microbial fermentation and obligatory thermogenesis, and thus equate conceptually to an energy intake estimated through the Atwater system, rather than just on available energy for ATP production.

As a blind test, if one converts the population's macronutrient intake into kcal/day using the Atwater system, the total energy values are close to the energy values calculated through the food composition tables. More specifically, the differences range from 25 kcal/day for the Luhya (Webuye) to 97 kcal/day for the Turkana. Note that this is the reason why macronutrient:energy percentages on Table 5.3 do not add to 100% for each population (since these percentages were calculated based on energy intakes from the food composition tables). Thus, the missing fraction of the macronutrient:energy percentage pertains to energy from other sources in food items, such as alcohol.

5.3.1. ENERGY

Broadly speaking, individuals require energy for two main processes, basal metabolism, i.e. the functions that maintain life, and physical activity. In addition, and

where applicable, individuals use energy for growth, pregnancy, and lactation. To meet these demands, individuals obtain energy from the macronutrient constituents of foods, mostly from carbohydrates and fat, but also from protein (when dietary energy intake is limited) (see Chapter 2, p. 43). Moreover, the exact energy requirements of an individual vary according to the individual's gender, age, body size, body composition, and habitual physical activity (WHO/FAO/UNU, 2004). As a result, energy recommendations of the WHO/FAO/UNU report (2004) take into account all of these factors.

Nonetheless, it quickly becomes clear that the calculated values of energy intake in this study are considerably lower than the WHO/FAO/UNU (2004) energy requirements for all populations, across all age and gender groups. In fact, the lowest reported WHO energy requirement for adults is around 1700 kcal/day and it pertains to a woman aged 18 to 60, with a sedentary lifestyle, a mean weight of 45 kg, and a height ranging from 134 to 156 cm. In parallel, the lowest value for a man with the same age and lifestyle, but with a weight of 50 kg, and height ranging from 142 to 164 cm, is 2100 kcal/day. Alternatively, one may use data collected by Jeffery (2018) on height and weight among the El Molo and the Turkana (and still assuming a sedentary lifestyle), to estimate more adequate energy requirements for these populations (WHO/FAO/UNU, 2004, pp. 41–45). Since height and weight data are not available for the other populations of this study, such level of precision is impossible to achieve for those groups, although it is still possible to compare it to the minimum energy requirements mentioned above. Nevertheless, the energy values reported in Table 5.2 are still considerably lower than the energy requirements estimated from the height and weight data on Table 5.4 (for Turkana and El Molo), and also lower than minimum energy requirements, suggesting that these populations do not obtain enough energy to meet their energy needs.

Table 5.4
Mean adult height and weight, by gender, for the El Molo and the Turkana, the correspondent WHO energy requirement based on these data, and the estimated energy intake. Height and weight data from Jeffery (2018).

		Height		Weight		WHO Energy Requirement	Estimated Energy Intake
Population	Age Group	*n*	mean (cm)	*n*	mean (kg)	(kcal/day)	(kcal/day)
El Molo							
male	25-99	39	170	39	56	2200	995
female	25-99	52	159	49	49	1800	1003
Turkana							
male	25-99	27	175	24	54	2200	1114
female	25-99	41	162	38	47	1800	853

As comparison, Galvin (1985) also reports low levels of energy consumption for the Turkana. Galvin (1985, pp. 222, 374) estimated an average daily intake of 981 kcal/day for women and of 1840 kcal/day for men when using twenty-four recalls, and of 1074 kcal/day for women and of 2028 kcal/day for men when using weighed records. Interestingly, recalled values were lower than measured ones, which might point to a similar underestimation of the values as in the present study. In addition, the energy expenditure (women: 1731 kcal/day; men: 2159 kcal/day) calculated by Galvin (1985) was higher than the energy intake. In other words, both men and women were in negative energy balance, and consumed less than what they spent in daily activities. This latter result was partly expected, considering recorded seasonal variations in body weight (Galvin, 1985, p. 271). However, since energy balance must be attained in the long term, it is possible that this was achieved through occasional days of high energy intake, not included in the study. This explanation is further supported by anecdotal accounts of binge eating among Turkana women (Galvin & Little, 1999). Finally, Galvin's (1985, p. 273) measurements of the daily energy expenditure of the Turkana are lower than those of the 'sedentary lifestyle' from which the lowest WHO/FAO/UNU (2004) energy requirements are calculated. In fact, it has long been suggested that FAO energy requirements are inflated (Durnin, Edholm, Miller, & Waterlow, 1973). And indeed, the equations used to calculate energy expenditure values are based on studies conducted in Western Europe and North America. Hence, they fail to represent the full geographical, cultural, and economic background of the

human population, and may be inadequate in some cases (WHO/FAO/UNU, 2004, pp. 8, 37).

In brief, it is likely that the estimated values of energy intake (gleaned from semi-quantitative food questionnaires) underestimate the actual energy intake of the populations of this study. Indeed, the estimated energy intake is inferior to WHO/FAO/UNU energy requirements. However, one must consider that these energy requirements probably reflect inflated estimations of energy expenditure. Despite the underestimation of energy intake in this study and the overestimation of energy requirements by WHO/FAO/UNU, the difference between these two values is so substantial that one may state that some of the groups studied here suffer from inadequate energy consumption, particularly the El Molo, the Turkana, and the Luhya (Webuye). In the cases of the El Molo and Turkana, this hypothesis is consistent with the levels of malnutrition detected by Jeffery (2018), which will be further explored in the discussion, and the results obtained by Galvin in her earlier study of the Turkana, where a negative energy balance was also observed.

5.3.2. PROTEIN

In simple terms, individuals require protein for all processes that consume amino acids and produce waste (see Chapter 2, p. 48). Similarly to energy, this includes maintenance activities, and special needs, such as growth, pregnancy, and lactation (WHO/FAO/UNU, 2007, p. 7). Also, as for energy, the amount of protein required to meet these demands depends on several factors. These must be discussed before addressing the protein requirement *per se*. First, protein requirements depend on the digestibility of the protein source, i.e. on the amount of food protein that is bioavailable to the body. For instance, plant protein has a lower digestibility than animal protein. In other words, the digestion and absorption of plant protein is less efficient. In addition, individual variation in metabolism also influences the amount of protein absorbed from food. In general, the digestibility of mixed diets varies between 50 and 80% of the actual intake of protein, although the higher value of 80% is generally used in

calculations (WHO/FAO/UNU, 2007, pp. 7, 113, 220). The second factor to affect protein requirements is the protein to energy percentage of the diet. To be more precise, since every step in protein digestion and absorption is energy dependent, adequate non-protein energy (from carbohydrates or fat) is indispensable to fuel these energy demands, and consequently, to ensure that dietary amino acids meet the metabolic demand. Conversely, if protein intake is high but energy intake is low, protein itself may be used as a source of energy. This process, however, is highly inefficient since a large portion of the energy in protein is lost as heat during deamination. In general, the protein:energy percentage is higher when energy requirements are lowest. Accordingly, protein:energy percentages also vary according to the individual's gender, age, body size, body composition, and physical activity. For example, young children need food with a higher energy concentration than that needed by older women (WHO/FAO/UNU, 2007, pp. 79–88). Nevertheless, a protein:energy percentage of about 10-15% is considered sufficient to meet the protein demands of both female and male adults (WHO/FAO, 2003, p. 56). Third and last, protein requirements depend on the body's demand of indispensable amino acids, often grouped with other micronutrients. As mentioned before, the human body cannot produce certain amino acids (the indispensable amino acids), which must be obtained from the diet. However, different protein sources contain different amounts of each amino acid. This means that an individual may consume the required amount of protein without obtaining the required amount of amino acids, a situation that will still result in ill health (WHO/FAO/UNU, 2007, pp. 7, 150). Unfortunately, considering that semi-quantitative food frequency questionnaires are unsuitable for the estimation of micronutrient consumption (Gibson, 1990), I cannot ascertain whether the diets of the study populations meet the amino acid requirements. Nonetheless, it is important to be aware of their existence.

Having discussed the factors influencing protein demand, one may now address whether the study populations meet protein requirements. The best estimate of a population *average* protein requirement for healthy adults is the median requirement of 0.66 grams per kilogram of body weight per day (g/kg/day). However, the *safe* level,

which meets the requirements of most (97.5%) of the population, is set at 0.83 g/kg/day (WHO/FAO/UNU, 2007, pp. 38, 126). At this point, one must use the weight and height data collected by Jeffery (2018) to be able to estimate protein requirements of the El Molo and the Turkana. Assuming a diet digestibility of 80%, one may calculate an *average* and *safe* protein requirement for El Molo women of 40 and 51 g/day, respectively, of 46 and 58 g/day for El Molo men, of 39 and 49 g/day for Turkana women, and of 45 and 56 g/day for Turkana men. When comparing these values to those on Table 5.2, it becomes clear that (1) the protein consumption of El Molo men (60 g/day) and women (60 g/day) exceeds both the average and the safe protein requirement, (2) the protein consumption of Turkana men (49 g/day) meets the average but not the safe requirement, and (3) the protein consumption of Turkana women (32 g/day) meets neither. These results for Turkana protein consumption present a sharp decrease to those reported by Galvin (1985). In this study, twenty-four hour recalls estimated protein intakes of 61 g/day for women and 79 g/day for men, while weighed records estimated protein intakes of 81 g/day for men and 44 g/day for women (Galvin, 1985, pp. 193, 376). Surprisingly, Galvin (1985) does not offer a clear explanation for the reported difference between recalled and measured protein intake for women, albeit it may stem from a high variability in dietary intake from day to day (Galvin & Little, 1999). Nevertheless, the observed discrepancy in protein intake between Galvin's and the present research may result from flawed estimates in one or both studies, or from a recent change in Turkana diet. On the one hand, it is to be expected that if the semi-quantitative food frequency questionnaires underestimated energy consumption, through incorrect portion size estimation, then protein consumption must also be underestimated. In other words, it is possible that actual protein intake is higher for all populations than that reported here, albeit the level of error across populations will be stable. On the other hand, the Turkana are going through a major economic transition with significant changes in diet, resulting from a process of sedentarisation (Chapter 3, p. 59). Besides the progressive sedentarisation of the population, the Turkana received international food aid in the form of maize flour for over three decades, and maize flour (recently also rice) has become the key staple

(Ellis, Galvin, McCabe, & Swift, 1987). During interviews with older Turkana men and women carried out by Marta Mirazón Lahr (personal communication), all individuals recall that the diet of their childhood and youth consisted almost entirely of milk and blood, supplemented with berries, nuts, and tubers in different seasons of the year, consistent with early ethnographies. Present-day Turkana of the Kerio Valley (where this study was carried out) only rarely consume the traditional milk and blood mixture that used to constitute the bulk dietary source of energy.

Following this, and since no weight and height data are available for the other populations, it is not possible to estimate directly the protein requirement of all groups studied. However, since this number increases only about 4 g for each 5 kg gain in body weight (WHO/FAO/UNU, 2007, p. 243), and considering that none of the participants were visibly overweight, it is reasonable to assume that requirements calculated for the El Molo and Turkana could be applied to the remaining populations. Using this reasoning, one may conclude that the Luhya (Webuye) are unlikely to meet the safe level of protein intake, albeit women may meet the average requirement. Strangely, this is the only population where estimates indicate that men consume less protein (38 g/day) than women do (44 g/day). Otherwise, both genders from the Luhya (women: 53 g/day; men: 70 g/day) and the Luo (Port Victoria) (women: 54 g/day; men: 62 g/day) are likely to have a protein intake that exceeds both the average and the safe protein requirement.

After discussing raw protein requirements, one may also address whether these populations meet the protein:energy percentage of 10-15%. To achieve this, one must use a conversion factor of 4 kcal/g to transform protein in g/day into kcal/day (FAO, 2003, p. 57), and once again consider a diet digestibility of 80%. The results of these calculations (Table 5.3) reveal that both genders from all populations have an adequate protein:energy percentage above 10%. In fact, several groups have protein:energy percentages above the recommended 15%, namely the El Molo men and women, the Luhya (Port Victoria) men and women, and the Luo (Port Victoria) men. Interestingly, if one applies the same conversion factor to the Turkana protein consumption reported

by Galvin (1985), then those diets fall just above the recommended protein:energy percentage, with values around 16%. These values, however, still represent a shift towards less protein consumption among the Turkana, consistent with the recent socioeconomic changes under way in this group.

Taken together, the above results suggest that the three fisher populations, the El Molo, and the Luhya and Luo (Port Victoria), have higher protein intake than the other two populations. In addition, and considering the low energy content and the high protein:energy percentages of these groups' diets, it is plausible that these groups, particularly the El Molo, may be using protein as a source of energy. On the contrary, the Turkana and the Luhya (Webuye) have relatively low protein consumption, albeit it still meets the average requirement in some instances, and it always meets the necessary protein:energy percentage. The Turkana result is particularly unexpected considering the ethnographic characterization of these people as pastoralists, but as noted above, the Turkana studied no longer consume a typical pastoralist diet. Although it should be noted that there are remote Turkana communities, particularly in the mountains near the Ugandan border and in the Ilemi, who still practise a traditional nomadic pastoralist economy, including a traditional pastoralist diet based on milk and blood as staple.

5.3.3. FAT

The role of fat in the body is par excellence that of energy storage. Nonetheless, in many developing countries, where carbohydrate intake is inadequate, fat is an important source for energy production (FAO, 2010, p. 13). However, and in contrast to protein, FAO (2010) fails to report on a specific fat requirement. Instead, these organisations simply recommend a fat:energy percentage that populations must meet in order to maintain health. Similarly to protein, and because it is dependent on energy consumption, this percentage varies according to age, gender, body size, body composition, and physical activity. Nonetheless, a fat:energy percentage of 15% is considered as the minimum to ensure adequate consumption of total energy, essential

fatty acids, and fat soluble vitamins for most individuals. The maximum recommended fat:energy percentage is set at 35% (FAO, 2010, pp. 11, 14). In addition, FAO establishes distinct requirements for different types of fatty acids (i.e. saturated fatty acids, monounsaturated fatty acids, and polyunsaturated fatty acids), since these have unique biological properties and health effects (FAO, 2010, pp. 9, 14). However, as for indispensable amino acids, it was not possible to estimate the consumption of these micronutrients, for the same reasons.

In order to estimate the fat:energy percentage of the study populations, a conversion factor of 9 kcal/g was used to transform fat in g/day into kcal/day (FAO, 2003, p. 58). These calculations (Table 5.3) show that all groups meet the fat:energy percentage of 15%, with the exception of Turkana women, who have a fat:energy percentage of 11%. Furthermore, most of the groups are very close to the 15% minimum, and none comes close to the 35% maximum. Unfortunately, Galvin (1985) does not report on total fat (or carbohydrate) intake for the Turkana, and so it is not possible to ascertain if the low fat intake of Turkana women is a recent, or an historical phenomenon. To sum up, all groups in the present work have an adequate fat consumption, with the possible exception of Turkana women.

5.3.4. CARBOHYDRATE

The final macronutrient, carbohydrates, acts as an easily available energy for the oxidative metabolism, and thus, is the main source of energy in the diets of most people. In addition, carbohydrates are an important vehicle of micronutrients, and maintain homeostasis and gastrointestinal integrity and function (FAO/WHO, 1998).

Regarding carbohydrate intake recommendations, the amount required to avoid ketosis is low, around 50 g/day. Ketosis is a metabolic state in which some of the body's energy comes from keto acids, produced through fat and protein digestion. This condition, however, is not detrimental to health *per se*, simply signalling that the individual is not obtaining enough energy from carbohydrates, and so must resort to other sources. Other than this comment on ketosis, the FAO/WHO provides no official

recommendation on absolute carbohydrate intake. Instead, and in a similar way to fat, they report on the amount of energy in diet that should come from carbohydrates, a carbohydrate:energy percentage. According to this, in an optimum diet, at least 55% of total energy should come from carbohydrates, while consuming more than 75% of the total energy in the form of carbohydrates may have adverse effects on nutritional status, mainly through exclusion of adequate quantities of protein and fat. Otherwise, this report neglects to set any recommendation on the consumption of different types of carbohydrates (i.e. sugars, oligosaccharides, and polysaccharides), simply stating that a varied source of carbohydrates is desirable.

In the light of these recommendations, all groups in this study meet the minimum carbohydrate amount of 50 g/day (see Table 5.2). Furthermore, and using a conversion factor of 4 kcal/g to transform carbohydrate in g/day into kcal/day (FAO, 2003, p. 58), it is clear that all groups meet, or nearly meet, the minimum carbohydrate:energy percentage of 55%, and none surpasses the 75% maximum (see Table 5.3). The only group that falls very close to the minimum is the El Molo, which might be explained by the high protein intake. Overall, all participants in this study have an adequate carbohydrate consumption.

To sum up, data from semi-quantitative food questionnaires probably underestimate energy intake, as well as all other aspects of dietary intake. This is likely the result of incorrect estimation of portion size. Nevertheless, these questionnaires indicate that the populations have a low energy intake when compared to international standards, but still meet the minimum macronutrient requirements, particularly in terms of macronutrient proportion in the diet.

5.4. DIETARY COMPOSITION

Having described dietary intake across populations and determined the dietary adequacy to nutritional demands, one may now characterise these populations' diet in terms of the composition of macronutrients (i.e. protein, fat, carbohydrate) and food

types (e.g. animal, C_3 plant, etc.). In particular, this section looks at the *proportions* of a categorical variable (i.e. macronutrients or food categories) across another categorical variable (i.e. populations) in relation to different measures of diet. Finally, the last section adds another categorical variable, gender, in order to study potential sexual dimorphism in energy intake.

By looking into proportions, instead of raw values, one evaluates the relative contribution of a diet component to the overall diet. For instance, if two populations eat the same amount of C_4 plants, but in the first, this is diluted in a lower amount of C_3 plants than in the second, then the impact of the C_4 diet component to isotopic or microwear data is bigger for the first population than for the second. In addition, and in parallel to Table 5.2, all tables in this section report mean and standard deviation, as well as medians and interquartile ranges. Thus, data are described by both classic and robust measures of central tendency and variability. Notably, these data aim, above all, to highlight the variability within these estimates. Due to this variation in dietary intake within populations, no statistical analyses were conducted. Performing such tests would be statistically sound, but would have no biological meaning because the contribution of different macronutrients or food types to dietary intake varies widely between individuals of the same population. Therefore, simply as a way of characterizing diet at a population level, the dietary composition is presented graphically and then described in text.

5.4.1. CONTRIBUTION OF MACRONUTRIENTS TO POPULATION DIETARY INTAKE

To begin, one should look at how macronutrient proportions differ across populations. For instance, does the El Molo population consume proportionally more protein than the other populations? To see this, one may plot dietary intake in terms of weight or energy. Here, both cases are included. Figure 5.2 illustrates the data for dietary weight, which corresponds to the values from Table 5.2, for each macronutrient, for population overall. Notably, the plot uses the dry weight (i.e. the weight of protein, fat, and carbohydrates, excluding water, fibre, and others), but does not consider the 80%

digestibility for protein. Additionally, Figure 5.3 and Table 5.5 present the data for the intake of dry energy (i.e. the energy of protein, fat, and carbohydrates, excluding water, fibre, and others), but take into account the 80% of protein digestibility. Hence, Figure 5.2 illustrates the macronutrient intake entering the body, while Figure 5.3 illustrates the macronutrient intake reaching the body energy metabolism. Nonetheless, the relation between macronutrient proportions between populations is maintained from Figure 5.2 to Figure 5.3, since the latter is simply a transformation of the former obtained by applying an 80% digestibility to protein consumption, and by transforming g/day into kcal/day using Atwater conversion factors. Accordingly, from Figure 5.2 to Figure 5.3 the contribution of protein to dietary composition decreases (due to the 80% protein digestibility), the contribution from fat increases (due to the higher 9 kcal/g conversion factor), and the contribution from carbohydrates changes as a result of the other two macronutrients.

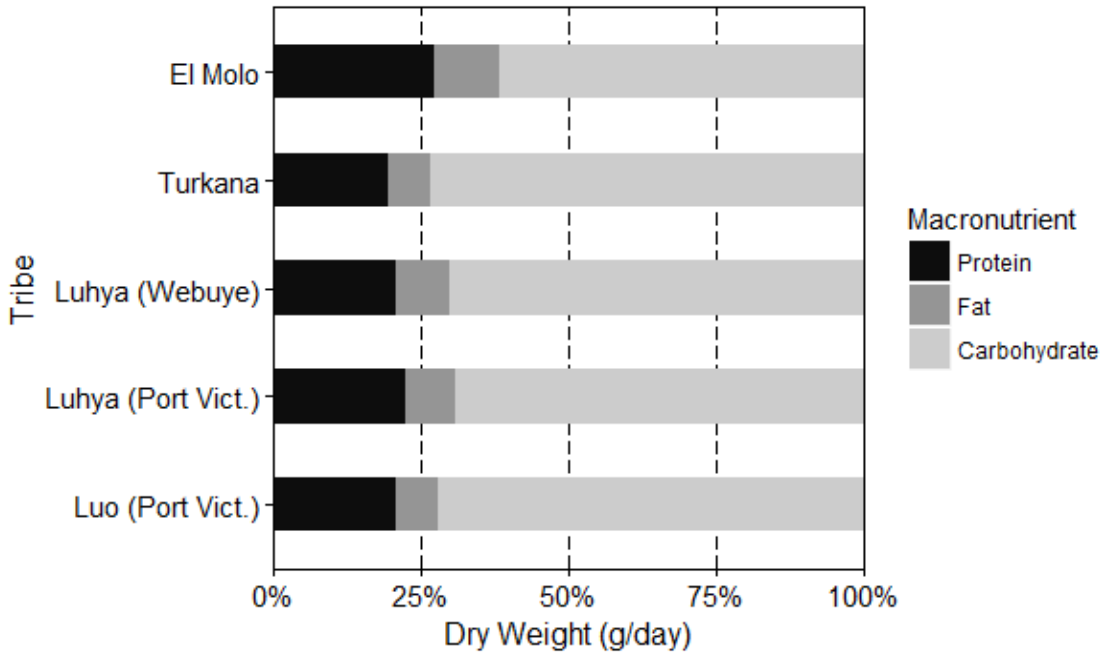


Figure 5.2 Macronutrient consumption expressed as percentage of the *dry weight* of consumed food, for all populations.

By looking at these two figures, one may see that the proportion of macronutrients in dietary intake, whether in relation to weight or to energy, does not differ widely across populations. In other words, the diet of all populations is consistent in macronutrient composition. If anything, perhaps the proportion of protein and fat is higher in the

El Molo. Therefore, and even though the El Molo, the Luhya (Port Victoria), and the Luo (Port Victoria) have a higher protein consumption in absolute terms (see before), this difference is not substantial when considering the contribution of protein to overall diet.

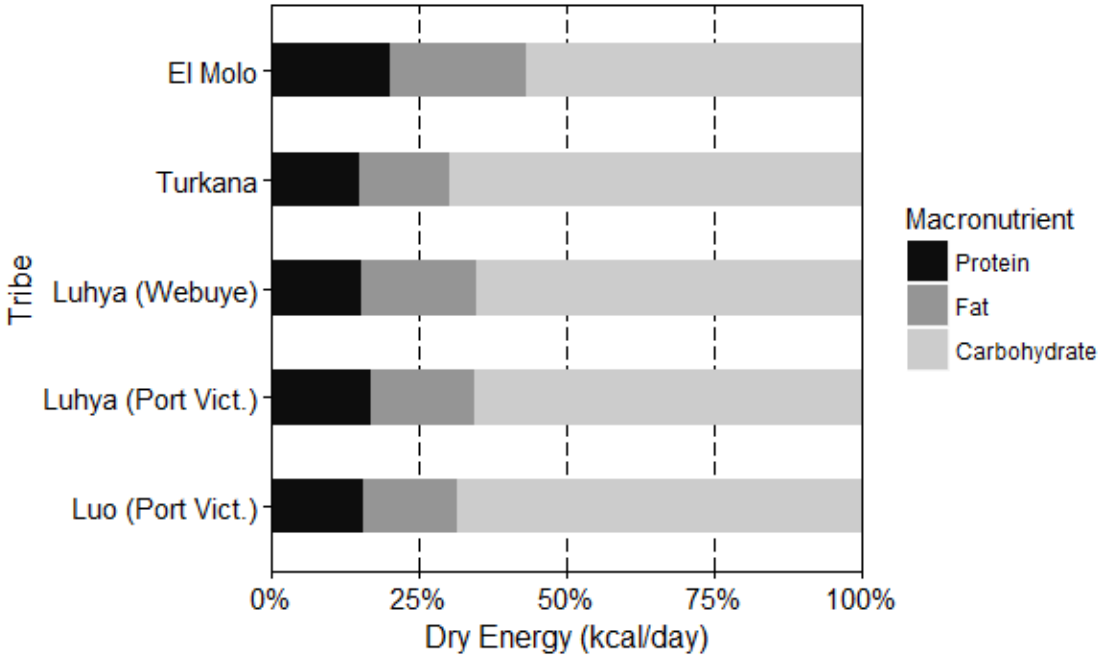


Figure 5.3 Macronutrient consumption expressed as percentage of the *dry energy* of consumed food, for all populations.

Table 5.5
Summary of dry energy intake according to macronutrient, across all populations.

		Macronutrient (kcal/day)					
		Protein		Fat		Carbohydrate	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	31	**193 (41)**	*201 (37)*	219 (64)	*198 (88)*	542 (147)	*537 (138)*
Turkana	37	**132 (91)**	*106 (68)*	**143 (151)**	*108 (71)*	608 (331)	*565 (439)*
Luhya (Webuye)	32	132 (47)	*127 (70)*	165 (63)	*161 (89)*	**554 (235)**	*533 (257)*
Luhya (Port Vict.)	31	194 (57)	*196 (82)*	202 (82)	*197 (127)*	748 (229)	*741 (314)*
Luo (Port Vict.)	29	185 (43)	*178 (41)*	185 (44)	*179 (68)*	803 (241)	*786 (371)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 330-331).

5.4.2. CONTRIBUTION OF FOOD CATEGORIES TO POPULATION DIETARY INTAKE

After looking at the macronutrient contribution to overall diet, one may study how diet partitions across food categories (i.e. animal, fish, C_3 plant, C_4 plant, and other). In

other words, which food types contribute the most to diet? For example, which populations eat relatively more animal or C_3 foods? However, before proceeding with this line of thought, one must keep two things in mind. First, and as acknowledged before, food categorisation is imperfect. In fact, there should be two categorization systems, one of C_3 versus C_4 plant foods, and one for plant versus animal versus fish foods. Yet, such approach is unfeasible, because it is difficult to ascertain what some animals eat, and thus, if they belong to the C_3 or C_4 category. Instead, it was opted to merge both classifications in a single system. Second, the only food items in the category 'Other' were honey and biscuits (Table 5.1), which made up only 1-2% of overall dietary intake. Considering this minimal contribution of the 'Other' category to diet, it was opted to exclude them from the graphical representations and interpretation.

Just as for macronutrient composition, one may look at the food composition in relation to different measures of diet. Figure 5.4 and Table 5.6 report on how wet weight (including water, ash, and others) distributes across food categories and populations. Following this, Figure 5.5 and Table 5.7 look at this same distribution using dry weight data (i.e. protein, fat, and carbohydrates only), while Figure 5.6 and Table 5.8 use wet energy data (i.e. the energy from the overall diet, as detailed in Table 5.2). Thus, Figure 5.4 illustrates the composition of foods entering the body; Figure 5.5 illustrates the composition of foods used by the energy metabolism; and Figure 5.6 illustrates the composition of foods entering the body in energy units. One could also look at this pattern in terms of dry energy (i.e. the energy from protein, fat, and carbohydrates). However, such plot would be very similar to the wet energy plot, due to the small differences between wet and dry energy, which vary from 25 kcal/day for the Luhya (Webuye) to 97 kcal/day for the Turkana.

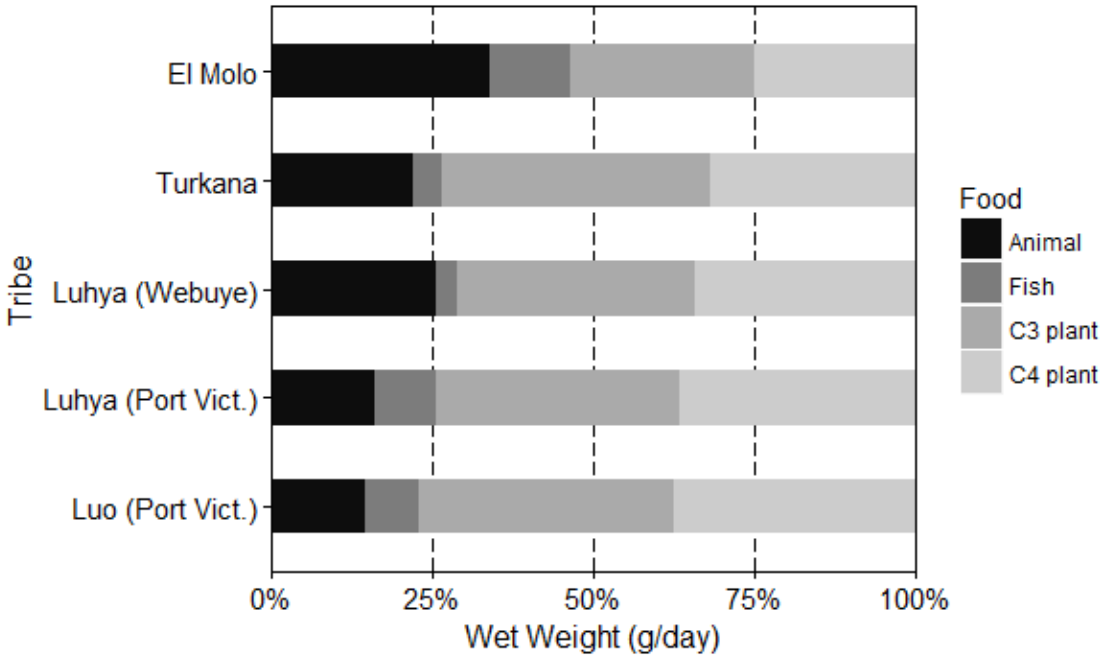


Figure 5.4 Food consumption expressed as percentage of the *wet weight* of consumed food. The 'Other' category was excluded, since its contribution to dietary intake was minimal, for all populations.

Table 5.6
Summary of wet weight intake according to food category, across all populations.

		Food Category (g/day)							
		Animal		Fish		C_3 plant		C_4 plant	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	31	**289**	*236*	**108**	*112*	243	*256*	214	*215*
		(151)	*(164)*	**(32)**	*(0)*	(79)	*(98)*	(84)	*(108)*
Turkana	37	**135**	*62*	**29**	*16*	254	*249*	**194**	*142*
		(138)	*(183)*	**(36)**	*(48)*	(148)	*(184)*	**(233)**	*(114)*
Luhya (Webuye)	32	**202**	*215*	**26**	*32*	288	*241*	267	*269*
		(76)	*(41)*	**(18)**	*(16)*	(136)	*(205)*	(95)	*(148)*
Luhya (Port Vict.)	31	**167**	*157*	**100**	*112*	391	*377*	**377**	*331*
		(101)	*(192)*	**(34)**	*(0)*	(177)	*(179)*	**(151)**	*(289)*
Luo (Port Vict.)	29	167	*188*	**94**	*112*	445	*458*	424	*455*
		(72)	*(131)*	**(39)**	*(0)*	(155)	*(245)*	(175)	*(309)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 332).

From Figure 5.4, several points become clear. To begin, food of plant origin is an important part of all the populations' diet. To be more precise, combined C_3 and C_4

foods make up close to 75% of all diets, apart from the El Molo, although among this group it is still above 50%. Between these two categories, the C_4 proportion is lower among the El Molo (around 25%), in comparison to the other populations. In parallel, C_3 foods contribute less to the El Molo diet than to others' diet. Excluding the El Molo, the C_4 proportion across the remaining populations looks relatively constant, while the C_3 proportion is lower for the Luhya (Webuye) than for the other populations. On fish consumption, the three fisher populations clearly consume more of this staple than the non-fishers do. However, for all populations, the proportion of fish is always smaller than the proportion of animal food, which may be explained by high consumption of milk. This food is the highest contributor of animal foods for all populations overall. In detail, the average consumption of milk (in wet weight) is 137 g/day for all populations (with eggs a distant second at 15 g/day). In addition, the average of milk consumption for El Molo is 255 g/day, followed by Luhya (Webuye) at 156 g/day, then Luo (Port Victoria) at 118 g/day, then Luhya (Port Victoria) at 104 g/day, and finally Turkana at 63 g/day. This seems to explain the pattern of animal food consumption seen in Figure 5.4, wherein the El Molo present the highest proportion, followed by the Luhya (Webuye), then the Luo and the Luhya (Port Victoria) with a similar proportion, and finally, the Turkana. Notwithstanding, the low milk consumption detected among the Turkana may be partly inaccurate. Although this population reported to consume milk less often than the others did, it is likely that they consumed bigger quantities when they had access to it. Thus, the estimated low milk consumption in the Turkana may partly result from incorrect estimation of portion sizes. As a final point on Figure 5.4, the food composition of the two populations from Port Victoria is very similar, pointing to an identical diet.

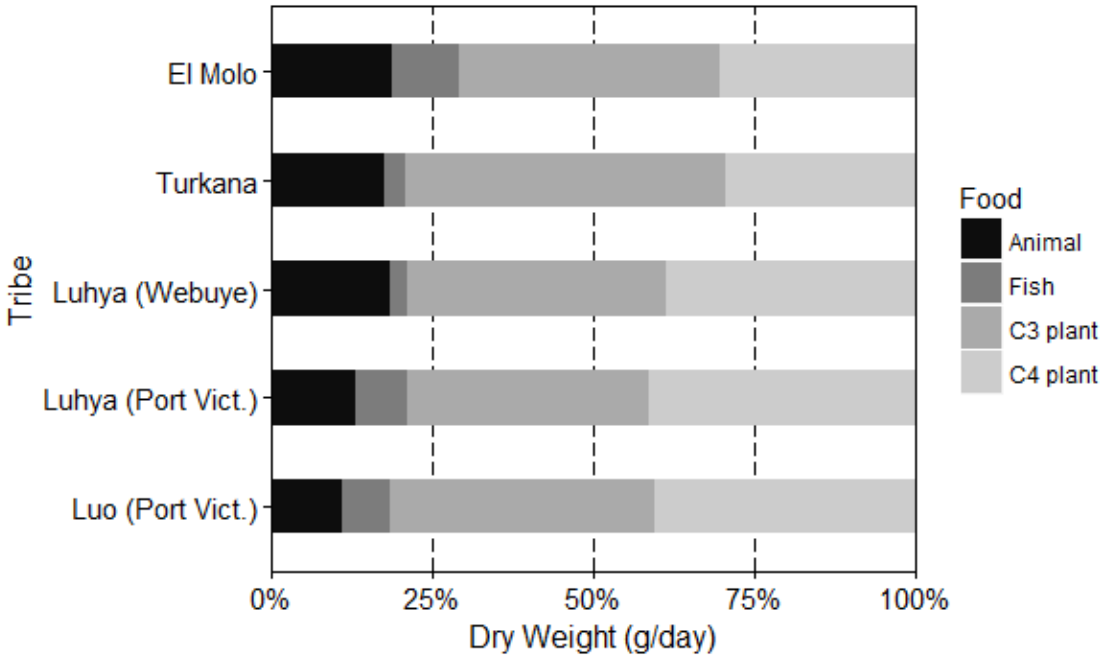


Figure 5.5 Food consumption expressed as percentage of the *dry weight* of consumed food.

Table 5.7
Summary of dry weight intake according to food category, across all populations.

		Food Category (g/day)							
		Animal		Fish		C_3 plant		C_4 plant	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	31	42 (19)	*38 (18)*	**23 (7)**	*24 (0)*	89 (32)	*81 (47)*	67 (25)	*66 (28)*
Turkana	37	**37 (39)**	*22 (36)*	**7 (12)**	*3 (10)*	104 (62)	*117 (93)*	**62 (42)**	*54 (53)*
Luhya (Webuye)	32	36 (16)	*34 (19)*	**5 (4)**	*7 (3)*	79 (40)	*68 (54)*	**76 (42)**	*68 (40)*
Luhya (Port Vict.)	31	35 (21)	*34 (35)*	**21 (7)**	*24 (0)*	100 (51)	*94 (68)*	**111 (40)**	*122 (78)*
Luo (Port Vict.)	29	31 (10)	*30 (13)*	**20 (8)**	*24 (0)*	115 (44)	*110 (61)*	113 (47)	*107 (59)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 333-334).

Turning to the food composition of dry weight (Figure 5.5), the same patterns observed in wet weight remain. However, there is one important difference: the contribution of combined animal and fish foods decreases considerably, while the contribution of

plant foods to diet increases. Once again, the milk in these populations' diet might explain this result. Milk has high water content and low macronutrient content (see Table 5.1), and so its contribution to dry weight is smaller than to wet weight. Despite this, it is likely that this food staple still contributes considerably to dry weight composition, considering the big gap between milk consumption per day and the closest animal food, eggs (see above). This could explain why the animal food share remains always higher than the fish share (even in fisher populations). An alternative explanation is interviewer bias. Since I, as the interviewer, was unfamiliar with the different types of fish consumed around Lake Turkana and Lake Victoria, participants were only asked about fish consumption once (Appendix A, Figure A.3), whereas they were questioned on many animal foods. Thus, even if participants reported eating a lot of fish, this might have been obscured by the added effects of being questioned on a long list of animal foods.

Regarding the plant part of the diet, it is important to realise that, for all populations, plant foods make up such an important part of the food reaching the body energy metabolism (close to 75% of the dry weight for the El Molo, and above for the remaining). A final point of note is that, in comparison to wet weight, the contribution of C_4 to dry weight increases across all populations, except the Turkana, where it decreases, while the C_3 contribution likely changes mostly because of the other components of diet. Thus, the contribution of C_4 to diet is lower, but similar, in the El Molo and Turkana, and higher in the Luhya (Webuye), and the Luhya and Luo (Port Victoria). It is not entirely clear what drives the decrease in the Turkana consumption of C_4 foods, although it may be explained by the high contribution of sweets and soda to the Turkana diet. Among the Turkana, sweets and soda constitute a crucial part of the C_4 consumption, while *githeri* and *ugali* are at least equally important sources among other populations. In detail, the Turkana consume on average 73 g/day of sweets and soda in terms of wet weight, but only 22 g/day in dry weight. Clearly, this results from the high water content of soda, and so, when analysing dry weight, the contribution of C_4 to the Turkana diet decreases visibly.

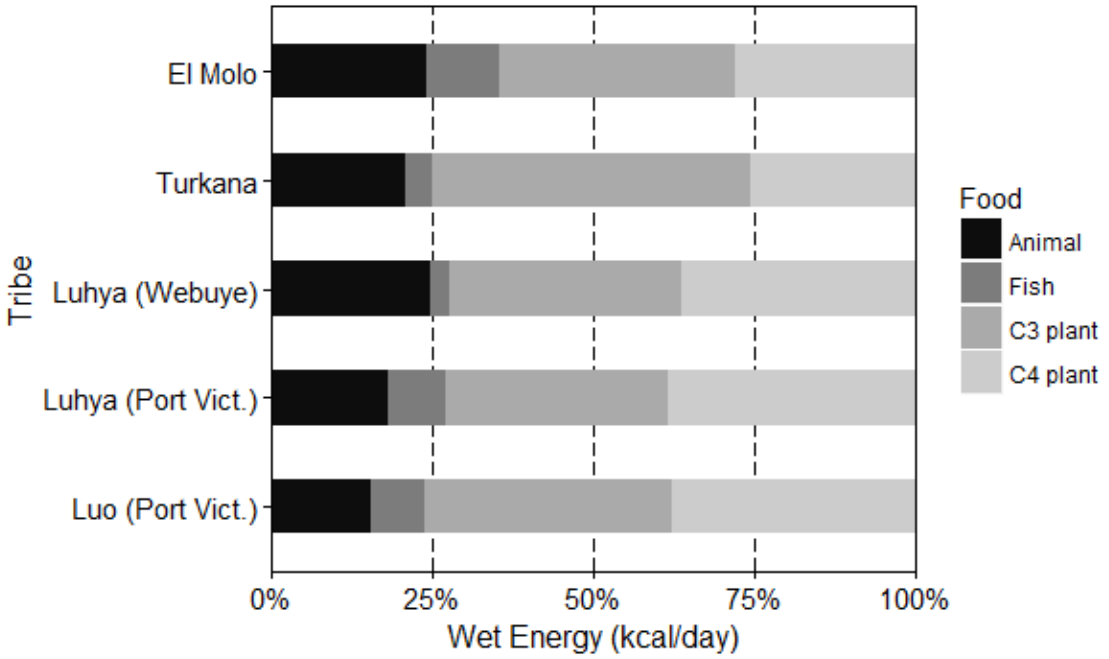


Figure 5.6 Food consumption expressed as percentage of the *wet energy* of consumed food, for all populations.

Table 5.8
Summary of wet energy intake according to food category, across all populations.

		Food Category (kcal/day)							
		Animal		Fish		C_3 plant		C_4 plant	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	31	241 (110)	*220 (101)*	**113 (34)**	*118 (0)*	376 (138)	*327 (193)*	278 (104)	*280 (119)*
Turkana	37	**205 (211)**	*136 (168)*	**39 (78)**	*17 (50)*	**486 (380)**	*521 (419)*	**250 (167)**	*209 (201)*
Luhya (Webuye)	32	217 (98)	*202 (121)*	**27 (19)**	*34 (17)*	314 (159)	*270 (220)*	**317 (166)**	*281 (167)*
Luhya (Port Vict.)	31	213 (122)	*204 (212)*	**105 (36)**	*118 (0)*	400 (204)	*387 (272)*	**448 (157)**	*485 (301)*
Luo (Port Vict.)	29	186 (64)	*176 (82)*	**99 (42)**	*118 (0)*	459 (180)	*436 (239)*	454 (184)	*423 (225)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 334-335).

Finally, Figure 5.6 expresses the distribution of food categories in terms of wet energy (i.e. the energy obtained from the whole diet, including macronutrients, but also other

food components). This figure is virtually the same as the above Figure 5.5 for dry weight, which confirms that the bulk of the diet energy comes from macronutrients: protein, fat, and carbohydrate.

Summing up, the above analysis informs that (1) plant foods form the bulk of the diet for all populations; (2) the El Molo and the Turkana have a lower C_4 diet component when compared to the other populations; (3) the El Molo, together with the Luo and the Luhya (Port Victoria), consume more fish than the other two populations; (4) the proportion of animal foods in the diet is always higher than the proportion of fish, which might be a combined result of high milk consumption and interviewer bias; (5) the El Molo and the Luhya (Webuye) have the highest proportion of animal foods, followed by the Turkana, and then by the Luo and Luhya (Port Victoria); and (6) the Luo and the Luhya (Port Victoria) have very similar dietary composition.

5.4.3. CONTRIBUTION OF FOOD CATEGORIES TO POPULATION PROTEIN INTAKE

The previous section considers how different food categories contribute to different measurements of overall diet, namely wet and dry weight, and wet energy. Subsequently, this section considers how these food categories contribute to just the protein part of the diet (i.e. the protein part of dry weight). This approach is particularly pertinent for the interpretation of isotopic results later on, since the carbon in the body is partly affected by protein routing, while the nitrogen hails almost exclusively from the protein part of the diet (Chapter 2, p. 48). Thus, Figure 5.7 and Table 5.9 summarise how food categories contribute to population protein intake.

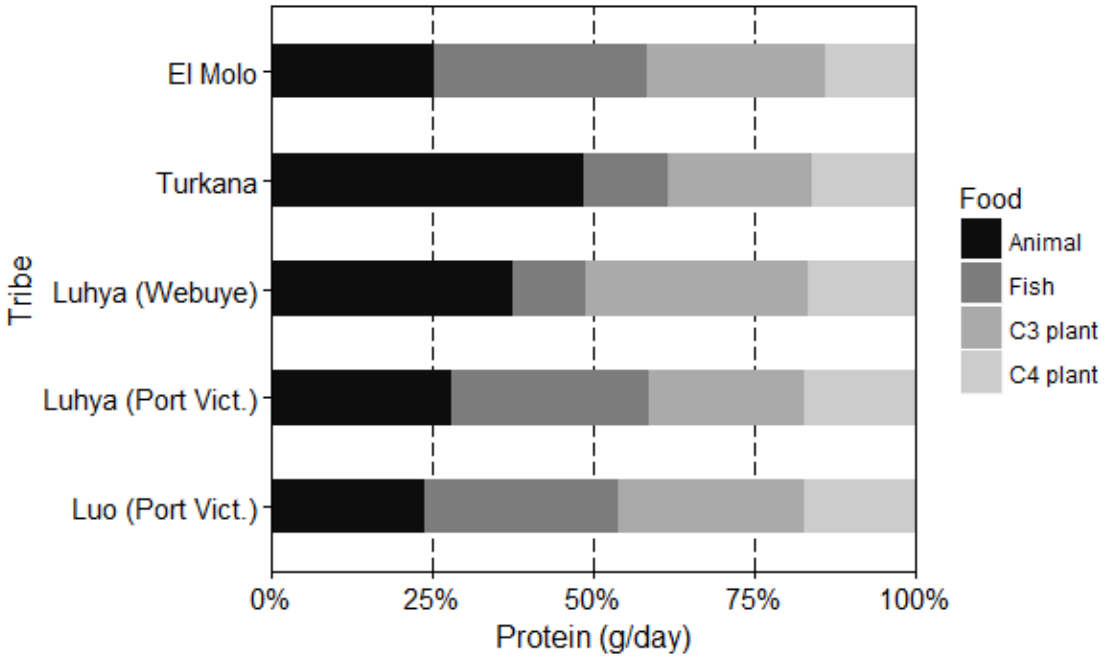


Figure 5.7 Food consumption expressed as percentage of the *protein* of consumed food, for all populations.

Table 5.9
Summary of protein intake according to food category, across all populations.

		Food Category (g/day)							
		Animal		Fish		C_3 plant		C_4 plant	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	31	15 (7)	*15 (6)*	**20 (6)**	*21 (0)*	**17 (6)**	*18 (4)*	**8 (3)**	*9 (3)*
Turkana	37	**20 (22)**	*12 (18)*	**6 (7)**	*3 (9)*	**9 (72)**	*7 (10)*	**7 (4)**	*5 (7)*
Luhya (Webuye)	32	15 (8)	*14 (11)*	**5 (3)**	*6 (3)*	**14 (8)**	*12 (9)*	**7 (3)**	*6 (7)*
Luhya (Port Vict.)	31	17 (10)	*16 (18)*	**18 (6)**	*21 (0)*	14 (8)	*15 (13)*	10 (5)	*11 (8)*
Luo (Port Vict.)	29	14 (5)	*13 (8)*	**18 (7)**	*21 (0)*	17 (8)	*17 (14)*	**10 (5)**	*11 (6)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 336).

These results reveal several points. First, foods of plant origin contribute less to protein than to overall diet. In fact, combined animal and fish foods make up more than 50% of protein intake for most populations (the exception are the Luhya (Webuye), among

whom it is nearly 50%). This result is expected, considering that plant foods have lower protein content than either animal or fish foods (see Table 5.1). Second, within the plant part of protein intake, the C_4 contribution is stable across all populations, while the C_3 contribution varies more. In detail, the Luhya (Webuye) have the highest proportion of C_3, followed by the El Molo and the Luo (Port Victoria), and then by the Turkana and the Luhya (Port Victoria). Third, the contribution of fish to protein intake among the fisher populations is more than double than that of the Turkana or the Luhya (Webuye). Fourth, the fish proportion in the fisher populations matches or surpasses the animal proportion, despite the (above mentioned) high milk consumption or interviewer bias. Finally, the Turkana have the highest proportion of animal protein, followed by the Luhya (Webuye), then the Luhya (Port Victoria), and lastly the El Molo and the Luo (Port Victoria). Furthermore, there are minimal changes in the protein distribution of food categories after removing Turkana individuals with high protein intake (those identified as outliers on Figure 5.1.b) (Appendix B, p. 337). Clearly, a consumption of animal foods with high protein content by the Turkana must explain this higher contribution of animal foods to protein intake in this group (when compared to measures of overall diet).

In brief, Figure 5.7 shows that (1) the contribution of plant protein to protein intake is higher among the Luhya (Webuye), than among the other populations; (2) fish protein contributes considerably to protein intake among the El Molo, the Luhya and the Luo (Port Victoria); and (3) the Turkana consume more animal protein than the remaining populations.

5.4.4. COMPARISON OF FOOD CATEGORIES ACROSS GENDER

At this point, gender is the last variable to consider in relation to the distribution of food types.

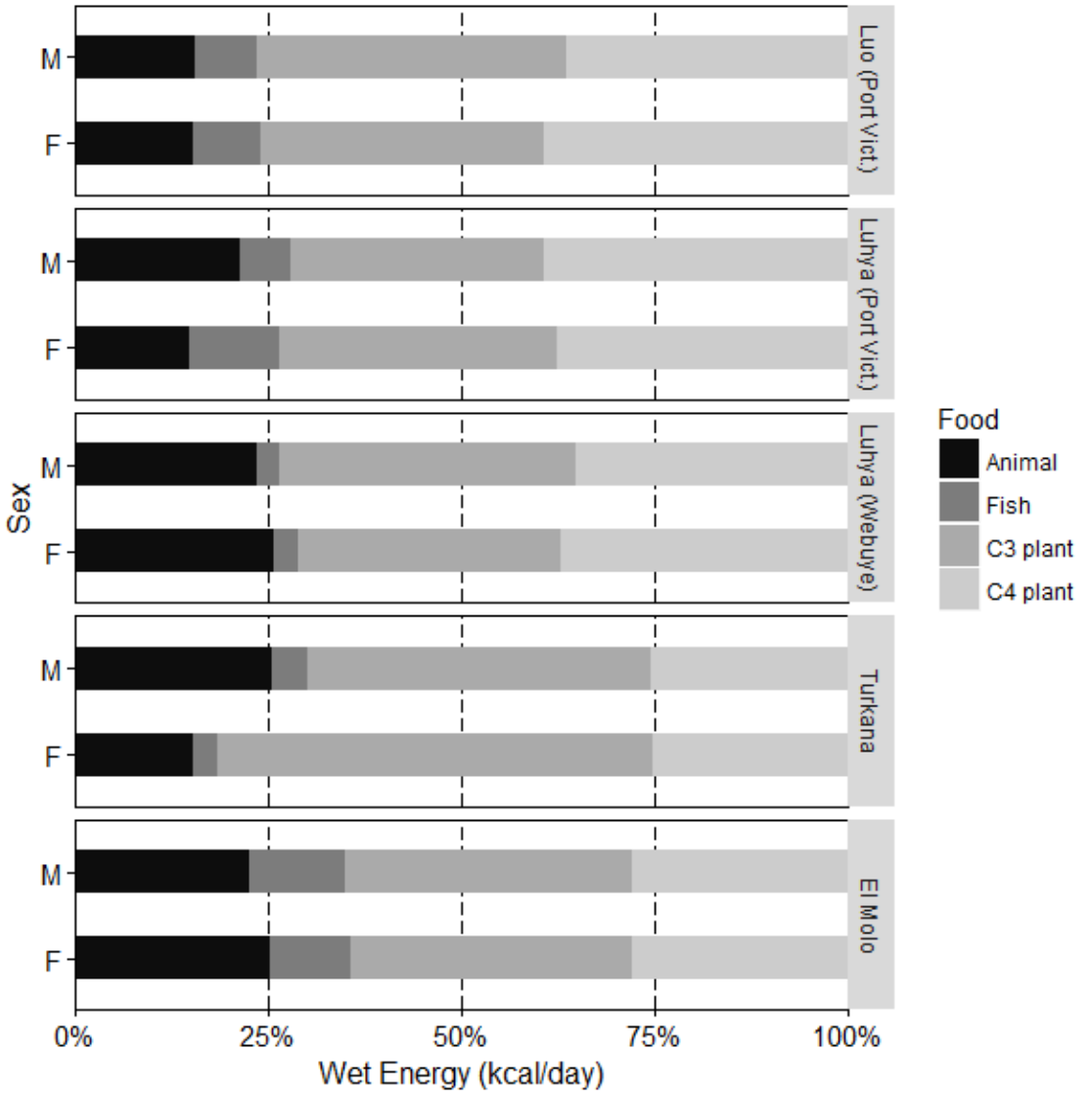


Figure 5.8 Food consumption expressed as percentage of the *wet energy* of consumed food, for all populations, but separated by gender. M = male, F = female.

Considering that energy requirements differ between the two genders (WHO/FAO/UNU, 2004), Figure 5.8 and Table 5.10 look at how wet energy distributes across food categories in each gender within each population. Conversely, the equivalent for macronutrient composition is not included because these data are presented on Table 5.3 and discussed throughout the section on dietary adequacy.

From Figure 5.8, it is possible to observe some differences between the two genders. For instance, Turkana men, in comparison to women, obtain more energy from animal foods, but less from C_3 foods. In a similar way, the men from Luhya (Port Victoria) obtain more energy from animal foods than the women do, while the women obtain

more energy from fish foods than the men do. Otherwise, differences between the genders for the other populations are too subtle to justify mention.

Table 5.10
Summary of wet energy intake according to food category, for all populations, by gender.

		Food Category (kcal/day)							
		Animal		Fish		C_3 plant		C_4 plant	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo									
male	14	225	*218*	**124**	*118*	369	*300*	277	*280*
		(95)	*(80)*	**(33)**	*(0)*	(190)	*(304)*	(118)	*(113)*
female	17	254	*233*	**104**	*118*	365	*367*	280	*300*
		(121)	*(115)*	**(32)**	*(0)*	(81)	*(104)*	(96)	*(304)*
Turkana									
male	18	**284**	*177*	**53**	*6*	492	*578*	285	*268*
		(262)	*(238)*	**(108)**	*(63)*	(287)	*(285)*	(183)	*(190)*
female	19	**131**	*87*	**26**	*17*	**480**	*333*	**216**	*168*
		(111)	*(140)*	**(29)**	*(32)*	**(458)**	*(409)*	**(148)**	*(214)*
Luhya (Webuye)									
male	16	192	*184*	24	*25*	311	*276*	287	*301*
		(117)	*(193)*	(17)	*(27)*	(149)	*(197)*	(91)	*(166)*
female	16	241	*234*	**30**	*34*	317	*245*	**348**	*276*
		(70)	*(102)*	**(20)**	*(17)*	(174)	*(234)*	**(217)**	*(176)*
Luhya (Port Vict.)									
male	14	289	*321*	90	*118*	441	*463*	**530**	*584*
		(108)	*(153)*	(50)	*(25)*	(211)	*(248)*	**(143)**	*(150)*
female	17	**150**	*122*	118	*118*	366	*353*	382	*361*
		(96)	*(177)*	(0)	*(0)*	(198)	*(190)*	(138)	*(230)*
Luo (Port Vict.)									
male	14	201	*199*	**102**	*118*	513	*532*	469	*414*
		(58)	*(93)*	**(39)**	*(0)*	(165)	*(155)*	(203)	*(275)*
female	15	171	*155*	**96**	*118*	408	*338*	440	*423*
		(67)	*(55)*	**(45)**	*(0)*	(184)	*(301)*	(172)	*(125)*

Note. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix B, p. 338-340).

5.5. DIET ASSESSMENT DISCUSSION

Throughout the above sections of the present chapter, the diet of the populations of this study were characterised and described using data obtained through semi-quantitative food questionnaires. At this point, it is pertinent to frame these results within the literature and particularly within the diet described ethnographically in Chapter 3. In this way, it will be possible to detect any recent changes in subsistence strategies and to refine further the diet characterizations. In particular, these questionnaires permitted to quantify the diet of the populations, which will prove useful to the interpretation of other data collected in this study.

In the context of energy and macronutrient intake, one must keep in mind that, at an individual level, there is a very high variability in food intake within the same population. Notwithstanding, at a population level, the diet of these groups may be characterised as containing low amounts of energy according to international standards, but meeting nutritional demands in terms of macronutrient composition. On one hand, this finding results in part from underestimated energy intake from diet questionnaires, due to incorrect estimation of portion size. On the other hand, Jeffery (2018) found that both the El Molo and the Turkana suffer from low levels of acute malnutrition, indicating inadequate energy intake in the short term. However, the Turkana have higher levels of chronic malnutrition than the El Molo do, as suggested by the rates of stunting, which implies low levels of energy intake in the long term. This is also consistent with the lower levels of energy intake found amongst this group. Unfortunately, there are no equivalent data available for the remaining populations of this study. Nonetheless, it is reasonable to posit that the Luhya (Webuye) suffer from similar levels of malnutrition (both acute and chronic) to those of the Turkana, while the populations from Port Victoria should have the lowest levels of malnutrition. Despite this, to ensure survival, energy intake must equal energy expenditure in the long term. This may be achieved by occasional high food intake, as already mentioned, but also by low levels of physical activity, which is consistent with

energy expenditure measurements in Galvin (1985), or by physiological adaptations to low energy intake, such as decreased basal metabolic rate (Waterlow, 1986).

Turning to the specifics of each population's diet, the ethnographic diet of the El Molo is similar to that detected through the diet questionnaires. Foremost, fish remains an important staple of the El Molo diet. Conversely, the questionnaires indicate an increase in the intake of animal foods. Although this finding probably results from the combination of high milk consumption and interviewer bias, it still reflects a recent greater assimilation into a wider Samburu community since traditionally the El Molo did not keep livestock and had no access to milk. Otherwise, in terms of wet weight, the main plant foods of the El Molo diet remain the same, i.e. beans (~200 g/day), maize (112 g/day), and rice (43 g/day), whereas the consumption of fruit and fresh vegetables is still low (~6 g/day) (Chapter 3, p. 59).

On the contrary, the diet assessment of the Turkana is in accordance with the fact that these individuals have become sedentary. Most strikingly, they have the lowest milk consumption of all populations (63 g/day), while previous work recorded that, when averaged across seasons, this staple provided at least 50% of Turkana intake of energy, and 70% of the intake of protein (Galvin, 1985, pp. 159, 173). In spite of this, the Turkana continue to be (among the populations of the present study) the ones with the highest proportion of animal protein in the diet, although they do not eat comparatively more protein. This finding likely results from the low total amount of food consumed by the Turkana, but also from the inclusion of many fisher groups in this study, which leads to low overall consumption of animal foods. Other important contributors to the wet weight of the Turkana diet continue to be maize (72 g/day), beans (62 g/day), doum palm fruit (58 g/day), and rice (58 g/day) (Chapter 3, p. 59). Therefore, these results indicate that, as the Turkana settle, they replace animal food by plant food. Consistent with this finding, the settled Ngilukumong and Ngiyapakuno along the Tarach River (different Turkana sections to the Ngisonyoka of this study) consume at least 45% of their food in the form of famine relief maize (Campbell et al., 1999; Ellis et al., 1987). Furthermore, the two genders of this group follow slightly

different diets, as indicated by the different proportion of animal food between the two. This is supported by Galvin (1985), who found that men obtained 73% of their energy from animal foods, while women obtained 59%. In other words, even though there has been a marked decrease in food of animal origin in the Turkana diet, the pattern of food distribution between the genders remains. Finally, the Turkana have a higher variability in dietary intake within the population than the other groups (see Table 5.2 and Figure 5.1). This may represent two phenomena, not mutually exclusive, (1) the inclusion of true pastoralists in the sample as they migrate into the area, particularly in view of the identified outliers in protein and fat consumption (Figure 5.1), and (2) differing levels of access to food across the population.

Regarding the Luhya (Webuye), this group has a diet with (1) a low fish proportion (just as the Turkana), (2) the second highest proportion of animal protein (after the Turkana), and (3) a higher C_4 proportion than the two previous group, but lower than the ones from Port Victoria. Moreover, in terms of wet weight, the main animal staples of the Luhya (Webuye) diet are cattle (20 g/day), eggs (13 g/day), and poultry (9 g/day), whereas the main plant staples are maize (149 g/day), beans (111 g/day), fruit (65 g/day), and cowpeas (60 g/day). This corresponds with the ethnographic record of this population, which characterises it as agriculturalist (Chapter 3, p. 70).

Subsequently, the Luhya and the Luo from Port Victoria follow similar diets, with the exception that the two genders of the Luhya differ in animal and fish protein intake, while the Luo do not. In detail, Luhya men have a higher intake of animal protein in relation to fish protein, whereas the women have the opposite pattern. Taken together, this indicates that location (Port Victoria and Webuye) is more important in determining the diet of a population than the ethnic affiliation (Luhya and Luo), although this variable may still play a role. In relation to the previous groups, the diet of the people from Port Victoria have (1) a high fish proportion, just like the El Molo, (2) a lower animal fraction, and (3) a higher C_4 fraction in relation to the other groups. Once again, the diet assessment endorses the diet described ethnographically. In particular, fish makes up an important part of the diet (Luhya: 100 g/day;

Luo: 94 g/day), while eggs (Luhya: 24 g/day; Luo: 25 g/day) and cattle (Luhya: 14 g/day; Luo: 10 g/day) are the most common animal foods. In addition, the main plant staples are maize (Luhya: 142 g/day; Luo: 140 g/day), beans (Luhya: 83 g/day; Luo: 79 g/day), and rice (Luhya: 71 g/day; Luo: 59 g/day).

Last, but not least, one must address the diet of the Baka. Considering that Dr. Ramírez-Rozzi was unable to conduct diet questionnaires among this population, it is only possible to discuss the Baka diet through other studies conducted on the subject. However, as pointed at the beginning of this chapter, data from other studies will not be fully comparable to the data from the present study due to differences in the data collection method.

To corroborate this idea that distinct methods provide different diet assessments, two studies among the Baka report different energy and protein intake (Sato et al., 2012; Yamauchi et al., 2000). Both studies were conducted in the same village of Ndgongo in south-eastern Cameroon, and both studies collected weighed records from two foraging trips in the forest. However, the foraging trips in Sato et al. (2012) lasted for 21 days each, while those in Yamauchi et al. (2000) lasted for 3 days each. Furthermore, the first study counted 16 participants on the first trip and 23 on the second trip, whereas the second study counted only four people on both trips. Thus, Sato et al. (2012) estimated an energy intake between 2500 and 2800 kcal/day, while Yamauchi et al. (2000) estimated an energy intake around 1700 kcal/day. Moreover, the former study determined a protein intake around 120 g/day, against the 50 g/day determined by the latter. Notwithstanding these considerable discrepancies, both studies agree that most of the energy (between 70 to 85%) in the Baka diet comes from starch and cereals (mostly yam tubers), followed by meat and fish (8-15%), and finally by fruits, vegetables, and other sources (7-15%). Equally important, both studies estimate similar contributions of protein of animal origin to total protein intake: 60% in Sato et al. (2012), and 52% in Yamauchi et al. (2000). Additionally, Sato et al. (2012) go on to estimate a 5% contribution of fish to total protein.

When comparing the Baka diet described above with the other populations in this study, the Baka have higher energy intake than all Kenyan populations in this study. However, weighed records usually estimate higher energy intakes than food frequency diet questionnaires or twenty four hour recalls (see the example of Galvin's study, p. 118), and it is possible that the Sato et al. (2012) study has inflated energy results, considering that the energy expenditure they estimate is relatively low for the high estimated energy intake, and that this population is not overweight, although they are certainly not malnourished. Nevertheless, it would be safe to assume that the Baka reach at least the same level of energy intake as the groups from Port Victoria. Furthermore, animal and fish do not contribute considerably to dietary intake. Instead, C_3 foods make up most of the Baka diet, with nearly no contribution from C_4 foods. In fact, the Baka should have the lowest intake of C_4 foods, considering that famine relief maize is not an important contributor to the Baka diet (Sato et al., 2012; Yamauchi et al., 2000). Finally, animal foods may contribute to the protein intake of the Baka as much as to the protein intake of the Turkana.

5.6. SUMMARY OF CHAPTER 5

To conclude this chapter, semi-quantitative food frequency questionnaires on the populations of this study inform that:

- The diet varies considerably between individuals of the same group;
- Nevertheless, in general terms, these groups have low energy intake but adequate macronutrient intake;
- They also have a similar diet in terms of macronutrient proportions;
- Table 5.11 summarises the proportion of food types in overall diet;

Table 5.11
Summary of proportions of food types in overall diet.

Diet	Animal	Fish	C_3 plant	C_4 plant
Higher	Luhya (Webuye)	El Molo	Turkana	Luhya and Luo (Port Vict.)
↑	El Molo	Luhya (Port Vict.)	Luo (Port Vict.)	
	Turkana	Luo (Port Vict.)	Luhya (Webuye)	Luhya (Webuye)
	Luhya (Port Vict.)	Turkana		
	Luo (Port Vict.)	Luhya (Webuye)	Luhya (Port Vict.) and El Molo	El Molo and Turkana
Lower				

  - the El Molo and the Luhya (Webuye) have the highest proportion of animal foods in respect to overall diet, but the Turkana obtain more *protein* from animal sources than the other populations;
  - plant foods form the bulk of the diet in all populations;
  - the Turkana and the Luhya (Port Victoria) are the only groups that differ between the genders in dietary intake. In detail, Turkana and Luhya (Port Victoria) men obtain more energy from animal foods than women do, whereas the women compensate this difference by consuming more C_3 foods and fish, respectively;

- The Baka may have a higher energy intake than the Kenyan populations, a protein intake of animal origin similar to the Turkana, and the lowest C_4 proportion in the diet, from all populations;

Overall, the diet questionnaires corroborate the diet described in the ethnography and in recent studies, apart from the Turkana, who seem to be changing into a sedentary diet, poor in animal foods and rich in C_4 foods, particularly maize.

CHAPTER 6

DENTAL MICROWEAR TEXTURE ANALYSES

RESULTS AND DISCUSSION

This chapter addresses the results of the *dental microwear texture analyses* (abbreviated to DMTA) conducted in this study. This technique is among the most effective ways of inferring past diets (Ungar et al., 2008). However, successfully studying the microwear patterns of living individuals has a few problems (e.g. the presence of biofilm), which are addressed in this chapter. Despite this, an attempt is made to compare populations across DMTA parameters, and to compare these results with previous studies. In addition, this chapter explores how the DMTA method may be improved in future work.

6.1. DMTA RESULTS

In Chapter 2 (p. 22, p. 27), it became clear that dental microwear studies are a direct means to study the diet of the individual during its lifetime. Indeed, microwear informs on the material properties of the food consumed, within a hard versus tough food continuum. In addition, DMTA offers a solution to the high intra-observer error found in more traditional microwear studies. Ultimately, a successful DMTA produces a data cloud representing the occlusal tooth surface, which, through scale sensitive fractal analyses, is then characterised by a few variables (*Asfc*, *Smc*, *epLsar*, *HAsfc*, *Tfv*). For ease of understanding, Table 2.2 (p. 34) is reproduced here in Table 6.1, which defines and gives examples for each DMTA variable.

However, and as described in detail in Chapter 4 (pp. 96-100), the moulding, casting, and scanning of teeth in living subjects for the purpose of DMTA is a complex process, involving many steps. Indeed, several obstacles may impede one from obtaining a usable microwear surface, such as the one depicted in Figure 6.1.

Before addressing these issues, note that no dental moulds were collected from the Baka, since another research group had already conducted those analyses (Romero et al., 2013). Instead, this chapter addresses the DMTA results of the El Molo, Turkana, Luhya (Webuye), Luhya (Port Vict.), and Luo (Port Vict.). However, Romero et al. (2013) used SEM to study the buccal microwear of the Baka, and hence, its results will not be fully comparable to those of the present study.

Table 6.1

Scale sensitive fractal analysis parameters used to describe a microwear surface. Adapted from Goodall, Darras, & Purnell (2015), Krueger et al. (2008), and Scott et al. (2005).

Parameter Name	Description	Example	Diet
Area scale fractal complexity (*Asfc*)	Measure of change in roughness with scale	A complex surface may have pits and scratches of different sizes overlaying one another	A complex surface indicates a hard diet
Scale of maximum complexity (*Smc*)	Scale at which the surface is most complex	Coarse features (large pits, no fine scratches) may dominate a surface with high *Smc*	High *Smc* indicates a hard diet
Exact proportion Length scale anisotropy of relief (*epLsar*)	Variation in lengths of transect lines measured at a given scale with orientations sampled at 5° intervals across a surface	An anisotropic surface may have a surface relief with similar orientation (i.e. more parallel scratches)	A tough diet will be more anisotropic
Heterogeneity of Area scale fractal complexity (*HAsfc*)	Variation of *Asfc* across a surface	High heterogeneity may indicate differing patterns of scratching and pitting across the surface	A mixed diet will be more heterogenic
Textural fill volume (*Tfv*)	Function of two components: (1) the shape of the surface, and (2) the texture of the surface	Deep features may dominate a surface with high *Tfv*	High *Tfv* indicates a hard diet

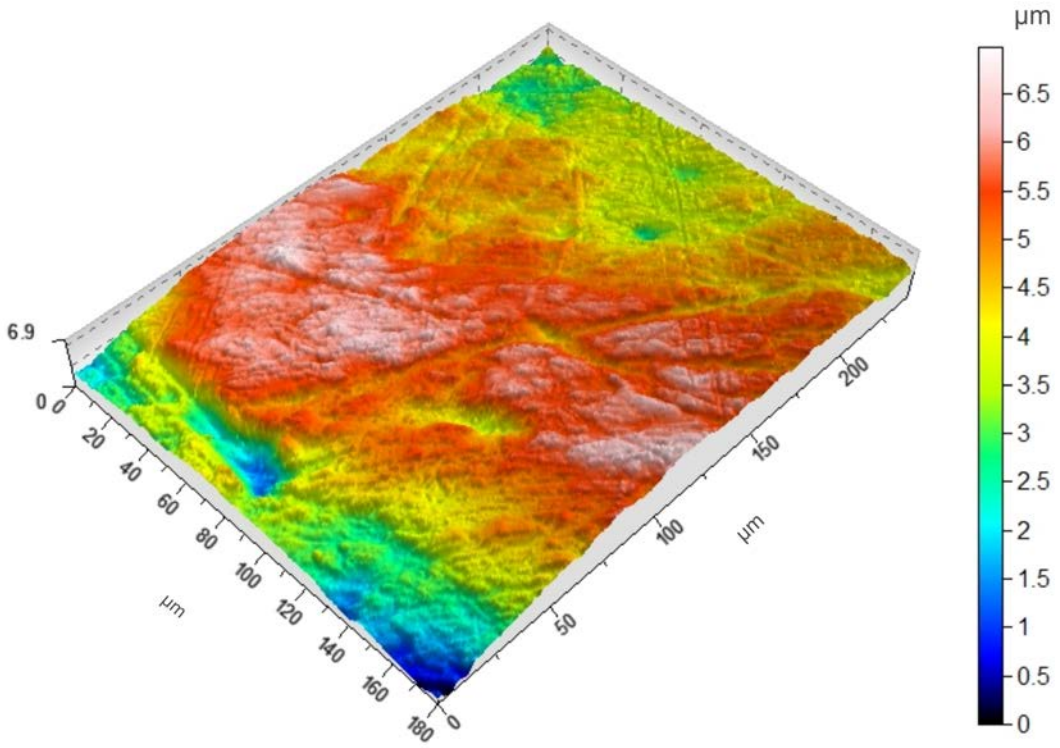


Figure 6.1 Example of a DMTA scan of good quality, with microwear features clearly visible. Individual and tooth surface: PT03 (Luhya, Port Vict.), R M_2 9 (see note on Table 6.3 for an explanation of the coding of tooth surfaces). Scale is in micrometres (µm).

From the initial 150 moulds, only 106 were successfully casted (Table 6.2). Casting of the remaining 44 moulds failed because (1) the epoxy seeped from the mould, regardless of the sturdiness of the wall of putty built around the mould, (2) the moulds rattled inside the centrifuge, causing it to become imbalanced and to halt, or (3), most often, bubbles became trapped in the silicone-epoxy boundary. In some cases, moulds were recast up to four times, but to no avail. Considering that cast quality decreases after the fourth consecutive replica from the original mould, recasting did not proceed beyond the fourth attempt (Galbany et al., 2006). Notwithstanding, casts affected by bubbles were still taken for scanning. Figure 6.2 (a, b) depicts a microwear surface resulting from scanning a cast with trapped bubbles. In Figure 6.2.a, the initial image obtained from scanning shows many spikes that indicate errors in the scanning process. By persisting through the correction of these defects, one obtains a surface such as the one in Figure 6.2.b, which is quite granular and substantially different from the image on Figure 6.1. Thus, a cast with trapped bubbles fails to reproduce accurately

the microwear surface and cannot be characterised through scale sensitive fractal analyses.

However, the main bottleneck in data collection took place upon scanning, with only 38 successful scans (Table 6.2). Besides the casts earmarked as containing bubbles during the casting process, a few casts presented microscopic bubbles which were not observable with the use of a light microscope that distorted microwear surfaces. Nonetheless, most scans failed due to the presence of biofilm produced by mouth bacteria. Figure 6.2 (c, d) illustrates a surface whose microwear was obscured by biofilm. When scanning such surfaces, spikes covered the resultant image, just as observable in the top left corner of Figure 6.2.c, or, most often, the microscope crashed and no image was produced. Indeed, the bottom right corner in Figure 6.2.c shows one of the few instances in which biofilm was successfully scanned, as indicated by the splotch-like appearance. In turn, diagonally across the image, one sees evidence of good reproduction of microwear features. Yet, correction of scanning defects on such images produced surfaces that looked flat and featureless, as in Figure 6.2.d.

Table 6.2
Number of successful stages in the DMTA process, by population.

Population	Moulds (*n*)	Casts (*n*)	Scans (*n*)	*Success Rate* [a] *(%)*
El Molo	31	23	10	*32*
Turkana	30	27	7	*23*
Luhya (Webuye)	30	16	7	*23*
Luhya (Port Vict.)	30	20	6	*20*
Luo (Port Vict.)	29	20	8	*28*
Total	**150**	**106**	**38**	***25***

[a] Success rate = Moulds/ Scans x 100

Therefore, despite the best efforts to remove biofilm, this issue severely affected data collection in the present sample. As highlighted in Table 6.2, DMTA had a success rate of only 25%. The reasons behind this are fully explored in the discussion of this chapter. Unfortunately, it is impossible to know beforehand (i.e. in the field) if a mould has successfully recorded microwear features. Notwithstanding, Table 6.3 reports on the microwear data obtained from successful scans, which includes surface complexity (*Asfc*), scale of maximum complexity (*Smc*), anisotropy (*epLsar*), heterogeneity (*HAsfc*),

and texture fill volume (*Tfv*). In addition, Appendix C contains the images of these scans (Figure C.1 to C.38, pp. 342-348).

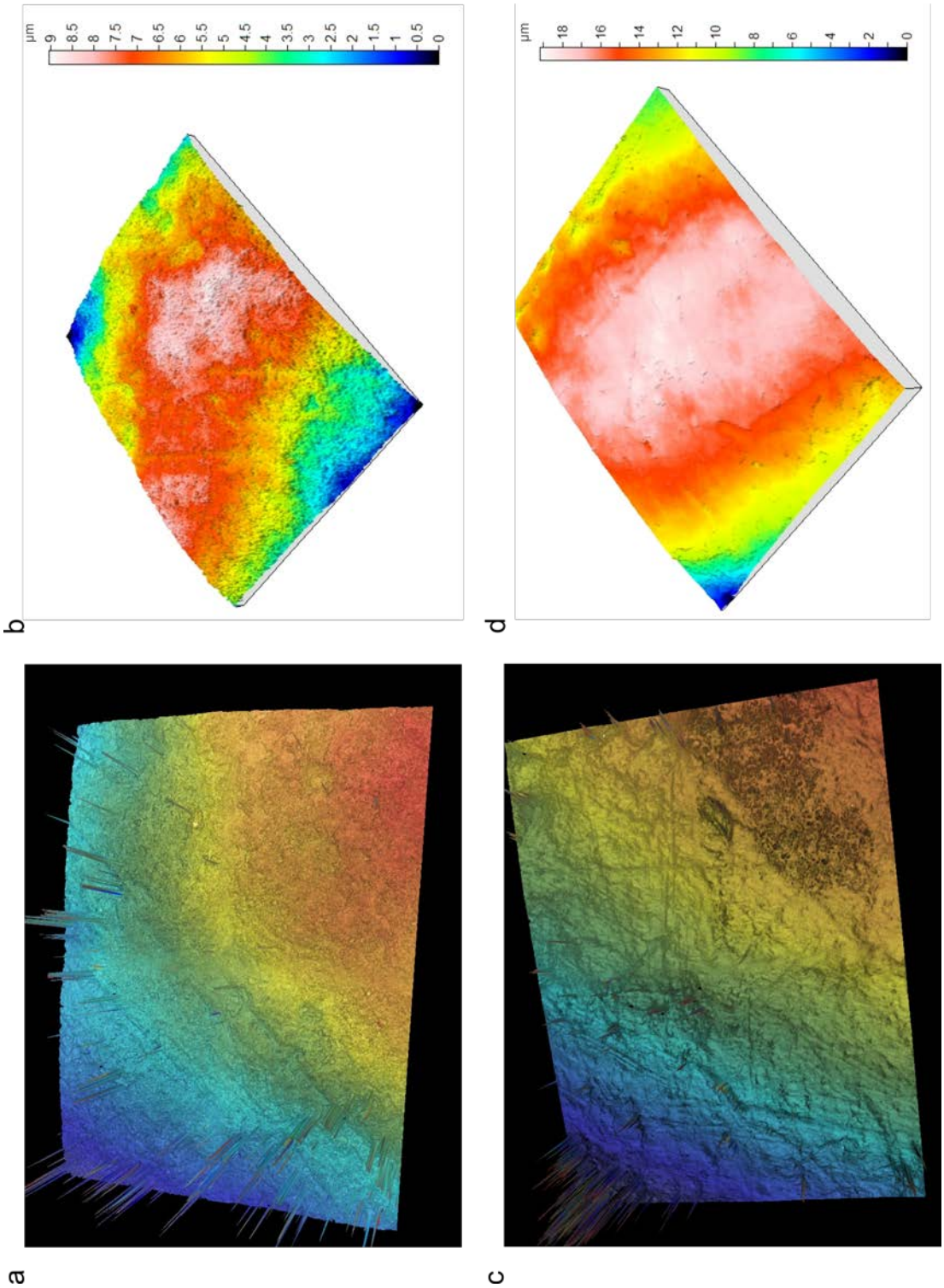


Figure 6.2 Example of DMTA scans of bad quality, with mostly obscured microwear features. Images (a, c) represent the initial scanning image produced by Sensofar® Plμ Neox, while (b, d) represent the image produced by SolarMap® after levelling and correction of defects (see Chapter 4, p. 98). In addition, images (a, b) depict a microwear surface affected by bubbles, whereas (c, d) depict an a microwear surface obscured by biofilm. Individual and tooth surface (a, b): TK13, L M1 x. Individual and tooth surface (c, d): TK03, R M1 9. Scale is in

Table 6.3
*Results of DMTA on the casts analysed in this study. DMTA parameters: Area scale fractal complexity (*Asfc*); Scale of maximum complexity (*Smc*); Exact proportion Length scale anisotropy of relief (*epLsar*); Heterogeneity of Area scale fractal complexity (*HAsfc*); Textural fill volume (*Tfv*).*

Population	Individual	Tooth Surface [a]	*Asfc*	*Smc*	*epLsar*	*HAsfc*	*Tfv*
El Molo	EM01	R M_1 9	3.78	0.675	0.0027	0.263	46999.7
	EM05	L M_1 9	3.79	0.675	0.0009	0.259	52229.1
	EM08	R M_1 9	5.89	0.675	0.0066	0.536	58888.9
	EM10	R M_1 x	3.92	0.882	0.0023	0.814	42790.8
	EM14	R M_3 x	4.24	0.675	0.0027	0.761	48165.8
	EM18	L M_1 10n	1.60	0.675	0.0028	0.530	65145.5
	EM19	L M_1 10n	1.76	0.882	0.0028	0.385	64703.8
	EM24	R M_1 9	4.78	0.675	0.0044	0.461	54709.4
	EM25	L M_1 x	1.33	1.667	0.0043	0.615	43630.4
	EM31	L M_1 x	3.21	1.116	0.0049	0.260	48665.8
Turkana	TK02	R M_1 10n	3.36	1.116	0.0007	0.538	45699.3
	TK03	R M_1 9	2.37	1.378	0.0033	0.180	64279.4
	TK09	R M_1 9	10.22	0.344	0.0007	0.213	46582.0
	TK16	R M_1 10n	3.16	1.378	0.0020	0.256	36872.2
	TK19	L M_1 9	4.40	0.675	0.0005	0.195	42242.4
	TK29	L M_1 9	2.09	1.116	0.0024	0.528	25951.8
	TK35	L M_1 10n	4.72	1.116	0.0007	0.228	44927.8
Luhya	WB01	L M_1 9	2.85	1.116	0.0017	0.192	58613.7
(Webuye)	WB02	L M_1 9	1.98	1.116	0.0041	0.310	57217.2
	WB05	R M_1 10n	6.31	1.116	0.0004	0.393	41188.3
	WB23	L M_1 9	0.97	0.675	0.0029	0.480	53447.3
	WB28	R M_1 10n	1.97	0.882	0.0036	0.515	36669.1
	WB29	R M_1 10n	4.54	1.116	0.0019	0.330	53880.5
	WB30	L M_1 9	2.50	0.882	0.0008	0.646	47174.3
Luhya	PT02	R M_1 x	3.09	0.675	0.0020	0.579	48900.8
(Port Vict.)	PT03	R M_2 9	3.42	2.328	0.0039	0.242	48077.2
	PT26	L M_1 9	1.76	1.116	0.0037	0.199	51100.5
	PT29	R M_1 x	0.99	1.667	0.0018	0.297	46389.4
	PT45	R M_2 9	1.43	1.116	0.0024	0.334	62460.6
	PT48	L M_1 9	2.57	1.667	0.0023	0.430	53076.1
Luo	PT10	L M_1 x	2.72	0.675	0.0028	0.430	50622.1
(Port Vict.)	PT25	L M_1 10n	3.38	1.116	0.0063	0.425	18847.2
	PT27	L M_1 10n	1.11	1.116	0.0047	0.666	45887.1
	PT32	R M_1 10n	2.78	1.116	0.0006	0.249	37526.1
	PT50	R M_2 9	3.44	0.675	0.0024	1.412	48984.0
	PT55	L M_1 9	2.94	1.116	0.0023	0.465	37740.8
	PT61	L M_1 x	1.78	1.116	0.0009	0.473	44632.6
	PT62	L M_1 x	1.64	1.984	0.0028	0.342	45552.0

[a] Identification of tooth surface by side (R or L = right or left), tooth (M_x = lower molar (position = 1, 2, or 3)), and facet (9,x, or 10n, see Figure 4.3, p. 99).

6.2. COMPARISON ACROSS POPULATIONS

Despite the low success rate and consequent low sample size, statistical analyses were conducted in order to test if these results were sufficient to distinguish the different populations. However, this was the only statistical analysis conducted. In other words, due to the low sample size, it was deemed prudent not to investigate the influence on DMTA parameters of other variables within each group, such as gender or age.

In addition, the parameter *Smc* was excluded from the analysis, since it often returned the same value for each surface (see Table 6.3). Goodall et al. (2015, p. 12) also report this phenomenon, but fail to explain it. In fact, several articles simply do not report *Smc* values (Arman et al., 2016; Schmidt et al., 2016). One possible explanation is that due to the material properties of food, similar features will often dominate the microwear surface, resulting in similar *Smc* values, which represent the scale at which the surface is most complex (Table 6.1).

Table 6.4
*Summary of microwear parameters for all populations. Area scale fractal complexity (*Asfc*); Scale of maximum complexity (*Smc*); Exact proportion Length scale anisotropy of relief (*epLsar*); Heterogeneity of Area scale fractal complexity (*HAsfc*); Textural fill volume (Tfv).*

		Asfc		*epLsar*		*HAsfc*		*Tfv*	
Population	*n*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*	mean (SD)	*median (IQR)*
El Molo	10	3.43 (1.48)	*3.79 (2.04)*	0.0034 (0.0016)	*0.0028 (0.0017)*	0.488 (0.201)	*0.495 (0.300)*	52593 (8105)	*50448 (10553)*
Turkana	7	**4.33 (2.77)**	*3.36 (1.79)*	0.0015 (0.0011)	*0.0007 (0.0015)*	**0.306 (0.159)**	*0.230 (0.195)*	43794 (11554)	*44928 (6583)*
Luhya (Webuye)	7	3.02 (1.82)	*2.5 (1.72)*	0.0022 (0.0014)	*0.0019 (0.0020)*	0.410 (0.153)	*0.390 (0.180)*	49742 (8328)	*53447 (11368)*
Luhya (Port Vict.)	6	2.21 (0.97)	*2.17 (1.45)*	0.0027 (0.0009)	*0.0023 (0.0013)*	0.347 (0.139)	*0.315 (0.150)*	51667 (5783)	*50001 (4299)*
Luo (Port Vict.)	8	2.47 (0.86)	*2.75 (1.30)*	0.0029 (0.0019)	*0.0026 (0.0013)*	**0.556 (0.365)**	*0.445 (0.120)*	**41227 (10192)**	*45092 (8974)*

Notes. *n* = number of individuals; SD = standard deviation; IQR = interquartile range. Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix C, p. 351-352).

Table 6.4 summarises the data for all microwear parameters and for all populations, whereas Figure 6.3 illustrates it. Regarding the statistical analyses themselves, each microwear parameter was treated as an independent variable, considering that each variable describes different attributes of the wear surface and that correlations between parameters are usually low (Scott et al., 2006). Therefore, to test for differences between groups for each parameter, a one-way analysis of variance (ANOVA) was used when the test assumptions were met, i.e. for *epLsar*, whereas a robust version of the same test was used for the remaining parameters that were not normally distributed (Table 6.4) (Field et al., 2012, p. 414). To be more precise, the robust ANOVAs conducted a one-way comparison of trimmed group means, with significance values estimated with bootstrapping (i.e. re-sampling with replacement) (Appendix C, p. 356-357) (Field et al., 2012, p. 443; Mair & Wilcox, 2017, p. 8).

Table 6.5 reports on the results of the statistical analyses, including effect sizes. Here, the effect size for robust ANOVAs is xi (ξ), specifically developed for this test (Wilcox, 2012, p. 166–169; 294), whereas the effect size for standard ANOVA is omega squared (ω^2), as advised by Field et al. (2012, p. 455).

Table 6.5
Results of the statistical analysis for the DMTA parameters.

Parameter	Statistic	Probability Value	Effect Size [a]
Asfc	$F_t = 1.59$	$p = 0.35$	$\xi = 0.59$
epLsar	$F(4,33) = 2.08$	$p = 0.11$	$\omega^2 = 0.10$
HAsfc	$F_t = 1.49$	$p = 0.32$	$\xi = 0.59$
Tfv	$F_t = 2.06$	$p = 0.22$	$\xi = 0.54$

Notes. Robust ANOVA's (F_t) used 5% trimmed means, and 999 bootstrap samples.
[a] Interpretation of ξ: small = 0.15, medium = 0.35, and large = 0.50 (Wilcox, 2012, p. 169);
Interpretation of ω^2: small = 0.01, medium = 0.06, and large = 0.14 (Field et al., 2012, p. 455).

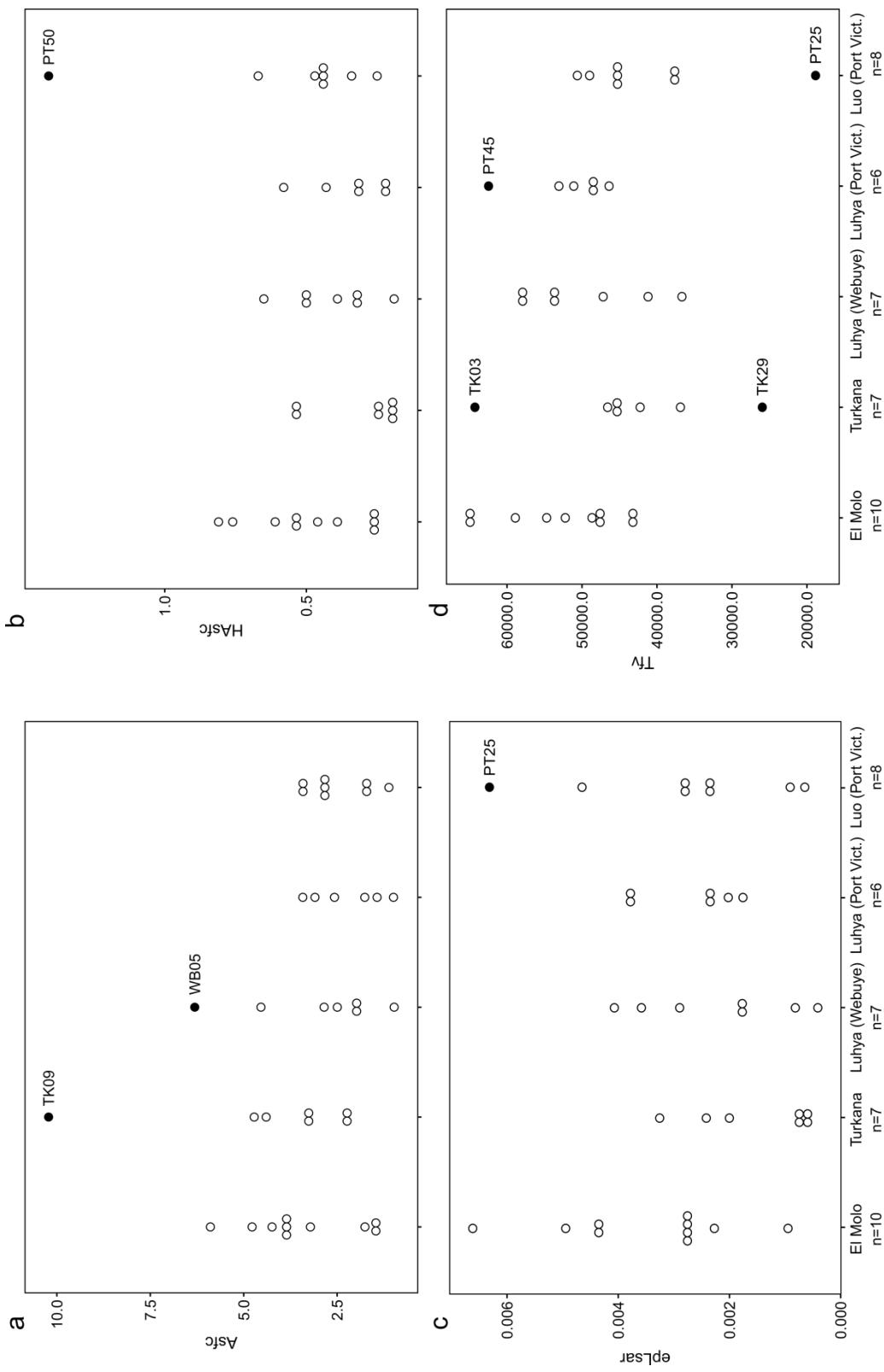


Figure 6.3 Dot plots of microwear variables for all populations (Box plots were not used due to the small sample size). Outliers are labeled with identification code of individuals and were estimated based on the interquartile range ± 1.5. Number of individuals per group are indicated below the x-axis label with "n =".

As it is clear from Table 6.5, the populations studied were not significantly different for any of the DMTA parameters. However, the effect sizes were large, except for the test on *epLsar*, where the effect size was medium, albeit close to large. A low sample size might explain this result. To confirm this, one may calculate a more adequate sample size, given the available data.

To calculate the sample size necessary to achieve a given level of power (usually 0.8), one simply conducts a power analysis resolved for sample size. The power of a test is the probability that a given test will find an effect, assuming one exists in the population (Field et al., 2012, p. 58). In this case, one hypothesises that DMTA parameters differ across populations (i.e. effect size is large), but that the power of the test is too low to find a significant difference, due to the low sample size. Note that power analysis for ANOVA uses a different effect size than those used previously. This effect size is termed Cohen's *f*, and is a biased measure, unlike ω^2, and it does not take into account the trimming and bootstrapping conducted in robust ANOVA's, which ξ does (Cohen, 1988, pp. 275–276; Field et al., 2012, p. 443; Wilcox, 2012, p. 169). Despite this, Cohen's *f* still calculates effect sizes that are large, or close to that, for all DMTA variables. Table 6.6 informs on the results of these analyses. From these, one may take the highest calculation of sample size as the minimum number to conduct DMTA successfully, i.e. the power analysis on *HAsfc*. In detail, these analyses calculated a sample size of 17 samples per group, adding up to 85 dental casts for the combination of the five populations.

Table 6.6
A priori *power analysis of ANOVA tests on DMTA parameters.*

Parameter	Effect Size (*f*) [a]	Sample Size (*n*)
Asfc	0.41	16
epLsar	0.45	13
HAsfc	0.39	17
Tfv	0.47	12

Notes. *n* = sample size per group. Power analyses based on 5 groups, a statistical power (1-β) of 0.8, and a significance level (α) of 0.05.
[a] Interpretation of Cohen's *f*: small = 0.25, medium = 0.25, and large = 0.40 (Cohen, 1988, p. 355).

Thus, this analysis highlights the need for a bigger sample size. In fact, it calculates a sample size that is more than double that obtained. However, if moulding, casting, and scanning had been entirely successful, the original sample size of 150 would have sufficed to reach and surpass the estimated sample size. Notwithstanding, one must also consider that in an ideal experimental set up, *post hoc* t-tests between pairs of groups would also reach significance. Such tests, as well as corresponding power analyses, were not conducted, since the initial ANOVA's were non-significant. Nevertheless, one may posit that the initial 150 samples, being considerably more than the estimated 85, would be enough to successfully detect any existing effects.

To sum up, formal tests failed to detect a difference in DMTA variables among the populations of this study. Nevertheless, the high effect sizes of the statistical analyses suggest that the diets of these groups affect DMTA variables. This hypothesis is consistent with the calculation of a more adequate sample size, which is higher than that available for analysis. In other words, there is no confidence in a true negative. Therefore, the results suggest that these populations differ in the material properties of consumed diet, but that it is not possible to exclude the possibility of a similar diet in terms of DMTA parameters, or the possibility that microbial film has affected to some degree the samples included in the analysis. Notwithstanding, one may still describe any trends observable in the data, as well as visually compare the DMTA results to those obtained through the food frequency diet questionnaires. The next section focuses on this effort.

6.3. EFFECT OF DIET ON MODERN HUMAN DMTA VALUES

In order to understand the DMTA values obtained in this study, one may describe the trends observable in Figure 6.3 and Table 6.4 for each parameter, and interpret them on the basis of the results from the diet questionnaires. Considering the low number of microwear results, no other statistical analyses, such as correlations, were conducted.

First, the surface complexity (*Asfc*) (Figure 6.3.a) differs very slightly between groups. Presenting the lowest values, the Luo and Luhya from Port Victoria, have similar complexity values. Next, if one excludes the very clear outlier in the Turkana group (TK09), then this group follows the two previous groups, with slightly more complex surfaces. Following this, the El Molo and the Luhya (Webuye) have wider ranges, but reach higher complexity values. These results suggest that the El Molo and the Luhya (Webuye) vary more in the material properties of their diet than the other three groups, whereas the higher mean values of these two groups indicate that these populations consume higher amounts of hard items, than the previous groups.

Second, the heterogeneity of surface texture (*HAsfc*) (Figure 6.3.b) seems to be similar across all groups, if once again one excludes the very distant outlier (PT50), which indicates that the complexity does not vary across the microwear surface (Table 6.1).

Third, the anisotropy (*epLsar*) (Figure 6.3.c) differs more substantially among groups. The Turkana have the lowest mean value, followed by the Luhya (Webuye), and then by the Luo and Luhya (Port Vict.), and finally, by the El Molo. However, the El Molo and the Luo (Port Vict.) present high variability. Therefore, these results confirm the complexity results to some extent: the Luo and Luhya (Port Vict.) have a tough diet, followed by the Luhya (Webuye), and then by the Turkana. In contrast, the El Molo have the toughest diet, which seems to contradict the previous result of a hard diet for this group, albeit the *epLsar* values also reveal a wide variability within this group.

Finally, the textural fill volume (*Tfv*) (Figure 6.3.d) also differs among groups. The lowest *Tfv* values belong to the Turkana and Luo (Port Vict.), followed by the Luhya (Webuye), the Luhya (Port Vict.), and last, the El Molo. Notably, this is the only microwear parameter for which the two Port Victoria populations do not have similar values. Compared to the other microwear features, the El Molo once again emerge as the group with the hardest diet. Similarly, the Turkana values indicate a somewhat tough diet, just as suggested by the complexity values, but not by the anisotropy values. Furthermore, the textural fill volume of the Luo (Port Vict.) confirms a tough

diet for this group, whereas the Luhya (Webuye) and the Luhya (Port Vict.) have harder diets.

Taken together, the microwear parameters suggest that, among the groups of this study, the El Molo have simultaneously the hardest *and* the toughest diet, although a considerable variability is present within the group. Subsequently, the Turkana, as well as the Luo and Luhya, from Port Victoria, have tougher diets than the Luhya (Webuye), which tend towards a harder diet.

To make sense of these results, one may consider the results of the semi-quantitative food frequency questionnaires. As a first point, the only common outlier between the microwear features and those identified in energy and nutrient intake is the individual TK03 from the Turkana (see Chapter 5, Figure 5.1, p. 117). In fact, this individual consumes a lot of protein and fat, while its teeth surfaces have high textural fill volume, which indicates a particularly hard diet. A high meat intake explains the high consumption of protein and fat for this individual (Appendix C.1); however, one expects meat to be a tough dietary item, rather than hard (El Zaatari, 2010; Scott et al., 2006). As way of explanation, it is possible that this individual follows a more traditional Turkana diet, which includes more abrasive plant products than the other more sedentary Turkana, which consume less abrasive agricultural produce. Alternatively, the processing of meat may add abrasives to the diet, although there is no compelling evidence to endorse this hypothesis (El Zaatari, 2010; Galvin, 1985).

Following this, the same principle that fleshy foods produce a tough signal explains why the Luo and Luhya, from Port Victoria, have tougher diets, as fish is also a fleshy food. The same reasoning applies to the Turkana, although, in this case, one must consider the TK03 case where higher meat consumption meant a harder diet, and not the contrary. In addition, the evidence for sedentarisation in this group, and concurrent reduction in the intake of animal products (Chapter 3, p. 65, and Chapter 5, p. 149), suggest that this group has a tougher signal than the other populations, but not a particularly strong one, probably through the consumption of tough plant foods. Conversely, one would expect the Luhya (Webuye) to have a tough signal, considering

that the diet questionnaires distinguish this group as the one consuming higher amounts of animal foods. This incongruence may be explained by an inaccuracy in the diet questionnaires, or by the consumption of a higher variety of hard plant items, which swamps the tough signal left by the tough animal foods.

Finally, the microwear features of the El Molo might arise from the food preparation techniques practiced by this group. In particular, the El Molo sun-dry fish in open racks along their beaches (Chapter 3, p. 59). This may result in the addition of sand to the food, leading to a particularly abrasive diet. El Zaatari (2010, p. 80) supports this explanation, since she found that a group that dried meat along beaches had microwear features indicative of a harder diet when compared to groups that did not process meat this way. Nevertheless, Schmidt et al. (2016) suggest that groups that consume less agricultural produce have diets that are simultaneously harder and tougher than the diets of those that consume such staples. Agricultural produce is usually more processed (e.g. turned into flour, or extensively washed) and thus, leaves less microwear marks than foods from more traditional diets. Consistent with this hypothesis, the diet questionnaires suggest that the El Molo consume less agricultural produce than the other groups in this work (Chapter 5, p. 145; Appendix C.1). This finding has important consequences to future DMTA studies, since food processing techniques can be a confounding factor in the interpretation of DMTA values.

6.4. COMPARISON WITH OTHER STUDIES

To further contextualise the results reported in Figure 6.3 and Table 6.4, one may compare them to the absolute values obtained by other DMTA studies from archaeological human groups with diverse diets. In particular, the microwear results of this work are compared to those of Schmidt et al. (2016) and El Zaatari (2010): the former discusses the microwear signal of grouped agriculturalists from England, Greece, and Nepal, and pastoralists from Mongolia, whereas the latter studies the microwear signal of different hunter gatherer groups, namely the Andamanese, from

the Andaman Islands, the Chumash, from an island in southern California, the Fueguians, from the southernmost part of South America, the Khoe San, from South Africa, and the Tigara, from Alaska. These two studies were chosen because they address the microwear signal of modern groups whose diets have been well studied, and because both studies used DMTA. Few other studies meet these two conditions. Accordingly, this section does not address the study on the microwear of the Baka (Romero et al., 2013) because this work used traditional microwear techniques, rather than DMTA. Instead, the results of this study will be considered in the discussion of this chapter.

Regarding complexity results, the populations from the present work presented *Asfc* values that were higher than those from either the agriculturalists (mean = 1.27; SD = 0.37; n = 46) or the pastoralists (mean = 0.98; SD = 0.31; n = 49) studied by Schimdt et al. (2016), and values that were lower than those of the Tigara (mean = 6.57; SD = 5.81; n = 25) and the Andamanese (mean = 6.34; SD = 0.4.22; n = 30), but close to those of the Khoe San (mean = 3.62; SD = 2.36; n = 43) and the Chumash (mean = 2.79; SD = 2.34; n = 13), and higher than those of the Fueguians (mean = 0.95; SD = 0.29; n = 6) (Table 4, El Zaatari, 2010, p. 78; Table 2, Schmidt et al., 2016, p. 8). Notably, the differences between the groups of this study were negligible when compared to the results from the two studies. Therefore, this suggests a harder diet for these groups than for those of Schmidt et al. (2016), and a diet with similar material properties to those of the Khoe San and of the Chumash. Surprisingly, these two groups had very different diets: whereas up to 80% of the Khoe San diet consisted of plant foods, the diet the Chumash relied heavily on fish sources. El Zaatari (2010, p. 81) proposes that the intermediate hard signal for both groups comes from the plant part of the diet, such as nuts, seeds, and tubers, for the Khoe San, and berries, seeds, and tubers for the Chumash. Thus, complexity results can be biased towards whichever part of the diet contains more abrasives, and not to whichever element makes up most of the diet.

Subsequently, the values of heterogeneity of surface texture of the present work are slightly higher than those from El Zaatari (2010). However, just like for this study,

El Zaatari (2010) fails to find a significant difference between the *HAsfc* values of different groups. In addition, Schmidt et al. (2016) do not report on *HAsfc* values, for reasons the authors do not mention. Considering that the small change between the values obtained here and by El Zaatari (2010) could result from a slight difference in microscope settings (Arman et al., 2016), then it is possible that, in modern humans, the complexity of microwear features is similar across a tooth surface, which might simply indicate that the particular constraints of jaw movements during mastication produce homogenous patterns across the tooth facets.

Then, when comparing the anisotropy values from this work to those of Schmidt et al. (2010), the El Molo have values similar to those from pastoralists (mean = 0.0034; SD = 0.0017; *n* = 49) and slightly lower than those from agriculturalists (mean = 0.0041; SD = 0.0015; *n* = 45); whereas the remaining populations from this work all had *epLsar* values that were lower than either pastoralists or agriculturalists. Next, for the El Zaatari (2010) study, the Turkana had lower *epLsar* values than all hunter gatherer groups, whereas the Luhya (Webuye), and the Luhya and Luo (Port Vict.) had values close to the Tigara (mean = 0.0029; SD = 0.0015; *n* = 25), Andamanese (mean = 0.0026; SD = 0.0012; *n* = 30), Khoe San (mean = 0.0026; SD = 0.0013; *n* = 43), and Chumash (mean = 0.0023; SD = 0.0007; *n* = 13), which, in turn, were undistinguishable from each other. In addition, the El Molo had higher values than the populations just mentioned, but still lower than the Fueguians (mean = 0.0044; SD = 0.0014; *n* = 6), the only distinct group in the anisotropy signal described in the El Zaatari (2010) paper (Table 4, El Zaatari, 2010, p. 78; Table 2, Schmidt et al., 2016, p. 8). These comparisons suggest that the populations from the present work have diets that are not very tough when compared to other groups. Indeed, the El Molo is the only group that presents a tougher diet, with anisotropy values similar to the Mongolian pastoralists, and to the fish eating Fueguians.

Finally, the texture fill volume values from this study were substantially higher than those obtained by either Schimdt et al. (2016) or El Zaatari (2010). In detail, the *Tfv* values from this work varied between the mean of 41227 for the Luo (Port Vict.) and

the mean of 52593 for the El Molo. In contrast, the agriculturalists from Schmidt et al. (2016) had a mean value of 37401 (SD = 90999; *n* = 41), whereas the pastoralists presented a mean value of 31723 (SD = 10523; *n* = 44). For the El Zaatari (2010) study, *Tfv* values varied from the lower mean of 5225 (SD = 3522; *n* = 6) for the Fueguians, and the higher mean of 11912 (SD = 4657; *n* = 25) for the Tigara. Schmidt et al. (2016, p. 10) state that past populations often have high textural fill volumes due to rapid rates of wear, which could also explain the values obtained in the present study since many of older participants had well-worn teeth.

In brief, compared to previous research, the groups from this study have relatively homogeneous diets in terms of its material properties. It is possible that this pattern results from a common reliance on the same food staples (e.g. maize); particularly considering that microwear features might be biased towards the dietary elements that more readily leave marks on the teeth. Hence, one finds that all the groups of this study have an intermediate hard diet, concurrent with abrasives being included in the diet, probably within the plant part of the diet. The comparison confirms the only exception to this pattern, an important tough component in the El Molo diet, possibly due to the inclusion of abrasives in fish during sun drying, and consequent repetitive and directional jaw movements.

6.5. DENTAL MICROWEAR TEXTURE ANALYSIS DISCUSSION

As stated at the start of this chapter, DMTA studies provides direct evidence on the food consumed by an individual during his or her lifetime, while controlling for the intra-observer error that characterised earlier microwear studies.

Notably, proof of concept for the DMTA technique derives from (1) support of DMTA findings by other means of dietary assessment in human archaeological remains (Grine et al., 2012), (2) DMTA studies conducted in both human and animal historical remains whose diets are well documented (El Zaatari, 2010; Schmidt et al., 2016; Scott et al., 2012), (3) one *in vitro* study (Hua et al., 2015), and finally, (4) agreement between

results obtained by different microwear texture laboratories (Arman et al., 2016). The clear gap in this scenario concerns the lack of evidence on how DMTA differs between living human populations with diverse, but quantifiable, diets. Closing this gap would permit a better understanding of the production of microwear features, and consequently increase the resolution of this technique. Furthermore, and considering that real-life mastication is complex and difficult to simulate (Hua et al., 2015), it is crucial to be able to consider microwear data of human archaeological material within a framework of modern human microwear. The present work aimed to start closing this gap.

With this goal in mind, Chapter 4 (p. 96) clearly outlined that the main obstacle to conducting DMTA in living subjects is the obscuring of microwear features by microbial activity in the mouth. To address this issue, the present work used a recently developed method that specifically focused on the removal of bacterial biofilm (Livengood, 2015, personal communication). However, Livengood trained me on this method remotely, entirely through online correspondence, and this might have severely affected the results. Indeed, the presence of biofilm remained the main cause of unsuccessful DMTA in this study. At this point, one may compare the method of teeth cleaning used here to that used in other published *in vivo* studies that involved dental moulding. This comparison has the objective of improving on the method used here, but also of detecting if any miscommunication that took place during the remote training could have caused the data loss. Table 6.7 summarises the characteristics of these *in vivo* microwear studies.

From this table, a few general features are important. First, the removal of biofilm may be mechanical (e.g. tooth brushing, use of oral irrigation device) or chemical (e.g. rinse with a weak solution of sodium hypochlorite). Second, to the best of my knowledge, *in vivo* dental moulding has only targeted primate species. And third, most studies analysed dental microwear, albeit one targeted dental topographic analysis. This second method aims to characterise variably worn teeth in three dimensions, and

consequently, can be said to study macrowear, as opposed to microwear (Chapter 2, p. 20) (Dennis et al., 2004).

Besides these general aspects, Table 6.7 evidences several differences in cleaning method that warrant attention. First, some studies administered atropine to subjects, usually 0.05 mg per kg of body weight (Teaford & Lytle, 1996; Teaford & Oyen, 1989b). With the exception of Teaford and Lytle (1996), this drug was used only in animal studies, in order to reduce the excess salivation caused by the anaesthetic ketamine hydrochloride. However, Teaford and Lytle (1996) simply state that they followed the method described in Teaford and Oyen (1989b), and so it is likely that they did not give atropine to their participant. Even though this pre-treatment could improve the quality of moulding, the use of this drug was required only because of the effect of the other drug. Thus, it is unlikely it would improve on a method targeting living people, not to mention the required training and ethical implications of using such a technique.

Second, a few studies used some form of chemical cleaning step, which included weak solutions of sodium hypochlorite, a solution of phosphoric acid, and a standard dental cleaning product (Lacer®). From these, Teaford and Oyen (1989b) report that the use of sodium hypochlorite was never 100% successful in biofilm removal, although there is evidence that it reduces biofilm in human subjects, although it requires repeated use (De Nardo et al., 2012). Subsequently, phosphoric acid could be dangerous to use without the appropriate expertise. This product is used as an etching solution, to clean and roughen the surfaces of teeth prior to the placing of dental appliances or fillings, and it may lead to enamel demineralisation (Knösel, Bojes, Jung, & Ziebolz, 2012). Finally, dental cleaning products are probably only as effective as a solution of sodium hypochlorite (De Nardo et al., 2012). Therefore, even though a chemical cleaning step could improve on the removal of biofilm, it would not guarantee it.

Table 6.7

Comparison of the cleaning method used in in vivo *studies. When reported, time spent on each step is provided; otherwise, only yes or no is included.*

Cleaning Method					Study Details				
Pretreatment	Tooth Brushing	Chemical Cleaning	Oral Irrigation	Air jet	Technique	Data Acquisition	Targeted Group	Mould Number	Reference
Atropine	No	1-3% NAClO solution	~2min	~2min	Occlusal microwear	SEM	Primates	NA	Teaford and Oyen (1989b)
					Occlusal microwear	SEM	Modern Humans	2[c]	Teaford and Lytle (1996)
Atropine	Yes	3% NAClO solution	Yes	~2min	Occlusal microwear	SEM	Vervet monkey	15	Teaford and Oyen (1989a)
Atropine	Yes	0.15% NAClO solution	~1min	~2min	Occlusal microwear	SEM	Mantled howler	33	Teaford and Glander (1991)
No	No	0.15% NAClO solution	No [a]	~30s	Occlusal microwear	SEM	Modern Humans	9	Teaford and Tylenda (1991)
Atropine	Yes	40% H_3PO_4 solution	Yes	~2min	Dental topography	LS	Mantled howler	56[c]	Dennis et al. (2004)
No	No	dental product (Lacer®)	No	~30s	Buccal microwear	SEM	Modern Humans	4[c]	Romero et al. (2007)
					Buccal microwear	SEM	Modern Humans	20[c]	Romero et al. (2012)
No	Yes	No	No	Yes	Buccal microwear	SEM	Modern Humans	36	Romero et al. (2013)
One cup of popcorn	Yes	No	~40s [b]	~2min	Occlusal microwear	WLCM	Modern Humans	150	Present study

Notes. NAClO = sodium hypochlorite; H_3PO_4 = phosphoric acid; min = minute; s = second; SEM = Scanning electron microscopy; LS = Laser Scanner; WLCM = white light scanning microscopy.

[a] gauze swab instead; [b] plus cotton swab; [c] some moulds are repetitions from the same individual.

Following this, and as mentioned above, one must consider that different studies targeted different data, and used different approaches to acquire it. One of the studies (Dennis et al., 2004), used a laser scanner to conduct dental topographic analysis. Since this method focuses on the three-dimensional shape of tooth surfaces, its results will be less affected by the presence of biofilm. In addition, three of the studies used SEM to study buccal microwear (i.e. the microwear on the cheek side of the tooth), as opposed to occlusal microwear (i.e. the chewing surface) (Romero, Galbany, De Juan, & Pérez-Pérez, 2012; Romero et al., 2007, 2013). Interestingly, these studies used simpler cleaning methods (e.g. no use of water irrigation device), and never refer to the need to remove biofilm. One may posit that buccal microwear is less affected by biofilm, although no evidence was found to support or oppose this hypothesis. Alternatively, it is possible that these studies erroneously count marks on biofilm as microwear. Ultimately, all the microwear studies on Table 6.7, apart from the present work, used SEM to collect data. This approach to microwear study involves the manual counting of features by a researcher, which means that one may analyse images only partially obscured by biofilm, such as the one in Figure 6.2 (c, d) (Teaford & Oyen, 1989a, 1989b), whereas the white-light scanning microscope used in DMTA cannot cope with this issue.

A final clear difference between cleaning methods in Table 6.7 is the use of popcorn in the present study. According to Livengood (2015, personal communication), popcorn is a soft material that proved successful in the removal of biofilm (consider the feeling of popcorn sticking to teeth). However, no other study used this step and the low success rate reported here might indicate that it does not improve on the results. Notwithstanding, Teaford and Oyen (1989b, p. 75) mention that animals that have eaten in the 4-8 hour period before an impression session yield better impressions than animals that have been fasted for 10-12 hours, which seems to credit the inclusion of this step in the procedure.

Besides comparing the method used in this study to other published methods, it was also possible to compare the details of the technique against the original method by

Livengood (2015, personal communication) that is used at the Ungar Laboratory. At this time, two main discrepancies stood out. First, in the original method, Livengood brushed the teeth of the participants, whereas in this study participants brushed their own teeth. In this case, the source of confusion originated from the word choice used upon the initial communication between us, in which the subject of the action was left ambiguous. However, considering that participants of the present work were observed, and even instructed, during the tooth brushing, there is no certainty that this factor could be responsible for the data loss, although it probably contributed to it. Second, the oral irrigation device used in the present work was from a different make of that used in the Ungar Laboratory. In particular, this study used a Panasonic® EW-DJ10 Travel Oral Irrigator, while Livengood used a Waterpik® Nano™ Water Flosser WP-250UK. When comparing the settings of these two devices, it was found that the first delivered 247 ml of water per minute at pressure of 77 psi, and had an operating time of 40 seconds per full tank, whereas the second delivered 500 ml per minute at a pressure of 90 psi, and had an operating time of 60 seconds per full tank. Therefore, since the Waterpik® is more powerful than the Panasonic®, it is plausible that the first is more efficient in biofilm removal than the second. However, Teaford and Oyen (1989b, p. 74) used a water irrigation device at its lowest setting, delivering 37 ml of water per minute at a pressure of 5 psi for about 2 minutes. Unfortunately, the other studies included in Table 6.7 that used such a device failed to report on its settings, rendering further comparison impossible. In addition, one must consider that the Panasonic® was chosen because it was the stronger such device that operated on batteries (unlike any other device from the Waterpik® brand). This was a sought characteristic due to the known difficulty in accessing electricity in the field. Summing up, the power of oral irrigation devices might play a factor in biofilm removal and mould quality, although there is some evidence that lower settings are also effective (Teaford & Oyen, 1989b, p. 74). This must be carefully considered prior to fieldwork, particularly when working in remote areas.

Apart from the cleaning of bacterial biofilm, one must also address other problems along the entire process, namely casting issues, such as the trapping of bubbles. It is

hard to know how to avoid this complication, although one solution might be to cut down moulds to a more manageable (and less wobbly) shape before centrifuging them. It is also possible that the trapping of bubbles arose from the rough surface produced by the presence of biofilm.

Overall, it is likely that all of the factors mentioned above contributed to some degree in hampering DMTA. Indeed, as it stands, there is no bulletproof way to completely remove bacterial biofilm, a fact already identified by Teaford and Oyen (1989b). Despite this, none of the articles included in Table 6.7 mention any success rates, as the one seen in this study, implying that they obtained microwear data from every tooth they casted. Although one must consider that it should be easier to work around biofilm when using SEM to study microwear, what arises from the combined factors is that adequate training and experience is crucial to properly mould and cast the teeth of living subjects. On this subject, it must be said that, prior to fieldwork, the entire procedure was repeatedly practiced on (open-minded and) willing office colleagues, with any doubts that arose throughout the process being clarified by Livengood, and that, at the time, moulding and casting seemed successful under a standard light microscope. Therefore, in the case of DMTA in living people, adequate training might require "shadowing" another experienced researcher throughout every step of the analysis. As DMTA moves forward, it might be that a need arises for field-directed workshops on this technique. Even though *in vivo* DMTA studies might remain a small subset of this type of analyses, it is also true that DMTA in archaeological samples can also be fraught with obstacles. For instance, post-mortem wear may be confused with microwear, and it requires an experienced eye to be able to tell the difference (Appendix C, p. 361-362). In other words, it is important to let the scientific community know of the pitfalls of attempting this technique.

Despite the considerate obstacles faced during DMTA in living people, the results included in this chapter suggest that occlusal microwear recorded differences in diet across the populations of this study, but that the analysis did not have enough power to detect those differences, due to the low sample size. Notwithstanding, the trends

observable in the data indicate that the El Molo have a harder *and* tougher diet than all the other groups, followed by the Turkana, and the Luo and Luhya (Port Vict.), whom have a tougher diet than the Luhya (Webuye). Based on data from the diet questionnaires, it is possible that the microwear signal of the El Molo results from the addition of sand to fish during sun drying, and to the consumption of little agricultural produce. In addition, the harder diet of the Luhya (Webuye) may derive from a higher consumption of agricultural produce among this group, and from a higher consumption of animal and fish products among the remaining groups.

Furthermore, when comparing these results to previous research, the microwear signals of the groups studied here point to an intermediate hard diet, apart from the El Molo, whom also have a considerable tough element in the diet. The similar microwear patterns across groups could emerge from the use of the same plant food staples, such as maize, beans, and rice, whereas the tough signal from the El Molo might, once again, derive from the technique used for fish preservation, which leads to the consumption of many abrasives.

Following this, exogenous grit introduced into food through either windblown dust (Ungar, Teaford, Glander, & Pastor, 1995) or during food processing (Teaford & Lytle, 1996) can change microwear patterns in such way that impedes the reverse engineering process of assessing diet from microwear. For instance, the amount of grit in the air may change seasonally (Ungar et al., 1995), whereas most modern human groups engage in some form of food processing. I suggest that these analyses may still be useful when considered vis-à-vis other techniques, and when addressing well delimated research questions.

To complete this discussion, one needs to consider the microwear results of the Baka reported in Romero et al. (2013). Clearly, the results from that study are not truly comparable to those reported here because Romero et al. (2013) look at buccal, rather than occlusal, microwear, and because they used SEM technology and individual counting of microwear features, rather than DMTA. In particular, since buccal microwear does not result directly from the act of chewing, and consequent contact

between teeth, one expects microwear features to be less pronounced than in occlusal surfaces. Nevertheless, the authors find that the Baka have low density patterns of buccal microwear, which is consistent with a diet with few abrasives that includes some meat and cooked foraged foods. In other words, one may see the Baka as having a microwear pattern opposite to the El Molo with few microwear features, and a diet that is neither particularly tough nor hard.

To conclude, one may state that DMTA remain a powerful tool of diet assessment (both in the present and in the past), although further research must be conducted in order to better understand the production and turnover of microwear. In detail, there is still much to understand on what DMTA variables mean biologically (e.g. what does a high value of *Asfc* translate to in terms of diet); on which food items leave traces in the teeth, and, as important, which do not; on how does buccal and occlusal microwear compare; and on what dietary composition corresponds to in microwear patterns. Furthermore, method wise, one may improve on the moulding teeth of living subjects (e.g. by adding a chemical cleaning step, or by setting up training workshops), but also on the standards to correctly identify true microwear.

6.6. SUMMARY OF CHAPTER 6

In this chapter, DMTA had a very low success rate (25%), mostly due to the incomplete removal of bacterial biofilm. Several aspects are likely to have contributed to this, such as the lack of a chemical cleaning step, the tooth brushing done by the individual instead of the researcher, and the weaker oral irrigation device (when compared to Livengood, 2015, personal communication). In addition, the laboratory set up might not have been ideal for the production of epoxy casts. However, the clearest obstacle was the lack of adequate training to conduct the data collection. Due to the inadequate cleaning of teeth, statistical analyses failed to detect a difference in microwear parameters across populations. Notwithstanding, high effect sizes suggest these differences exist, and power analysis confirm that the power of the statistical analysis

was too low to detect possible differences because of the low sample size. In addition, calculation of adequate sample sizes indicates that the initial sample size would have sufficed to detect microwear patterns, suggesting the initial research design was adequate.

Despite these difficulties, the microwear results presented in this chapter indicate that the El Molo have a diet that is both hard and tough, whereas the Luhya (Webuye) tend towards a hard diet, and the remaining groups tend towards a tough diet. In contrast, when compared to previous studies, the differences among groups dilute, with almost all groups having microwear values that indicate an intermediate hard diet. The only exception, the El Molo, have a tough component in the diet that distinguishes them from the other populations. In addition, the Baka have the opposite pattern of the El Molo, based on data from previous research. As expected, the specificities of the diet of each population explain these patterns: for the El Molo, the consumption of hard plant items and the inclusion of sand abrasives in fish explain the combined tough-hard signal; for the Luhya (Webuye), the consumption of a high variety of plant products could explain the harder signal; for the remaining groups in this work, the consumption of animal and fish products explains the tough signal; and last, the consumption of well-cooked and processed plant foods may explain a lack of microwear features among the Baka, although this last finding might result from a difference in technique and not from a true difference in diet. Furthermore, when compared to other studies, all the groups except the El Molo emerge as having an intermediate hard diet, possibly due to the consumption of the same food staples across groups; whereas the tough signal for the El Molo is likely once again due to the inclusion of abrasives during fish processing.

Thus, one may conclude that DMTA can be an effective method to assess diet, as well as food processing techniques, but that it must be carefully used, since microwear features are easily obscured by biofilm in living people, and erased by post-mortem wear in archaeological teeth. In addition, considering that it can be difficult to distinguish which microwear patterns are produced by foods directly, and which are a

consequence of exogenous grit, DMTA in modern human groups may be useful only within specific research investigations, and not as a general method for diet assessment in the past.

CHAPTER 7

STABLE ISOTOPE RATIO ANALYSIS RESULTS AND DISCUSSION

Having discussed the results of the semi-quantitative food frequency questionnaires and of the dental microwear texture analysis, the last set of data left to examine are stable isotope ratio analyses. This chapter addresses these analyses, and it encompasses: (1) the quality assessment of isotopic results; (2) the analysis of the effect of population, gender, and age on the isotopic results; (3) the comparison of isotopic results between the body pools analysed; (4) the juxtaposition between the isotopic results and the dietary information available for each population; and (5) the comparison to other isotopic studies on hair and nail.

7.1. STABLE ISOTOPE RATIO RESULTS

This section addresses the quality of stable isotope results and discusses any important points that arise from this issue, whereas section 7.2 (Comparison across Populations) interrogates these results statistically. For clarity, the different types of samples analysed (i.e. food, hair, nail, and breath) are discussed separately. And for brevity, the full sample details and isotope results are presented in Appendix D.

7.1.1. FOOD RESULTS

Figure 7.1 and Table 7.1 inform on the stable isotope results of food samples. Regarding the fish foods, all samples yielded reliable results. As explained in Chapter 4 (p. 106), this assessment is based on (1) C/N ratios, which fell within the range of 2.9 to 3.6 acceptable for collagen, with the exception of the *Dagaa* (whole) sample; (2) the errors of replicate measurement, which were always below 0.3‰ for $\delta^{13}C$ and 0.4‰ for $\delta^{15}N$; and (3) carbon and nitrogen yields, which were well above the minimum of 13% and 5%, respectively (Ambrose, 1990; DeNiro, 1985). Regarding the C/N ratio of the *Dagaa* (whole), one must consider that this sample is a mix of several tissues (besides bone collagen), and thus, that it is acceptable for the C/N ratio to diverge slightly from that expected for pure bone collagen. In contrast, it is difficult to evaluate the quality of plant isotopic results, since there is no equivalent to the C/N ratios of proteic tissues. In addition, plant isotope values can vary considerably within the same plant, particularly in nitrogen isotopic values (Lightfoot et al., 2016), as is clear from the standard deviations in Table 7.1. Therefore, this section includes the isotopic results of all plant samples. However, one must keep in mind that some of these might be unreliable.

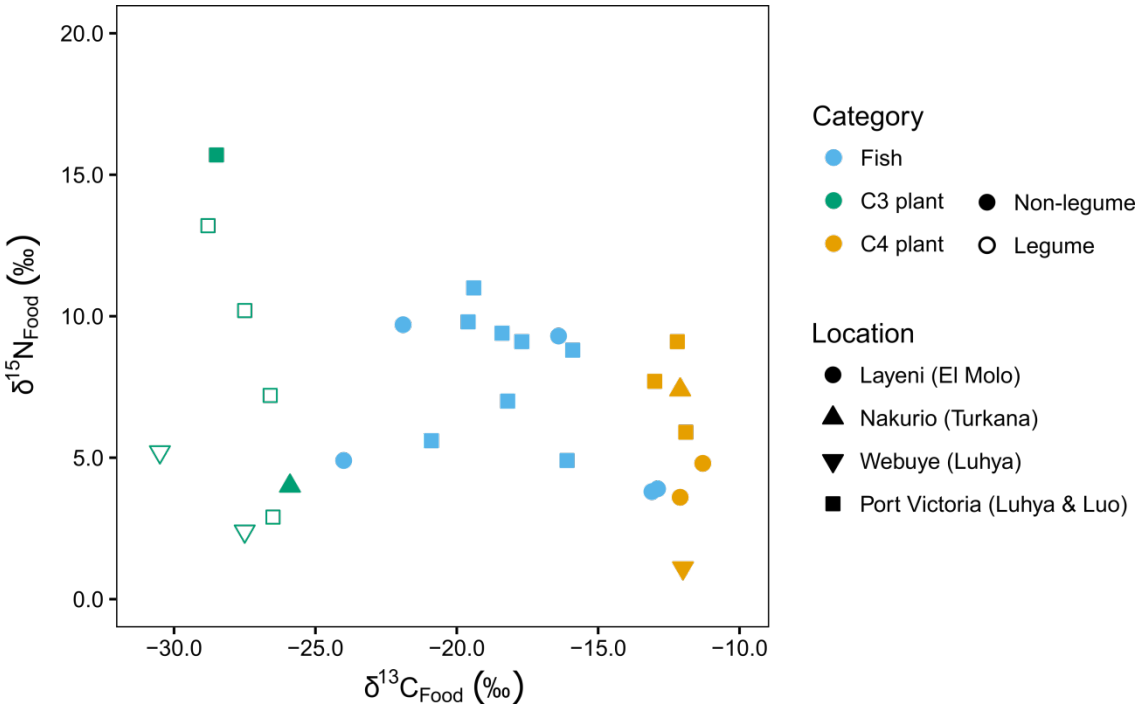


Figure 7.1 Carbon and nitrogen isotope ratios for individual food samples.

Following this, Figure 7.1 illustrates different points about the food isotope results. First, food samples from Webuye have relatively low nitrogen isotopic values, although this finding is tentative since the sample size is three. Otherwise, there is no other pattern in either carbon or nitrogen isotopic values in respect to the location where the samples were collected. In contrast, a pattern is present when considering the different types of foods. In detail, the carbon isotope values clearly distinguish between the three food categories, wherein the C_4 foods have the highest $\delta^{13}C$ values and group together (ranging from −13.0‰ to −11.3‰), while the C_3 foods have the lowest $\delta^{13}C$ values, but do not group as tightly (ranging from −30.5‰ to −24.1‰). These carbon isotopic values fall well within the expected ranges of –20‰ to –6‰ for C_4 plants and −35‰ to –21‰ for C_3 plants (Smith & Epstein, 1971). In turn, the carbon isotopic values of fish samples (ranging from −24.0‰ to −12.9‰) span the entire space between C_3 and C_4 foods. One predicts that the $\delta^{13}C$ values of animals (or fish) primarily track the plant diet at the base of the food chain (DeNiro & Epstein, 1978; McCutchan et al., 2003). Accordingly, these values fall within the ranges of −50‰ to −11‰ reported for freshwater plants, albeit one must consider that this range is quite wide (see Figure 2.5, p. 37) (Keeley & Sandquist, 1992).

Table 7.1
Carbon and nitrogen isotope ratios for all food samples.

Location (Ethnic Group)	Category	Sample	$\delta^{13}C$ (‰) mean (SD)	*n*	$\delta^{15}N$ (‰) mean (SD)	*n*	C/N ratio
Layeni (El Molo)	C_4 plant	Sorghum	−12.1	2	+3.6	2	NA
		Maize	−11.3 (0.1)	3	+4.8 (0.1)	3	NA
	Fish	Tilapia	−13.1 (0.0)	3	+3.8 (0.0)	3	3.0
		Golefish	−16.4 (0.0)	3	+9.3 (0.1)	3	3.0
		Flatfish	−24.0 (0.1)	3	+4.9 (0.1)	3	3.2
		Catfish	−21.9 (0.1)	3	+9.7 (0.1)	3	3.0
		Mudfish	−12.9 (0.0)	3	+3.9 (0.1)	3	3.0
Nakurio (Turkana)	C_3 plant	Doum palm fruit	−25.9 (0.2)	3	+4.0	2	NA
	C_4 plant	Sorghum	−12.1 (0.8)	3	+7.4	2	NA
Webuye (Luhya)	C_3 plant	Beans (legume)	−30.5 (0.1)	3	+5.2	2	NA
		Cow peas (legume)	−27.6 (0.3)	3	+2.4	2	NA
	C_4 plant	Maize	−12.0 (0.4)	3	+1.1	2	NA
Port Victoria (Luhya and Luo)	C_3 plant	Soya beans (legume)	−28.8 (0.3)	3	+13.2	2	NA
		Beans (legume)	−26.6	2	+7.2 (1.5)	3	NA
		Cassava [a]	−24.1 (0.1)	3	—	—	NA
		Rice	−28.5 (0.1)	3	+15.7	2	NA
		Cow peas (legume)	−27.5 (0.3)	3	+10.2	2	NA
		Green peas (legume)	−26.5 (0.2)	3	+2.9 (1.0)	3	NA
	C_4 plant	Sorghum	−12.2	2	+9.1 (1.2)	3	NA
		Finger millet	−13.0 (0.1)	3	+7.7	2	NA
		Yellow maize	−11.9 (0.3)	3	+5.9	2	NA
	Fish	*Dagaa*	−18.4	2	+9.4	2	3.4
		Dagaa (whole)	−19.4 (0.1)	3	+11.0 (0.1)	3	3.7
		Ofulu	−17.7 (0.1)	3	+9.1 (0.1)	3	3.4
		Ofulu (whole)	−19.6 (0.0)	3	+9.8 (0.1)	3	3.6
		Nile perch	−15.9 (0.1)	3	+8.8 (0.0)	3	3.1
		Helicopter fish	−18.2 (0.1)	3	+7.0 (0.0)	3	3.2
		Esebu	−20.9 (0.0)	3	+5.6 (0.1)	3	3.2
		Tilapia	−16.1 (0.1)	3	+4.9 (0.1)	3	3.1

Notes. *n* = number of replicates; SD = standard deviation. No Shapiro-Wilk tests were conducted since replicate number is low.
[a] Isotopic analysis to determine the $\delta^{15}N$ of cassava failed systematically, possibly as a result of its low nitrogen content. Consequently, this datum point is not plotted in Figure 7.1.

Turning to the nitrogen isotope results, the plant values range from around 0.0‰ to +15.0‰. This range of values is wider than that suggested for either leguminous (−2‰

to +2‰) or non-leguminous plants (0‰ to +6‰) (Pate, 1994; Sharp, 2007). However, this is not entirely surprising considering the many factors that affect the $\delta^{15}N$ values of plants (Chapter 2, p. 39) (Robinson, 2001). In particular, high nitrification rates in soil, within a wet agriculture context, could explain the very high $\delta^{15}N$ values of rice and soya beans (Reddy, Patrick, & Lindau, 1989; Robinson, 2001), whereas low precipitation and high temperatures in some areas could explain the remaining high $\delta^{15}N$ values (Amundson et al., 2003; Hartman, 2011; Hülsemann et al., 2015). On the other hand, the nitrogen isotopic values of the fish pool more closely, ranging from +3.8‰ to +11.0‰, not reaching values as low as those found in plants. Nevertheless, within these samples, one spots two disparate groups, one around +5‰ and another around +10‰. Interestingly, the tilapia, one of the fish in the first group, is pelagic and feeds mostly on higher plants and algae, although it is also known to eat insects and crustaceans (Awaïss et al., 2010), whereas in the second group, the *Dagaa* feed on zooplankton and surface insects (FishBase team RMCA & Geelhand, 2016), and the Nile perch is a "voracious predator" (Azeroual et al., 2010, p. 4). This suggests that two groups correspond to a difference in diet, where the first group is one or more trophic levels below the second. Note that it was not possible to obtain dietary information for most of the other fish species, due to a lack of reliable ecological data or to uncertainty in the species classification (see Table 4.2, p. 90).

In contrast, no clear trophic shift is detectable between the fish foods in relation to the plants analysed. Two, not mutually exclusive, explanations for this phenomenon are that (1) the analysed plants are enriched in ^{15}N, which has been discussed above, and (2) the base of the food chain in the two lakes (Lake Turkana and Lake Victoria) is depleted in ^{15}N, an effect that is then relayed up the food chain (Campbell, Hecky, & Wandera, 2003; Gownaris, Pikitch, Ojwang, Michener, & Kaufman, 2015). Consistent with the latter, the isotopic values obtained for aquatic fauna in this study are close to those found in other studies. Among these, three studies investigated the carbon and nitrogen isotopic signature of some of the same fish species in Lake Turkana and Lake Victoria included in this study. From Lake Turkana, Kiura (2005, p. 292) analysed the bone collagen of catfish ($\delta^{13}C$: mean = −12.7‰, SD = 2.9‰, n = 5; $\delta^{15}N$: mean = +5.3 ‰,

SD = 1.5‰, *n* = 5) and mudfish ($\delta^{13}C$: mean = −13.4‰, SD = 5.3‰, *n* = 9; $\delta^{15}N$: mean = +5.3 ‰, SD = 2.6‰, *n* = 9), whereas Gownaris et al. (2015, p. 10) analysed muscle samples from tilapia ($\delta^{13}C$: mean = −17.2‰, SD = 1.7‰, *n* = 114; $\delta^{15}N$: mean = +5.0‰, SD = 2.6‰, *n* = 114) in the same lake. From Lake Victoria, Campbell et al. (2003, p. 246) analysed muscle samples, namely tilapia ($\delta^{13}C$: mean = −18.1‰, SD = 0.8‰, *n* = 11; $\delta^{15}N$: mean = +5.9‰, SD = 0.8‰, *n* = 11), Nile perch ($\delta^{13}C$: mean = −18.9‰, SD = 0.5‰, *n* = 12; $\delta^{15}N$: mean = + 8.0‰, SD = 1.1‰, *n* = 12), *Dagaa* ($\delta^{13}C$: mean = −17.0‰, SD = 0.7‰, *n* = 8; $\delta^{15}N$: mean = +8.1‰, SD = 0.6‰, *n* = 8), and *Ofulu* ($\delta^{13}C$: mean = −18.3‰, *n* = 2; $\delta^{15}N$: mean = +9.2‰, *n* = 2).

A final point of interest regarding food results is the difference in isotope values between the fishes (*Ofulu* and *Dagaa*) when just the bone was analysed versus when analysed whole. First, the difference in carbon isotope results is easy to explain, considering that collagen is enriched in ^{13}C relative to other tissues due to its high glycine content (Chapter 2, p. 43) (Bocherens & Drucker, 2003; DeNiro & Epstein, 1978; O'Connell et al., 2001). Second, the discrimination factor in fish $\delta^{15}N$ depends on tissue analysed and on diet isotopic ratios in such a way that, for higher diet isotopic ratios, the difference between the whole body and the individual tissues is also higher (Caut et al., 2009).

7.1.2. HAIR RESULTS

Figure 7.2 and Table 7.2 inform on the isotopic results of the human hair samples. The assessment of these isotopic results was based on (1) a C/N ratio within the range of 3.0 to 3.8, and (2) replicate measurement errors below 0.3‰ for $\delta^{13}C$ and 0.4‰ for $\delta^{15}N$. Based on these criteria, nine samples were removed from the data set, resulting in 134 hair samples. In particular, samples belonging to the individuals TK10, TK12, TK16, and TK17 exceeded the 3.8 maximum C/N ratio, whereas the individuals TK27, WB16, WB26, PT48, and PT50 exceeded the acceptable replicate measurement errors. Unfortunately, in these cases, it was impossible to repeat the analysis since the hair samples were short and no sample was left after the initial cycle of analysis. It is worth

mentioning that all nine of these samples belonged to women, and from these, three were breastfeeding, although, at this point, it is unclear if this fact might have any relevance.

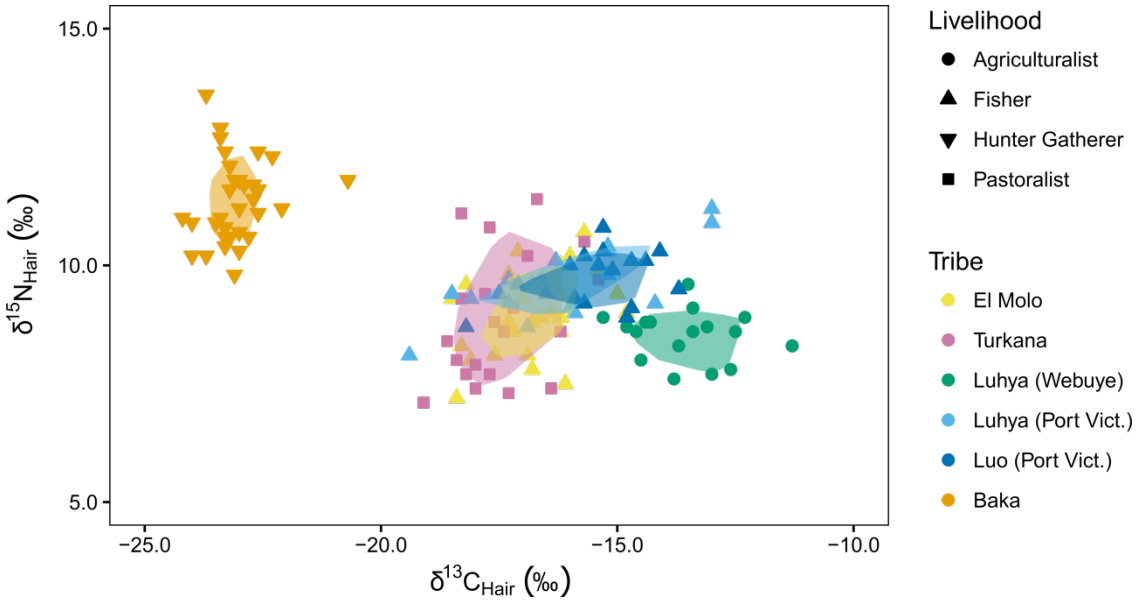


Figure 7.2 Carbon and nitrogen isotope ratios of hair samples, by livelihood and population. Shaded areas represent bagplots (two dimensional boxplots) that contain 50% of the data distribution.

Table 7.2
Summary of hair isotope results, by population.

		δ¹³C (‰)		δ¹⁵N (‰)	
		mean	*median*	mean	*median*
Population	*n*	(SD)	*(IQR)*	(SD)	*(IQR)*
El Molo	27	−16.8 (1.0)	*−16.8 (1.1)*	+9.0 (0.9)	*+9.0 (1.2)*
Turkana	21	−17.5 (1.0)	*−17.7 (1.3)*	+8.9 (1.3)	*+8.6 (2.0)*
Luhya (Webuye)	17	−13.6 (1.0)	*−13.5 (1.4)*	+8.5 (0.5)	*+8.6 (0.5)*
Luhya (Port Victoria)	18	−16.1 (1.8)	*−16.1 (2.0)*	+9.7 (0.8)	*+9.6 (0.9)*
Luo (Port Victoria)	18	−15.6 (1.2)	*−15.4 (1.3)*	+9.7 (0.6)	*+9.8 (0.8)*
Baka	33	**−23.1 (0.6)**	*−23.2 (0.6)*	+11.4 (0.9)	*+11.2 (1.1)*

Notes. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, p<0.05, Appendix D, p. 370).

From a visual analysis, the Baka hunter gatherers have the lowest carbon isotopic signal, the highest nitrogen isotopic signal, and are thus, the most distinct group. Following this, the agriculturalists Luhya, from Webuye, have the highest carbon isotopic signal, and distinguish themselves from the remaining groups. The remaining four groups show a considerable overlap in both the carbon and nitrogen isotopic signal. However, the pastoralist Turkana, have the lowest carbon isotopic signal among

them, and the widest nitrogen isotopic range. On the other hand, all fisher groups exhibit considerable overlap, albeit the El Molo pool closer to the Turkana.

7.1.3. NAIL RESULTS

In parallel with the previous section, Figure 7.3 and Table 7.3 inform on the isotopic results of nail samples. Once again, the assessment of these isotopic results was based on (1) a C/N ratio within the range of 3.0 to 3.8, and (2) replicate measurement errors below 0.3‰ for $\delta^{13}C$ and 0.4‰ for $\delta^{15}N$. Based on these criteria, one sample was removed from the data set, resulting in 82 nail samples. The sample belonging to the individual PT45 exceeded the acceptable replicate measurement error. As for hair, no material was left for a second cycle of analysis in this sample.

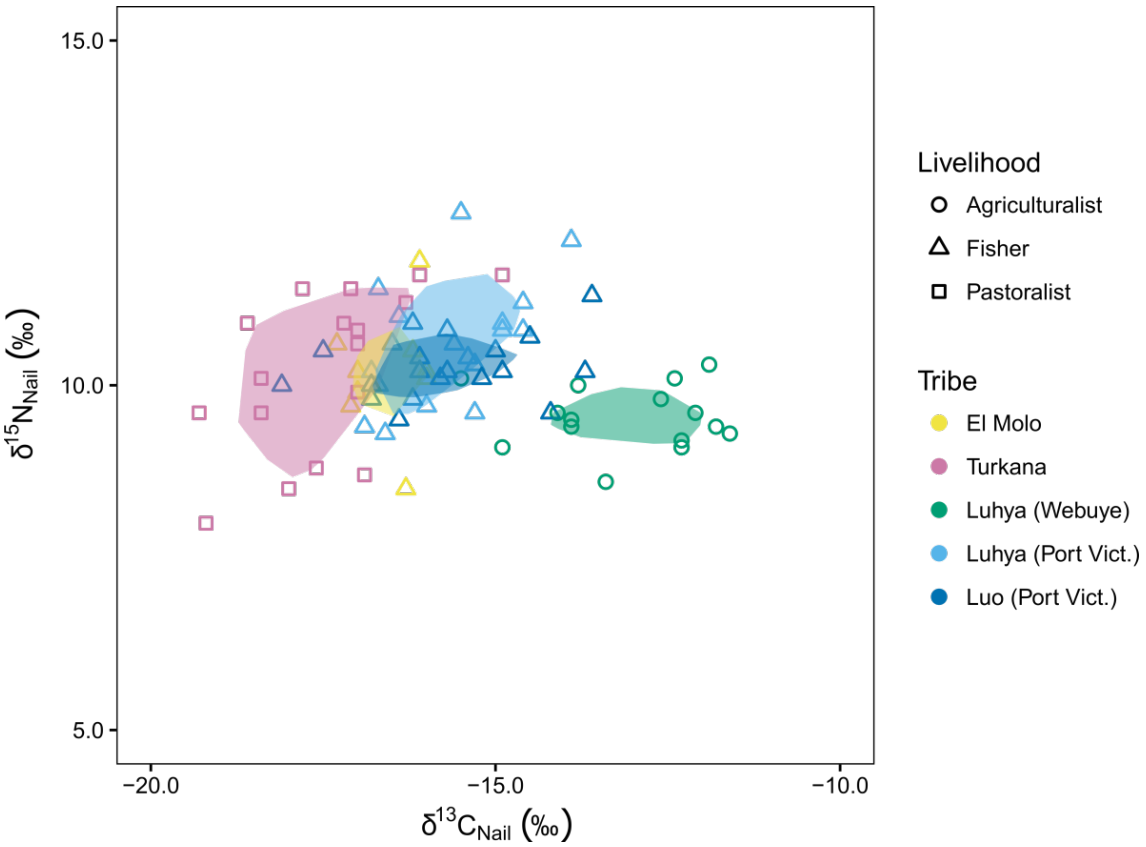


Figure 7.3 Carbon and nitrogen isotope ratios of nail samples, by livelihood and population. Shaded areas represent bagplots (two dimensional boxplots) that contain 50% of the data distribution.

As reported in Chapter 4 (Table 4.1, p. 89), no nail samples were collected from the Baka. Therefore, Figure 7.3 and Table 7.3 present the isotopic signal of the five other groups. The same patterns observed in the hair results are present in the nail data.

First, the Luhya (Webuye) form a distinct group, with the highest carbon isotopic signal, and the lowest nitrogen isotopic signal, whereas the other groups pool closer together. The Turkana have the lowest carbon isotopic signal and a wide range of nitrogen isotopic values (overlapping with the agriculturalists Luhya). Finally, the three fisher populations overlap substantially. However, the El Molo do not overlap as much with the Turkana as they do in hair.

Table 7.3
Summary of nail isotope results, by population.

		$\delta^{13}C$ (‰)		$\delta^{15}N$ (‰)	
		mean	*median*	mean	*median*
Population	*n*	(SD)	*(IQR)*	(SD)	*(IQR)*
El Molo	9	−16.6 (0.5)	*−16.3 (0.9)*	+10.2 (0.9)	*+10.2 (0.6)*
Turkana	20	−17.5 (1.1)	*−17.1 (1.4)*	**+10.4 (1.1)**	*+10.9 (1.8)*
Luhya (Webuye)	15	−13.1 (1.2)	*−12.6 (1.7)*	+9.5 (0.5)	*+9.5 (0.7)*
Luhya (Port Victoria)	17	−15.6 (0.9)	*−15.5 (1.6)*	+10.6 (0.9)	*+10.6 (0.8)*
Luo (Port Victoria)	21	−15.8 (1.2)	*−16.1 (1.5)*	+10.3 (0.4)	*+10.2 (0.5)*

Notes. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
Values in **bold** indicate a sampling distribution that deviates significantly from normality (Shapiro Wilk test, $p<0.05$, Appendix D, p. 372).

7.1.4. BREATH RESULTS

As before, Figure 7.4 and Table 7.4 inform on the isotopic results of breath samples. For this body pool, carbon isotope results are plotted against the time of day the sample was collected because (1) nitrogen is absent from breath samples, and (2) breath $\delta^{13}C$ varies throughout the day according to the proportions of protein, carbohydrates, and fat being oxidised for energy (Chapter 2, p. 43) (DeNiro & Epstein, 1977; Hedges et al., 2009; Hedges & van Klinken, 2000; O'Brien, 2015; Schoeller et al., 1984). As in plant samples, it is difficult to evaluate the contamination of breath samples. Instead, one can only look at the concentration of CO_2 in the sample. Accordingly, and as detailed in Chapter 4, EM10 and TK15 were removed from the data set because they had insufficient CO_2 (< 0.1%) for analysis, whereas PT27, PT44, and PT50 (from the Luo), as well as BK02, BK09, BK22, and BK30 were earmarked as having low CO_2 concentration (< 3%). Thus, the final data set included 184 breath samples.

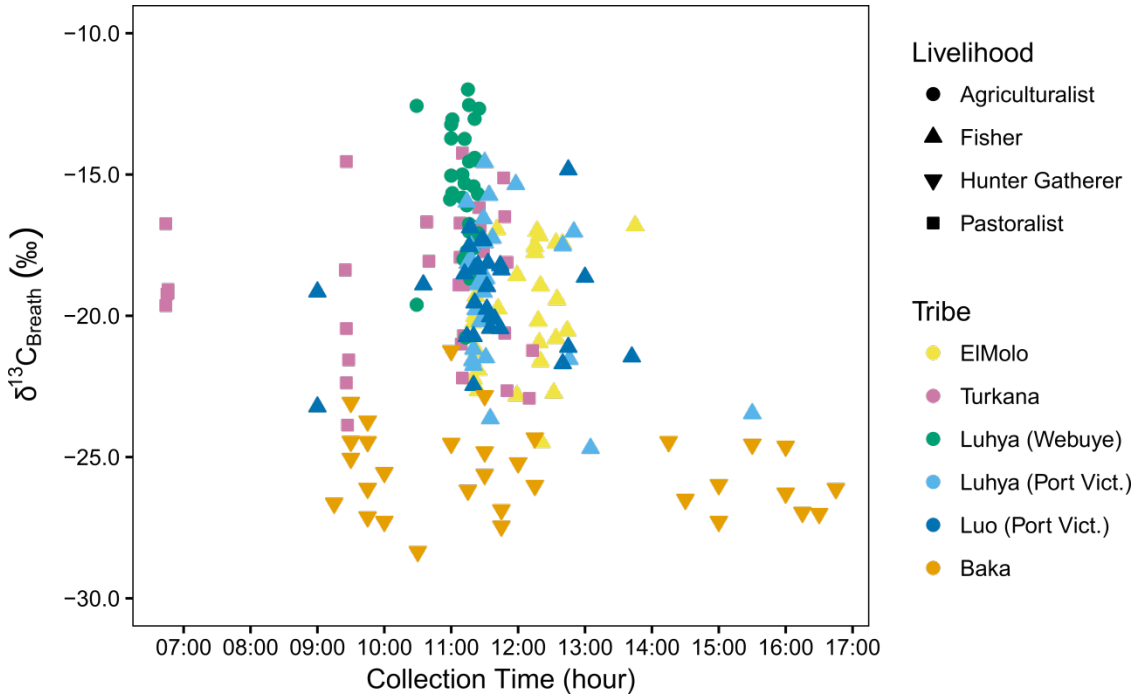


Figure 7.4 Carbon and nitrogen isotope ratios of breath samples, by livelihood and population, and plotted against the time of collection.

Table 7.4
Summary of breath isotope results, by population.

		δ13C (‰)	
		mean	*median*
Population	*n*	(SD)	*(IQR)*
El Molo	29	−19.9 (2.1)	*−20.0 (3.5)*
Turkana	32	−19.0 (2.5)	*−18.9 (4.0)*
Luhya (Webuye)	32	−15.6 (2.2)	*−15.5 (3.4)*
Luhya (Port Victoria)	30	−19.1 (2.5)	*−18.8 (3.5)*
Luo (Port Victoria)	29	−19.3 (1.8)	*−19.0 (2.2)*
Baka	32	−25.5 (1.6)	*−25.8 (2.2)*

Notes. *n* = number of individuals; SD = standard deviation; IQR = interquartile range.
No sampling distribution deviated significantly from normality
(Shapiro Wilk test, p<0.05, Appendix D, p. 374).

Furthermore, Figure 7.4 and Table 7.4 show similar patterns to those observed in the hair and nail isotopic data. In detail, the Baka have the lowest carbon isotopic signal, the Luhya (Webuye) have the highest signal, and the remaining groups overlap considerably, with the Turkana having the lowest carbon isotopic values among these. In addition, Table 7.4 reports higher standard deviations and interquartile ranges than previous tables for hair and nail. To explain this result, one must understand that the isotopic signal of breath responds more promptly to dietary changes than the isotopic

signal of hair or nail, simply because in the latter the body isotope pools act as buffers against short-term fluctuations in diet (Chapter 2, p. 43) (Ayliffe et al., 2004; O'Connell & Hedges, 1999b). Thus, one expects this higher variability within a population in breath isotopic results compared to the other tissues.

As mentioned in Chapter 4, p. 102, breath collection ideally took place at 11 am, or at least three hours after a meal. However, as is clear from Figure 7.4, even though most of the results group around the same time of day (between 10 am and 12 pm), it was not always possible to keep this collection standard. This is particularly true for the Baka.

In fact, the time of day is significantly correlated with overall breath $\delta^{13}C$, $\tau = -0.13$, $p < 0.01$. Note that this correlation reveals a weak negative relationship ($|\tau| \leq 0.30$), and that a non-parametric Kendall's tau (τ) correlation was used, because overall breath $\delta^{13}C$ follows a non-normal distribution ($W = 0.98$, $p = 0.007$) (Field et al., 2012, p. 225). This correlation, however, might result from the differences in breath $\delta^{13}C$ between populations, and indeed, the relationship does not hold when looking at each population individually. In other words, no significant correlation is found between the time of day and breath $\delta^{13}C$ for the El Molo, $r = 0.35$, $p = 0.20$, for the Turkana, $r = 0.02$, $p = 0.90$, for the Luhya (Webuye), $r = -0.04$, $p = 0.90$, for the Luhya (Port Vict.), $r = -0.36$, $p = 0.20$, for the Luo (Port Vict.), $r = 0.11$, $p = 0.85$, or for the Baka, $r = -0.20$, $p = 0.53$. In this case, a Pearson's correlation was used since the distribution of breath $\delta^{13}C$ within groups was normally distributed (Table 7.4). Furthermore, to correct for multiple comparisons, this test used the Benjamin-Hochberg False Discovery Rate (FDR), which controls the familywise error rate (Field et al., 2012, p. 430). Notwithstanding, some of these correlations denote a medium effect ($|r| \geq 0.30$) (Field et al., 2012, p. 58); however, it is possible that the sample sizes are too small or that the data are too close together to detect this effect. Overall, the tests suggest that breath $\delta^{13}C$ decreases throughout the day, but that this effect is small and does not hold across populations. This means that the effort taken in standardizing time of collection was

largely successful, but that a residual effect of time of day on breath $\delta^{13}C$ remains, which must be taken into consideration when further analysing these data.

7.2. COMPARISON ACROSS POPULATIONS

Having discussed the quality of the isotope results, one may begin the scrutiny of these data by comparing the different populations included in the study. In this section, these statistical analyses are separated according to the different body pools sampled: hair, nail, and breath. Conversely, the results from food samples were not examined statistically due to the low sample size. Appendix D includes the full details of all statistical analysis.

7.2.1. STATISTICAL CONSIDERATIONS

At this point, one must consider the statistical reasoning that underpins the analysis. In the case of breath, there is one dependent variable, $\delta^{13}C$, and one independent variable, population affiliation. On the other hand, for hair and nail, there are two dependent variables, $\delta^{13}C$ and $\delta^{15}N$, and the same independent variable, population affiliation. Therefore, the adequate test for the first case is an analysis of variance (ANOVA), whereas a multivariate analysis of variance (MANOVA) is adequate to investigate the second case, since it takes into consideration any relationship between the two dependent variables.

For the first analysis, one must consider the effect of the weak negative correlation between the time of collection and breath $\delta^{13}C$. To achieve this, the correct approach is an analysis of covariance (ANCOVA), with time of collection as covariate. Fortunately, no distributional or heteroscedasticity problems arose (Table 7.4, Appendix D, p. 387), since the current robust ANCOVA in R is limited to two categories, and cannot deal with the 6 populations included in this study (Field et al., 2012, p. 448; Mair & Wilcox, 2017, p. 24).

Furthermore, MANOVAs are only required if the dependent variables are correlated moderately, i.e. a null or weak correlation between such variables indicates that no relationship exists and that one may proceed with univariate analysis (although considerable debate exists about this side of the argument), whereas a very strong correlation indicates that the variables are essentially the same and should be combined (Field et al., 2012, p. 698). Here, hair $\delta^{13}C$ was significantly correlated with hair $\delta^{15}N$, $\tau = -0.27$, $p < 0.001$, while nail $\delta^{13}C$ was not significantly correlated with nail $\delta^{15}N$, $\tau = -0.07$, $p = 0.38$. Note that, as before, a non-parametric Kendall's tau (τ) correlation was used, because both overall hair and nail $\delta^{13}C$ followed a non-normal distribution (Appendix D, p. 374). Despite finding only a weak negative correlation for hair ($|\tau| \geq 0.3$), and no correlation for nail, and in view of the debate around such correlations, this work erred on the side of caution, and proceeded with MANOVA for both hair and nail. In addition, a robust MANOVA was used whenever the test assumptions were not met (Field et al., 2012, p. 717). In detail, this test used robust estimations of multivariate location and scatter, calculated through the minimum covariance determinant (MCD) estimator (Todorov & Filzmoser, 2009).

To follow up significant MANOVAs, this study used both univariate ANOVAs and linear discriminant analysis (LDA). These tests have two different, complementary goals: whereas the first checks how an individual variable differs between groups, the second checks which linear combination of the two variables leads to maximum group separation. Once again, robust versions of these tests were used whenever the test assumptions were not met. More exactly, a robust ANOVA with 5% trimmed means and 999 bootstrap samples was used for univariate tests (Field et al., 2012, pp. 440–441; Mair & Wilcox, 2017, p. 8), whereas robust discriminant analysis were done using both the MCD estimator (similarly to the robust MANOVA) and the minimum volume ellipsoid (MVE) (Todorov & Filzmoser, 2009).

Subsequently, any significant ANOVAs (or the ANCOVA) were followed by *post hoc* tests. For these, two types of tests were used, (1) Tukey honest significant difference (HSD) tests followed standard ANOVAs, and (2) robust pairwise t-tests with 20%

trimmed means, 5000 bootstrap samples, and an FDR correction, followed robust ANOVAs (Field et al., 2012, p. 432).

Finally, and whenever possible, this study reports effect sizes for all tests conducted. In the case of standard MANOVAs, the effect size is eta squared (η^2), which is equivalent to r^2, while for standard ANOVAs, the appropriate effect size is omega squared (ω^2), an unbiased measure of eta squared. In addition, the effect size of the ANCOVA is partial omega squared (partial ω^2), which looks at the proportion of variance that is not explained by other variables in the analysis (Field et al., 2012, p. 454–455; 491). To the best of my knowledge, no measure of effect size exists for robust MANOVAs (Mair & Wilcox, 2017; Todorov & Filzmoser, 2009), whereas the effect size for robust ANOVA is xi (ξ), which is based on Cohen's *d* (Wilcox, 2012, p. 166–169; 294). Nevertheless, all the above tests address a general hypothesis, and so the most relevant effect sizes are those for the pairwise comparisons. Here, the correlation *r* is used for standard *post hoc* tests, while xi (ξ) is used for robust *post hoc* tests (Field et al., 2012, p. 455; Mair & Wilcox, 2017, p. 6). Lastly, no simple and computationally efficient effect size exists yet for LDA (Klaus, 2013).

7.2.2. HAIR DATA

The combined $\delta^{13}C$ and $\delta^{15}N$ hair results were tested using robust tests, considering (1) the presence of both uni- and multivariate outliers, (2) the lack of multivariate normality in the Baka group, and (3) the lack of uni- and multivariate equality of variances (Table 7.2, Appendix D, p. 376-378) (Field et al., 2012, p. 717).

Using Wilk's lambda statistic, the robust MANOVA showed that there was a significant effect of the population on the carbon and nitrogen isotopic results, $\Lambda = 0.04$, $X^2(9.07) = 275.7$, $p < 0.001$.

Following this significant result, robust ANOVAs were used for carbon and nitrogen isotope values, separately. These tests found a significant effect of population on both hair $\delta^{13}C$, $F_t = 481.5$, $p < 0.001$, $\xi = 0.95$, and hair $\delta^{15}N$, $F_t = 42.8$, $p < 0.001$, $\xi = 0.79$.

Table 7.5

Robust post hoc *results for hair $\delta^{13}C$ and $\delta^{15}N$. The cells above the diagonal represent the pairwise comparisons between groups for hair $\delta^{15}N$, whereas the cells below the diagonal represent the pairwise comparisons between groups for hair $\delta^{13}C$. Values in each cell are the robust effect size (ξ) for each pairwise comparison. Shaded cells are significant comparisons ($p < 0.05$).*

		Hair $\delta^{15}N$					
		El Molo	Turkana	Luhya (Webuye)	Luhya (Port Vict.)	Luo (Port Vict.)	Baka
Hair $\delta^{13}C$	El Molo	—	0.2	*0.4*	**0.5**	**0.6**	**0.9**
	Turkana	**0.5**	—	0.1	**0.5**	**0.6**	**0.9**
	Luhya (Webuye)	**1.0**	**1.0**	—	**0.9**	**0.9**	**1.0**
	Luhya (Port Vict.)	*0.4*	**0.6**	**0.8**	—	0.1	**0.9**
	Luo (Port Vict.)	**0.7**	**0.9**	**0.9**	0.3	—	**0.9**
	Baka	**0.9**	**0.9**	**0.9**	**1.0**	**1.0**	—

Notes. Interpretation of ξ: small = 0.15, *medium* = 0.35, and **large** = 0.50 (Wilcox, 2012, p. 169).

Subsequently, Table 7.5 informs on the robust pairwise t-tests that followed significant robust ANOVAs. This table shows that all significant comparisons have large effect sizes, whereas non-significant comparisons have small effect sizes, with the exception of the comparison between the El Molo and Luhya (Port Vict.) for carbon, and the comparison between the El Molo and Luhya (Webuye), which have a medium effect size for nitrogen isotopic results. In addition, almost all pairwise comparisons for carbon are significant. In detail, the Turkana, the Luhya (Webuye), and the Baka all distinguish themselves from every other group. In contrast, the two groups from Port Victoria (Luhya and Luo) do not differ in their hair $\delta^{13}C$ signature, and neither do the El Molo and the Luhya (Port Victoria). In contrast, there are less significant pairwise comparisons in nitrogen. In detail, the Baka still differ from all other groups, while the Luhya and the Luo from Port Victoria again do not differ from each other. Conversely, this test failed to distinguish between the El Molo, the Turkana, and the Luhya (Webuye). Therefore, for hair $\delta^{13}C$, *post hoc* tests detect four groups: more depleted in ^{13}C, the Baka, followed by the Turkana, then the El Molo, very similar to the Luhya

(Port Victoria), whom, in turn, are similar to the Luo (Port Victoria), and finally the Luhya (Webuye). Then, for hair $\delta^{15}N$, *post hoc* tests find three groups: more depleted in ^{15}N, the Turkana, the El Molo, and the Luhya (Webuye), followed by the Luhya and the Luo from Port Victoria, and finally, the Baka with the higher $\delta^{15}N$ values.

It is intriguing to realise that the closer these populations are to each other geographically, the closer they are isotopically. For instance, the Luo and the Luhya from Port Victoria cannot be distinguished based on the carbon or nitrogen isotopic signature. Similarly, the El Molo and the Turkana, which inhabit opposite margins of Lake Turkana, have similar nitrogen isotope results, albeit they differ in carbon isotope results. Nevertheless, for each isotope, one of the not significant comparisons came from populations that are far from each other in space (El Molo versus Luhya (Port Vict.) for carbon; and El Molo versus Luhya (Webuye) for nitrogen). However, they have medium effect sizes, suggesting that populations differ but that this difference is not big enough for these tests to detect it. Finally, the Luhya (Webuye) versus the Turkana are the only pairwise comparison in nitrogen that is not significant, has a small effect size, and whose members are quite far away from each other in space. This might simply reflect a true similarity in diet. This is consistent with the shift from a fully pastoralist diet to a largely maize-based diet among the sedentary Turkana, thus approaching that of the agriculturalist Luhya from Webuye.

Following *post hoc* tests, robust LDA using MVE extracted two discriminant functions. The first explained 99.0% of the variance in the combined carbon and nitrogen isotopic values, whereas the second explained 1.0%. The coefficients of the first discriminant function differentiated hair $\delta^{13}C$ ($b = -1.2$) from $\delta^{15}N$ ($b = 0.7$), whereas the second discriminant function discriminated hair $\delta^{13}C$ ($b = -0.3$) and $\delta^{15}N$ ($b = -1.0$) in the same direction. The discriminant function plots (Figure 7.5) show that the first discriminates the Baka group, and to some extent the Luhya (Webuye), while the second function does not seem to differentiate any group particularly well. These results are maintained when using other robust statistics, such as MCD (Appendix D, p. 381). In addition, when using these functions to predict group affiliation, 35% of the

individuals are wrongly assigned. This misclassification rate is relatively high, although it is still more accurate than simply guessing the observation's group, which would have an error rate of 83% ($1 - {}^{1}/_{6}$). Therefore, the LDA analysis does not add to the previous results, i.e. the combined carbon and nitrogen isotopic signal provides less information about population of origin than each isotope individually.

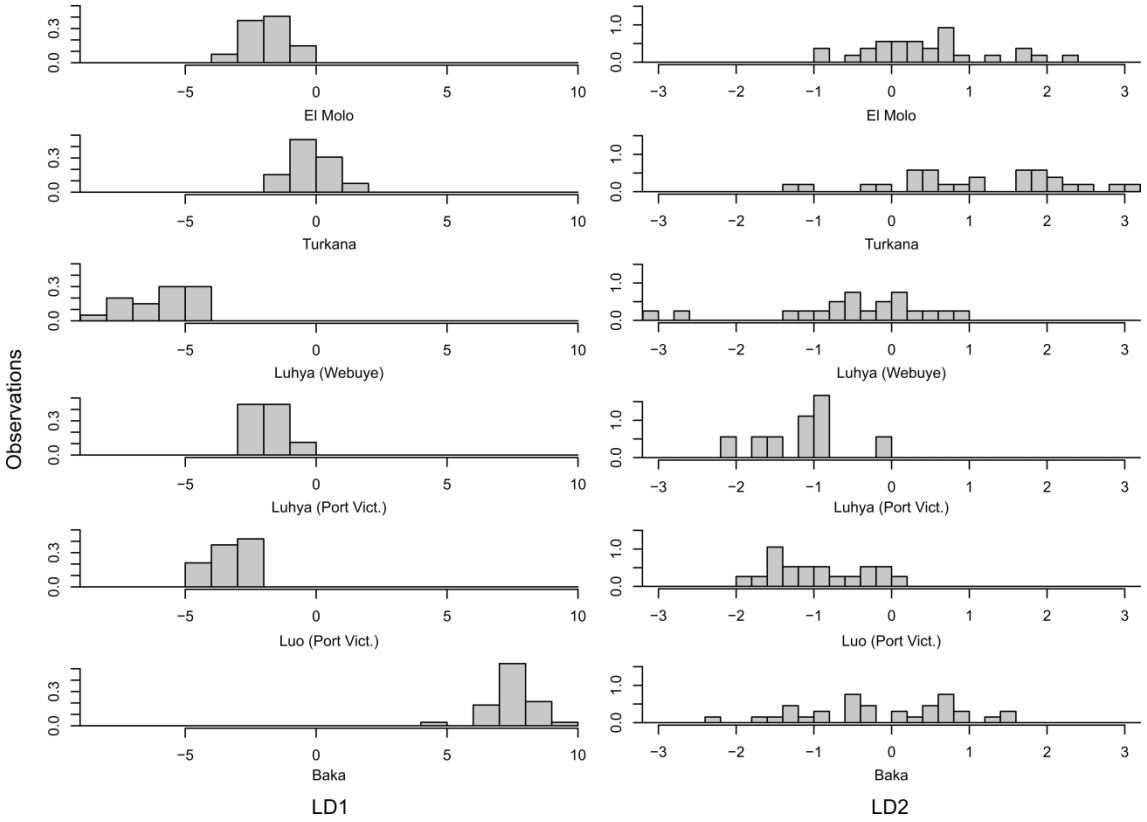


Figure 7.5 Histograms for the observations in each group on the two linear discriminant dimensions of hair (LD1 and LD2). For each dimension, if the histograms of a group do not overlap with the histograms of the other group, then the function discriminates that group well; conversely, if the histograms overlap, then the function fails to discriminate the groups successfully (Appendix D, p. 381).

7.2.3. NAIL DATA

The analysis of nail $\delta^{13}C$ and $\delta^{15}N$ used a combination of standard and robust tests. In detail, a robust MANOVA was used in view of the lack of multivariate equality of variances (Appendix D, p. 382-383). A standard ANOVA was used for carbon, considering that the distribution within groups is normally distributed, and that the variances are equal across groups, whereas a robust ANOVA was used for nitrogen, considering that the data violated both the normality and homocedasticity

assumptions (Table 7.3). Accordingly, standard *post hoc* tests followed the standard ANOVA, whereas robust *post hoc* tests followed the robust ANOVA. Finally, a robust LDA was used, due to the lack of multivariate equality of variances (Appendix D, p. 383).

Using Wilk's lambda statistic, the robust MANOVA showed that there was a significant effect of population of origin on the carbon and nitrogen isotopic results from nails, $\Lambda = 0.22$, $X^2(6.61) = 70.2$, $p < 0.001$.

Subsequently, the standard ANOVA found a significant effect of population on nail $\delta^{13}C$, $F(4, 77) = 38.5$, $p < 0.001$, $\omega^2 = 0.65$, as did the robust ANOVA on nail $\delta^{15}N$, $F_t = 7.48$, $p < 0.01$, $\xi = 0.50$.

Table 7.6 reports on the *post hoc* tests that followed the significant ANOVAs. Note that the Baka are not present in this table because no nails were collected from this population (Table 4.1, p. 89). In the same way as hair results, this table shows that all significant comparisons have large effect sizes. However, in contrast to the hair results, more of the non-significant comparisons have medium effect sizes. In fact, the only small effect size in carbon isotope pairwise comparisons is between the Luo and Luhya (Port Vict.), whereas the nitrogen isotope comparisons report three small effect sizes: the El Molo and the Turkana, the El Molo and the Luo (Port Vict.), and the Turkana and the Luhya (Port Vict.). In addition, the significant *post hoc* tests on the carbon isotope results of nail are also those identified as significant for hair. However, the *post hoc* tests of nail carbon isotopes detect fewer significant pairwise comparisons than the parallel tests carried on hair results. Analysis of the carbon isotope results of nail failed to distinguish between the El Molo and the Turkana, and the El Molo and the Luo (Port Vict.), whereas the equivalent hair analysis did. Otherwise, the two remaining comparisons, El Molo versus Luhya (Port Vict.), and Luo versus Luhya (Port Vict.), are not significant in both tissues. This lower number of significant comparisons might result from the lower sample size of nail, and consequent lower statistical power (for the same five groups, there were 101 isotope results of hair, but only 82 of nail). In fact,

the two comparisons that were significant in hair, but not in nail, involve the group El Molo, whom had the lowest nail sample size (n = 9).

Table 7.6
Post hoc *results for nail: standard for $\delta^{13}C$ and robust for $\delta^{15}N$. The cells above the diagonal represent the pairwise comparisons between groups for hair $\delta^{15}N$, whereas the cells below the diagonal represent the pairwise comparisons between groups for hair $\delta^{13}C$. Values in each cell for carbon pairwise comparisons are the standard effect size,* r, *while values for nitrogen comparisons are the robust effect size, ξ. Shaded cells are significant comparisons (*p < *0.05).*

		Nail $\delta^{15}N$				
		El Molo	Turkana	Luhya (Webuye)	Luhya (Port Vict.)	Luo (Port Vict.)
Nail $\delta^{13}C$	El Molo	—	0.2	**0.7**	*0.4*	0.1
	Turkana	*0.4*	—	**0.7**	0.1	*0.4*
	Luhya (Webuye)	**0.9**	**0.9**	—	**0.9**	**0.9**
	Luhya (Port Vict.)	**0.5**	**0.7**	**0.8**	—	0.4
	Luo (Port Vict.)	*0.3*	**0.6**	**0.7**	0.1	—

Note. Interpretation of *r*: small = 0.10, *medium* = 0.30, and **large** = 0.50 (Field et al., 2012, p. 58); Interpretation of ξ: small = 0.15, *medium* = 0.35, and **large** = 0.50 (Wilcox, 2012, p. 169).

On the contrary, *post hoc* tests in the nitrogen isotopes of nail have quite different results from their equivalents in hair. Indeed, the only significant results simply single out the Luhya (Webuye) as different from all other groups. Furthermore, of the 10 comparisons, only 4 have the same result as their hair equivalent. Both hair and nail distinguish between the Luhya (Webuye) and both the Luhya and the Luo from Port Victoria, but fail to differentiate between the El Molo and the Turkana, and the Luhya and the Luo from Port Victoria. Conversely, nail nitrogen isotopes discriminate the Luhya (Webuye) from both the El Molo and the Turkana, while hair does not. On the other hand, nail cannot differentiate the four combinations between the El Molo and the Turkana, and the Luhya and Luo (Port Vict.), but hair does. It is not entirely clear what is driving these differences in hair and nail $\delta^{15}N$. In this case, the effect sizes do not help in explaining the results, since they are quite different between hair and nail for these comparisons. Nonetheless, of the 6 differing results, 3 involve the El Molo, and another 3, the Turkana. Hence, the lower El Molo sample size might partially

explain these results, but this reasoning cannot be applied to the Turkana (n = 20). A metabolic difference between the two tissues offers an alternative explanation. The section comparing the different body pools delves again into this issue.

In summary, for nail $\delta^{13}C$, *post hoc* tests detect two, maybe three, groups: more depleted in ^{13}C, the Turkana and the El Molo, and then the Luo and the Luhya from Port Victoria (although the two groups from Port Victoria can only be distinguished from the Turkana, and not from the El Molo), and finally the Luhya (Webuye), more enriched in ^{13}C. In turn, for nail $\delta^{15}N$, only two groups are found: the Luhya (Webuye) are more depleted in ^{15}N than all other groups, whom all group together.

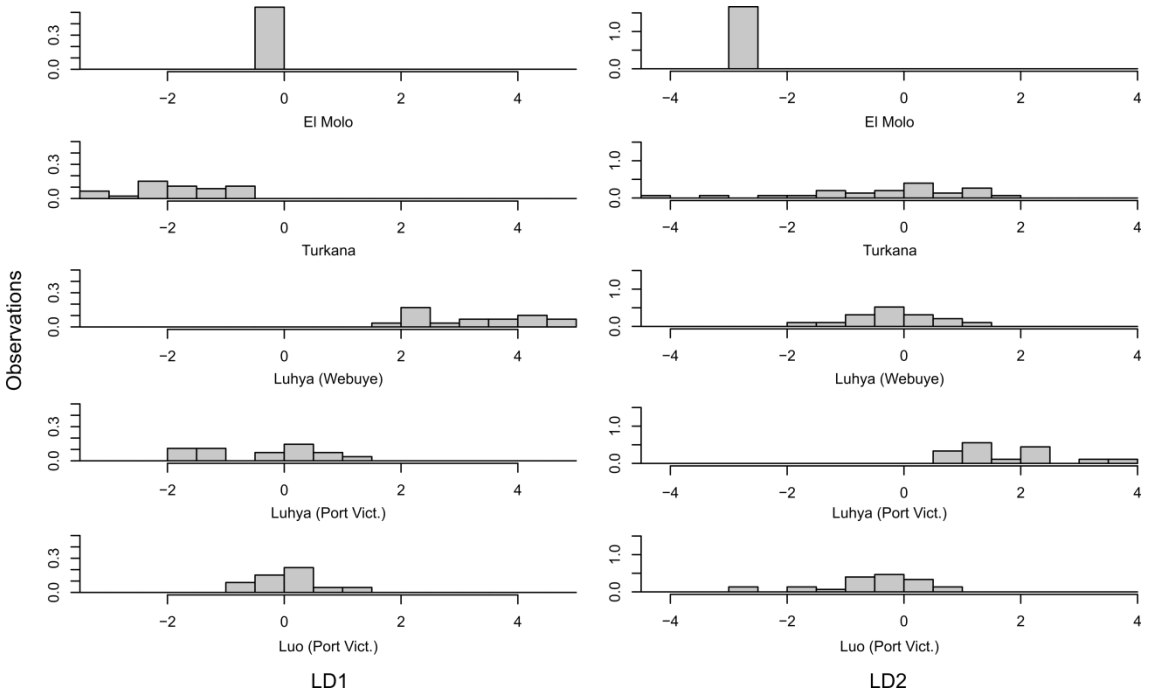


Figure 7.6 Histograms for the observations in each group on the two linear discriminant dimensions of nail (LD1 and LD2). For each dimension, if the histograms of a group do not overlap with the histograms of the other group, then the function discriminates that group well; conversely, if the histograms overlap, then the function fails to discriminate the groups successfully.

As a final point on the analysis of nail, robust LDA using MVE found two discriminant functions. The first explained 95.3% of the variance, whereas the second explained 4.7%. The coefficients of the first discriminant function differentiated nail $\delta^{13}C$ (b = 1.0) from $\delta^{15}N$ (b = −0.9), whereas the second discriminant function affected nail $\delta^{13}C$ (b = 0.3) and $\delta^{15}N$ (b = 1.5) in the same direction. In addition, the discriminant function

plots (Figure 7.6) show that the first discriminates the Luhya (Webuye), and to some extent the Turkana, while the second function seem to differentiate the Luhya (Port Vict.). Thus, the combined carbon and nitrogen isotopes differentiate the Luhya (Webuye) from the other groups, but also partially the Turkana and the Luhya (Port Vict.). Therefore, the LDA analysis of nail confirms the results of the *post hoc* tests, once again suggesting that it is not the interaction between the two isotopes that improves the discrimination between groups. In addition, and once more, at 42%, the error rate in this analysis is high, but not as high as the 80% random classification ($1 - {}^{1}/_{5}$).

7.2.4. BREATH DATA

As laid out at the beginning of this section, the effect of the group on breath $\delta^{13}C$ may be investigated using an ANCOVA, with the time of collection as covariate. However, the time of collection differs significantly across groups, $F(5, 177) = 7.7$, $p < 0.001$, $\omega^2 = 0.16$, which breaks the ANCOVA assumption that the covariate and treatment effect must be independent. This is problematic because, in this situation, the covariate reduces the effect of the group, by explaining some of the variance that would be otherwise attributable to the group (Field et al., 2012, p. 464).

This lack of independence between the time of day and the group is a spurious side effect of data collection, and not a true effect. An option to deal with this issue is to proceed with the ANCOVA, and if the time influences the values of breath $\delta^{13}C$, simply discuss how including the covariate could bias the results. Conversely, if time has no effect, then a simple ANOVA (without considering the time of collection) will suffice to understand the differences between groups in breath carbon.

Thus, in the ANCOVA, the covariate, i.e. time of collection, was not significantly related to breath $\delta^{13}C$, $F(1, 176) = 0.46$, $p = 0.5$, partial $\omega^2 = 0.00$. This suggests that the significant correlation found between breath $\delta^{13}C$ and the time of collection was indeed a result of the influence of population of origin on the isotopic signal. In other words, since the time of collection is not evenly distributed across populations, the correlation was explaining a variance that was, in fact, attributable to the population. Thus, this

work proceeds with the standard ANOVA, which found a significant effect of population on breath $\delta^{13}C$, $F(5, 178) = 71.5$, $p < 0.001$, $\omega^2 = 0.66$.

Table 7.7 presents the results of *post hoc* tests for breath $\delta^{13}C$. As before, all significant comparisons have large effect sizes, whereas non-significant comparisons all have small effect sizes. Overall, the results show that both the Baka and the Luhya (Webuye) differ from all other populations, while the remaining four populations all group together.

Table 7.7
Standard Post hoc *results for breath $\delta^{13}C$. Values in each cell for carbon pairwise comparisons are the standard effect size,* r. *Shaded cells are significant comparisons (*p < *0.05).*

BREATH $\delta^{13}C$	El Molo	Turkana	Luhya (Webuye)	Luhya (Port Vict.)	Luo (Port Vict.)	Baka
El Molo	—					
Turkana	0.2	—				
Luhya (Webuye)	**0.7**	**0.6**				
Luhya (Port Vict.)	0.2	0.0	**0.6**	—		
Luo (Port Vict.)	0.2	0.1	**0.7**	0.1	—	
Baka	**0.8**	**0.8**	**0.9**	**0.8**	**0.9**	—

Notes. Interpretation of *r*: small = 0.10, *medium* = 0.30, and **large** = 0.50 (Field et al., 2012, p. 58).

Compared to the statistical results of hair, less significant differences are found in breath $\delta^{13}C$, with all significant differences in breath also being observed in hair. In further detail, the analyses of breath carbon isotopes fail to distinguish between the El Molo and the Luo (Luhya), between the Turkana and the El Molo, the Turkana and the Luhya (Port Vict.), and the Turkana and the Luo (Port Vict.). Furthermore, when considering these results vis-à-vis the nail results, a similar pattern of significance is present, except that breath carbon cannot distinguish between the Turkana and the Luo (Port Vict.), and the Turkana and the Luhya (Port Vict.), while the nail results can.

This dearth of significant comparisons in breath $\delta^{13}C$ might partly result from the larger variability within groups, which dilutes the differences between groups (Table 7.4). Alternatively, it could indicate a true difference in diet between the groups. If one considers that: (1) hair and nail are biased towards the protein part of the diet, whereas breath records the overall diet (Chapter 2, p. 42), and (2) from the diverging comparisons between breath and the two other body pools, all but one of them involve the Turkana, then this might indicate that the protein part of the Turkana diet differs in its carbon isotopic signal from the rest of the Turkana diet. This aspect of the interpretation will be further explored later in this chapter.

In summary, the *post hoc* tests on breath $\delta^{13}C$ find three groups: more depleted in ^{13}C, the Baka, followed by the El Molo, the Luo (Port Vict.), the Luhya (Port Vict.), and the Turkana, all grouped together, and finally, more enriched in ^{13}C, the Luhya (Webuye).

7.3. COMPARISON TO DEMOGRAPHIC INFORMATION

After examining the differences in isotope results between the different groups of this study, it is also important to analyse the influence on isotopic signatures of other demographic data, namely gender and age. Accordingly, this section splits into these two variables.

7.3.1. STATISTICAL CONSIDERATIONS

To understand whether isotopic signatures change according to an individual gender within populations, but also overall across populations, one may use a factorial ANOVA to compare the main effects of gender and tribe, as well as the interaction between the two, whereas, to understand the same for age, one may use an ANCOVA, considering age as a covariate. Note that, since the multivariate approach (i.e. MANOVA and LDA) failed to produce compelling results in the previous section, it was opted to skip this step and to proceed directly with univariate tests of carbon and nitrogen separately for hair and nail. Furthermore, and as before, if any distributional

and heteroscedasticity problems arose, this work used robust factorial ANOVA with 5% trimmed means and 999 bootstrap samples, albeit no such equivalent is available for 6 group ANCOVA (Field et al., 2012, pp. 440–541; Mair & Wilcox, 2017, p. 8; 11-12; 24).

Afterward, significant interaction effects in factorial ANOVAs (e.g. the effect of gender differed according to group affiliation) were followed by simple effects analysis, i.e. t-tests (robust, if following a robust factorial ANOVA) between the two genders within each population, with an FDR correction to control for the multiple comparisons. Subsequently, a significant main effect of gender (i.e. gender affected isotopic results overall) was interpreted only if no significant interaction effect was found (Field et al., 2012, p. 525).

On the other hand, a significant interaction effect within an ANCOVA test breaks the assumption of homogeneity of regression slopes, which can only be addressed within a multilevel model (Field et al., 2012, p. 467). Provided this situation did not arise, the covariate was interpreted and reported (Field et al., 2012, p. 432; 480).

Finally, the effect size for both factorial ANOVA and ANCOVA is partial omega squared, which is split across the main and interaction effects. No equivalent is reported for the robust versions of the factorial ANOVA (Field et al., 2012, p. 491; 542; Mair & Wilcox, 2017). Following this, the effect size used for standard *post hoc* tests was the correlation *r*, while for robust *post hoc* tests, it was xi (ξ) (Field et al., 2012, p. 455; Mair & Wilcox, 2017, p. 6).

7.3.2. GENDER

Before addressing the main analysis, two points must be made. First, the sample included some pregnant and breastfeeding women (Table 4.1, p. 89). This must be taken into consideration, since the $\delta^{15}N$ of women decreases in these conditions (Fuller et al., 2004b; Fuller, Fuller, Harris, & Hedges, 2006). However, these numbers were small: out of the 99 women participants, only six were pregnant, while 31 were

breastfeeding. Although no effect of pregnancy or breastfeeding is apparent upon visual inspection of the results (Figure 7.7 for hair, Figure 7.8 for nail, and Figure 7.10 for breath), robust factorial ANOVAs were conducted only within the female population. These analyses combined pregnant and breastfeeding women into one category and compared them to the rest of the women, for each combination of body pool and isotope, except for nail (since fewer women provided this tissue, rendering the analysis impossible). None of these analyses produced a significant main or interaction effect, i.e. being pregnant, breastfeeding, or neither, does not influence significantly the isotopic results of women, overall, or within each population (Appendix D, p. 390-398). Therefore, the isotopic results of all women were grouped together when compared to the results of men.

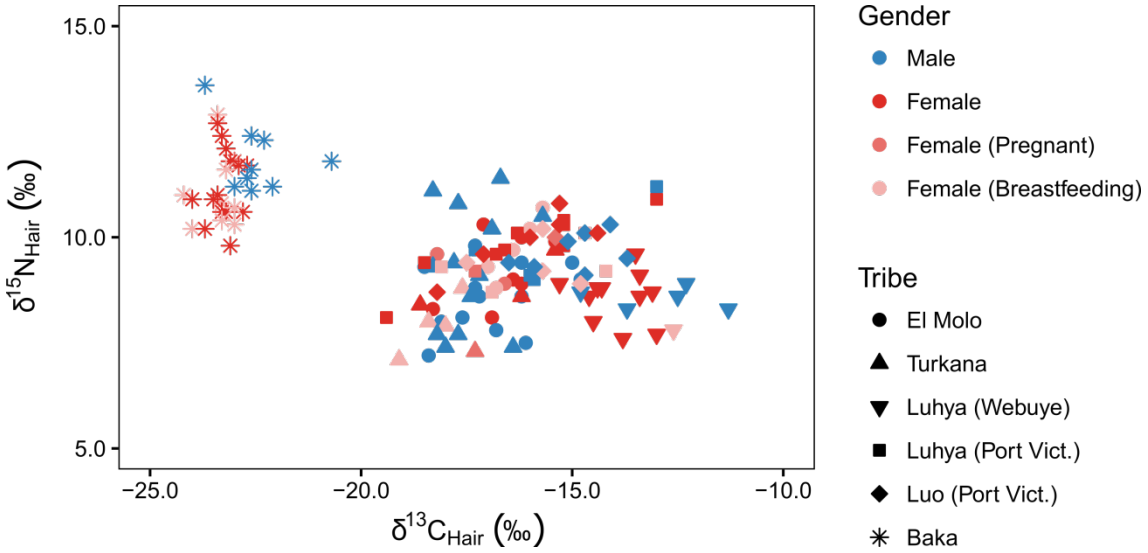


Figure 7.7 Isotopic results of hair samples, by gender and population.

The second point arises from which gender provided which tissues. In detail, from the 99 women, 87 provided hair, but only 22 provided nail, and 10 provided both, whereas from the 85 men, only 47 provided hair, 60 provided nail, and 22 provided both. The particulars of data collection explain this issue: more men than women were concerned about witchcraft (and refused to give hair), although when men gave hair, they were also more willing to give nail, whereas women preferred to give only one tissue. This leads to some gender bias within the analysis, which must be taken into consideration when interpreting any significant results, particularly within populations (Figure 7.7

and Figure 7.8). As an example, among the Luhya (Webuye), 2 men and 16 women gave hair, whereas 13 men and 4 women gave nail.

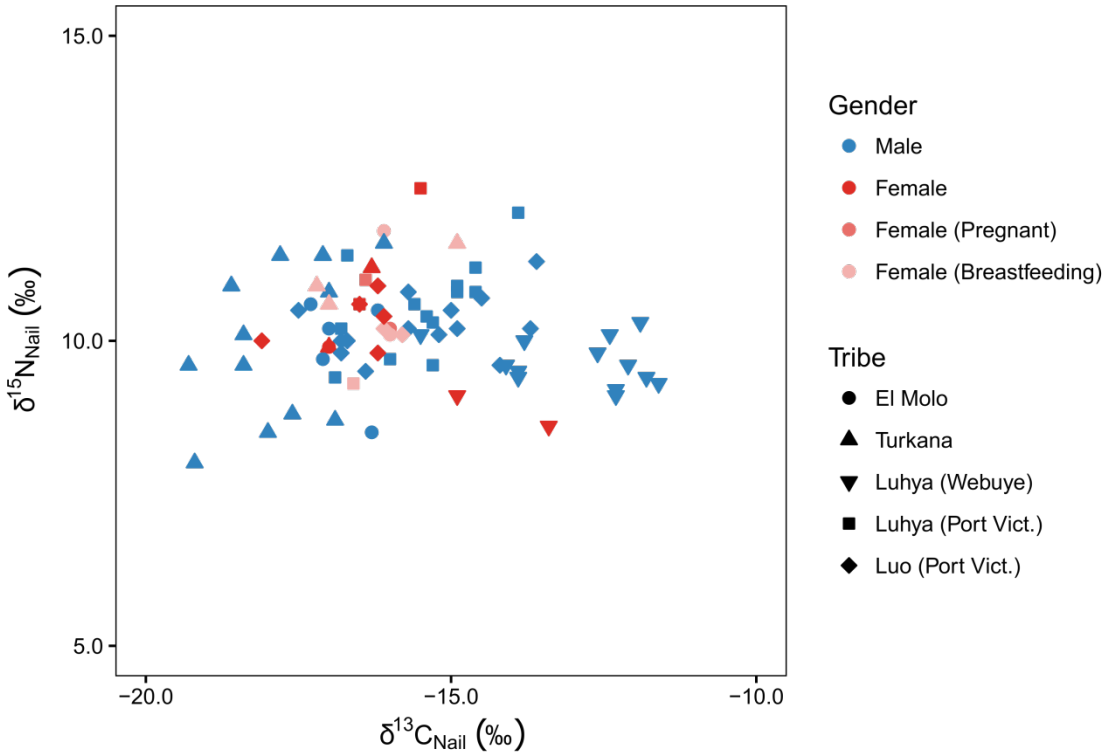


Figure 7.8 Isotopic results of nail samples, by gender and population.

Table 7.8

Results of factorial ANOVAs on the main effects of gender and tribe and the interaction between them on isotopic results, divided by tissue and isotope (carbon and nitrogen). Shaded cells are significant comparisons (p < 0.05).

		Statistic	Probability Value	Effect Size [a]
Interaction Effects				
Hair	$\delta^{13}C$	$F(5, 122) = 1.55$	$p = 0.18$	$\omega^2 = 0.01$
	$\delta^{15}N$	$F_t = 17.19$	$p = 0.06$	NA
Nail	$\delta^{13}C$	$F(4, 72) = 3.98$	$p = 0.006$	$\omega^2 = 0.05$
	$\delta^{15}N$	$F_t = 11.73$	$p = 0.10$	NA
Breath	$\delta^{13}C$	$F(5, 172) = 35.5$	$p = 0.18$	$\omega^2 = 0.001$
Main Effects				
Hair	$\delta^{13}C$	$F(1, 122) = 0.006$	$p = 0.94$	$\omega^2 = 0.001$
	$\delta^{15}N$	$F_t = 1.02$	$p = 0.39$	NA
Nail	$\delta^{13}C$	$F(1, 72) = 0.59$	$p = 0.45$	$\omega^2 = 0.0005$
	$\delta^{15}N$	$F_t = 0.52$	$p = 0.48$	NA
Breath	$\delta^{13}C$	$F(1, 172) = 0.00$	$p = 0.95$	$\omega^2 = 0.001$

Notes. NA = Not Applicable.

[a] Interpretation of ω^2: small = 0.01, *medium* = 0.06, and **large** = 0.14 (Field et al., 2012, p. 455).

Turning to the main analysis (Table 7.8), only nail carbon had a significant interaction effect between gender and population. Otherwise, none of the tests found a significant main effect of gender on isotope values. Notably, the main effect of the population (i.e. population affected isotopic results overall) was always significant, as expected, but it was not reported, since its results are included in the previous section.

The simple effects analysis on nail carbon isotopes failed to find any significant difference between men and women within each population (Table 7.9). However, this non-significance is only reached after applying the FDR correction. Prior, the Turkana, and the Luo (Port Vict.), differed between genders in nail $\delta^{13}C$. Furthermore, all of the effect sizes are medium, suggesting that the sample size was too small to detect any differences. Notwithstanding, Figure 7.9 demonstrates that the difference between men and women in nail carbon isotopes is indeed small. It is possible that the imbalance in sample size between the two genders explains the nearly significant results obtained, as well as the medium effect sizes.

Table 7.9
Results of simple effect analysis of gender on nail $\delta^{13}C$, by population. Shaded cells are significant comparisons (p < 0.05).

	Statistic	Probability Value	Effect Size [a]
El Molo	$t(6.64) = -1.53$	$p = 0.21$	*r = 0.44*
Turkana	$t(6.78) = -2.74$	$p = 0.09$	*r =* **0.57**
Luhya (Webuye)	$t(1.41) = 1.48$	$p = 0.32$	*r = 0.47*
Luhya (Port Vict.)	$t(9.77) = 2.21$	$p = 0.09$	*r = 0.38*
Luo (Port Vict.)	$t(17.86) = 2.16$	$p = 0.09$	*r = 0.41*

Notes. Interpretation of *r*: small = 0.10, *medium* = 0.30, and **large** = 0.50 (Field et al., 2012, p. 58).

Overall, the isotopic results of this study did not differ between the two genders. A significant interaction effect was found in nail $\delta^{13}C$, but this difference could not be detected within each population.

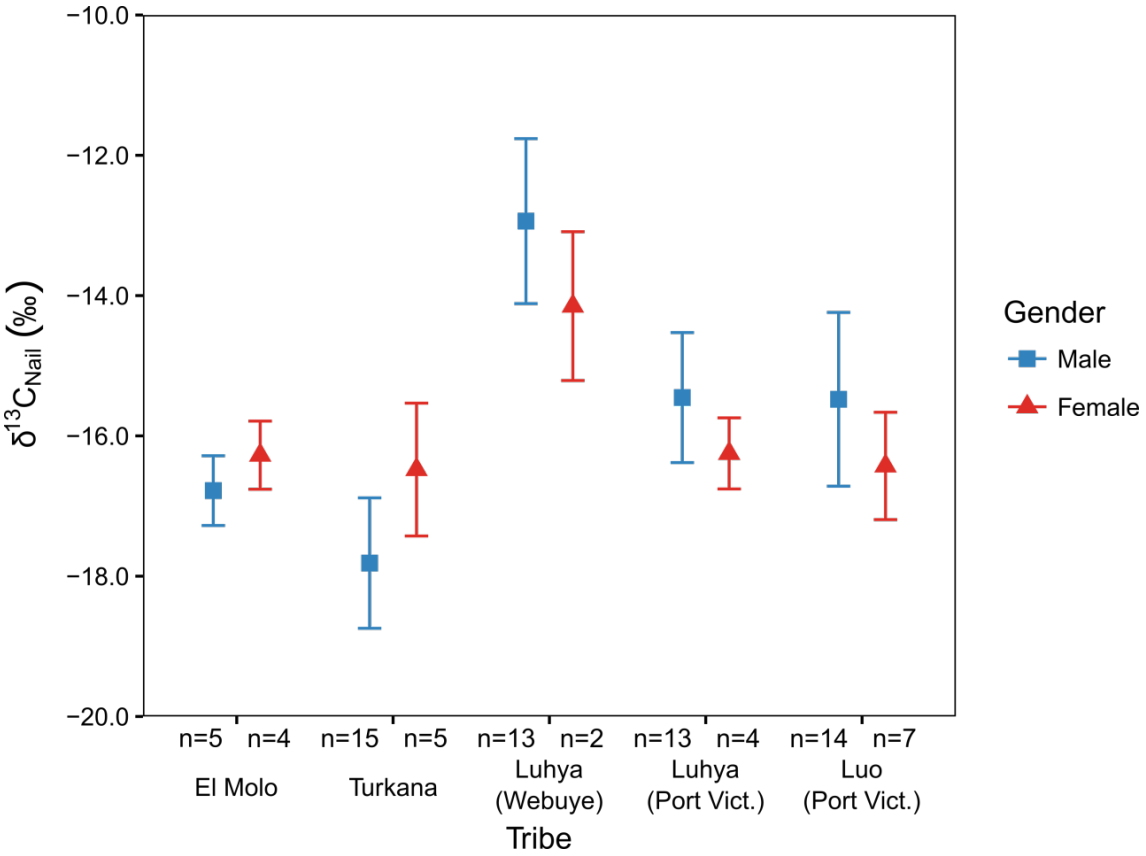


Figure 7.9 Bar plot of nail $\delta^{13}C$, grouped by population and gender. Error bars represent the standard deviation.

7.3.3. AGE

As mentioned before, the robust ANCOVA in R is limited to two categories and thus, non-normality and heteroscedasticity issues cannot be addressed through this method. Nevertheless, these do not seem to be important factors within this analysis (Appendix D, p. 398). Instead, once again, the problem arises from the lack of independence between the covariate and the treatment, since the distribution of individual ages differs among groups, $F(5, 187) = 6.22$, $p < 0.001$, $\omega^2 = 0.12$. Therefore, if any significant results are found, one must keep in mind that some variance attributed to the covariate, i.e. age, is, in fact, attributable to the population affiliation.

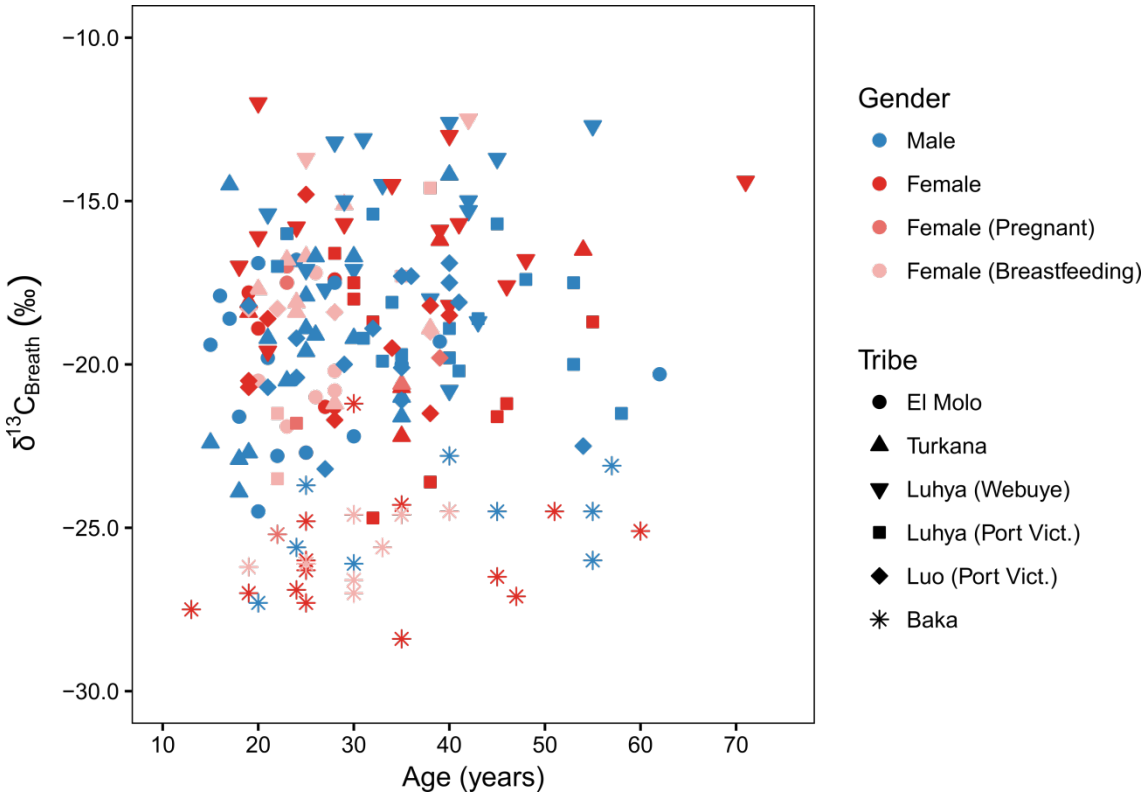


Figure 7.10 Isotopic results of breath samples, by gender and population, and plotted against the age of the individual.

Figure 7.10 shows that breath $\delta^{13}C$ does not differ with the age of the individual. Accordingly, the main analysis found no significant interaction effect, indicating that these tests did not break the assumption of homogeneity of regression slopes (Appendix D, p. 399) (Field et al., 2012, p. 467). Furthermore, the covariate was not significantly related to the isotopic results in any of the body pools or isotopes (Table 7.10). One of these tests, for nail $\delta^{15}N$, came close to significance ($p = 0.06$), however, this might be explained by the influence of the population isotopic variation. In addition, the effect size for this test is small, further confirming that age does not influence the nitrogen isotopic signal of nail.

To sum up, the isotopic results of this study do not differ significantly with either the gender or the age of the participants.

Table 7.10
*Results of the main effects of the ANCOVA for the covariate (age), by body pool and isotope. Shaded cells are significant comparisons (*p < *0.05).*

		Statistic	Probability Value	Effect Size [a]
Hair	$\delta^{13}C$	$F(1, 122) = 2.737$	$p = 0.10$	$\omega^2 = 0.04$
	$\delta^{15}N$	$F(1, 122) = 1.65$	$p = 0.20$	$\omega^2 = 0.04$
Nail	$\delta^{13}C$	$F(1, 72) = 0.01$	$p = 0.93$	$\omega^2 = 0.003$
	$\delta^{15}N$	$F(1, 72) = 3.80$	$p = 0.06$	$\omega^2 = 0.007$
Breath	$\delta^{13}C$	$F(1, 172) = 1.80$	$p = 0.67$	$\omega^2 = 0.004$

Notes. NA = Not Applicable. Values in **bold** are significant ($p < 0.05$).
[a] Interpretation of ω^2: small = 0.01, *medium* = 0.06, and **large** = 0.14 (Field et al., 2012, p. 455).

7.4. COMPARISON BETWEEN BODY POOLS

After discussing the effect of population, gender, and age on the isotopic results, this section discusses how the different body pools relate to each other isotopically. One may study this by applying general linear models between each combination of tissues, and for each isotope.

The justification for the exact type of regression chosen is given with each analysis, since it depended on the best fit to the distribution of data. Nevertheless, two points were common to all models. First, as a measure of fit of the model, the adjusted coefficients of determination (r^2) are reported. This adjusted version of r^2 informs on how much variance in the outcome variable would be accounted for if the model had been derived from the population, rather than a sample. In addition, this statistic also controls for the fact that r^2 will increase as more variables are added to the model, giving a false sense that the model is improving (Field et al., 2012, pp. 263, 273). Second, the treatment of outliers and influential cases was always the same. In detail, a case was only removed from the data set prior to modelling if its Cook's distance were bigger than one. This statistic considers the effect of a single case on the model as a whole, and thus, it identifies data points that have a large effect on the regression analysis (Field et al., 2012, p. 271).

At this point, one must recall some of the information on the isotopic fractionation in the human body, reviewed in Chapter 2 (pp. 42-51). To begin, some of these isotopic

differences between body pools have been previously investigated. In detail, O'Connell et al. (2001) found a non-significant (0.21‰) enrichment in hair ^{13}C, in relation to nail, whereas Hülsemann et al. (2015) chose to use an offset of 0.4‰ to transform the carbon isotopic signal of nail to hair. In addition, hair has been estimated to be about 0.6‰ depleted in ^{15}N in relation to nail (Hülsemann et al., 2015; O'Connell et al., 2001). Note that these studies suggest a simple offset to transform the isotopic signal of one body pool into another. Following the same reasoning, one could hypothesise that the spacing between breath $\delta^{13}C$ and hair or nail $\delta^{13}C$ will be around 4-6‰, by adding the estimated depletion in breath ^{13}C with the enrichment in hair and nail ^{13}C (Hedges et al., 2009; O'Connell & Hedges, 1999a; O'Connell et al., 2001).

However, the carbon in body protein (including hair, nail, and collagen) is biased towards the isotopic signal of dietary protein, whereas other body pools (including breath and bioapatite), record the isotopic signal of the overall diet (Fogel & Tuross, 2003; Hedges et al., 2009). In other words, these relationships will likely be more complex than the one between hair and nail. Backing this idea, Froehle et al. (2010) found two different offsets, depending on the protein source, to model the relationship between collagen and apatite $\delta^{13}C$. Namely, this study found two linear regressions, with similar slopes, but different intercepts, of which the first applies to individuals whose protein source is C_4 or marine based, whereas the second applies to individuals with a protein source based on C_3 foods. In a follow up study, the same authors used cluster analysis and discriminant function analysis to include $\delta^{15}N$ in the model, and thus, distinguish between protein sources based on C_4 and marine foods (Froehle, Kellner, & Schoeninger, 2012).

A final point concerns the sample size available in this study for such analyses. From the total number of participants, 32 provided both hair and nail, 130 provided hair and breath, and 78 provided nail and breath. Therefore, the comparison between hair and nail, in particular, might not be appropriate for generalization to other studies.

7.4.1. HAIR-NAIL COMPARISON

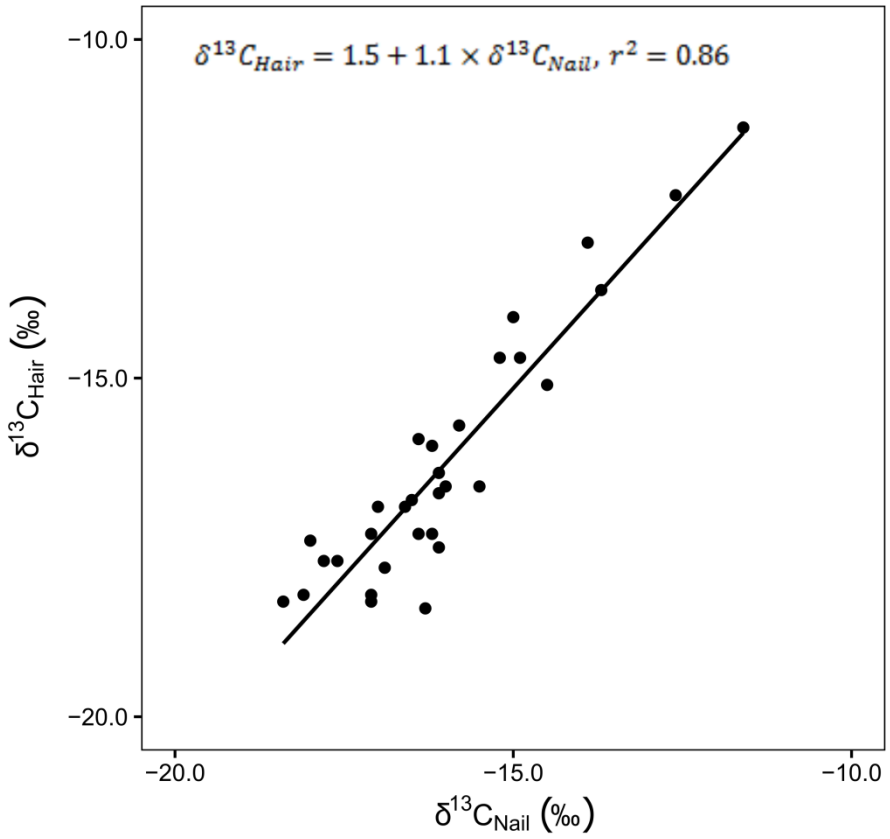


Figure 7.11 Comparison of carbon isotopic values of hair and nail. The plot includes the regression line and equation, as well as the coefficient of determination (r^2).

Regarding the carbon isotopic signal of hair and nail, a simple linear regression was calculated to predict participant's hair δ¹³C based on the nail δ¹³C (Figure 7.11). From the original data set, the participant TK29 was removed based on its Cook's distance. A significant regression equation was found, $F(1, 29) = 181.1$, $p < 0.0001$, with an r^2 of 0.86, where the predicted hair δ¹³C of participants is equal to $1.5 + 1.1 \times \delta^{13}C_{Nail}$. Conversely, to predict nail δ¹³C, one can simply invert this equation, which is preferable to the transposed equation (Pryor, Stevens, O'Connell, & Lister, 2014). Thus, the predicted nail δ¹³C of participants is equal to $(\delta^{13}C_{Hair} - 1.5)/1.1$. Furthermore, when examining the first equation in detail, it is possible to observe that the slope ($b_1 = 1.1$) of the equation is very close to one, with confidence intervals (CI) falling close to this value, 95% CI [0.9, 1.3], which indicates that a fixed offset could explain the differences between the two tissues. This offset could be interpreted as the intercept ($b_0 = 1.5$) of the equation. However, this value was not significantly different from zero,

$t(30) = 1.15$, $p = 0.27$, and indeed, the confidence interval around this estimate includes zero, 95% CI [−1.2, 4.2]. Taken together, this confirms that hair will be slightly enriched in ^{13}C in relation to nail, but that overall, the isotopic signal of carbon in hair and nail are very similar to each other and hence, that this enrichment is negligible, just as suggested by O'Connell et al. (2001). This is also consistent with the very high r^2 of 0.86, which suggests a near perfect linear relationship between the two variables (Field et al., 2012, p. 271). However, one must consider the low sample size available to this analysis, which might explain the uncertainty in the estimates.

Following this, another linear regression predicted participant's hair $\delta^{15}N$ based on the nail $\delta^{15}N$ (Figure 7.12). Once again, a significant regression equation was found, $F(1, 30) = 29.5$, $p < 0.0001$, with an r^2 of 0.50, where the predicted hair $\delta^{15}N$ of participants is equal to $2.5 + 0.7 \times \delta^{15}N_{Nail}$. In this equation, the slope ($b_1 = 0.7$) of the equation is not close to 1, albeit its confidence intervals, 95% CI [0.4, 0.9], do not fall as close to the estimate as they do in the previous equation. The intercept ($b_0 = 2.5$) was not significantly different from zero, $t(31) = 1.96$, $p = 0.06$. Accordingly, the confidence interval around this estimate includes zero, 95% CI [−0.1, 5.1]. Therefore, this equation agrees with previous literature that hair is depleted in ^{15}N in relation to nail (Hülsemann et al., 2015; O'Connell et al., 2001), but it disagrees that this relationship can be modelled with a simple fixed offset (O'Connell et al., 2001). Instead, it suggests that the average hair $\delta^{15}N$ increases 0.7‰ for each 1‰ increase in nail $\delta^{15}N$. However, in view of the wide confidence intervals around this estimate, this equation might not be reliable. Accordingly, a relatively low r^2 (0.50) suggests that the model accounts for only 50% of the variation in hair $\delta^{15}N$. Nevertheless, this relationship could clarify the differences found in $\delta^{15}N$ between hair and nail. In detail, as the nitrogen isotopic values of hair increase, the values of nail increase in a higher proportion, which might explain why, for hair, the El Molo and the Turkana pool with the Luhya (Webuye), with lower nitrogen isotopic values, whereas, for nail, these two groups pool with the Luo and Luhya, from Port Victoria, with higher nitrogen isotopic values.

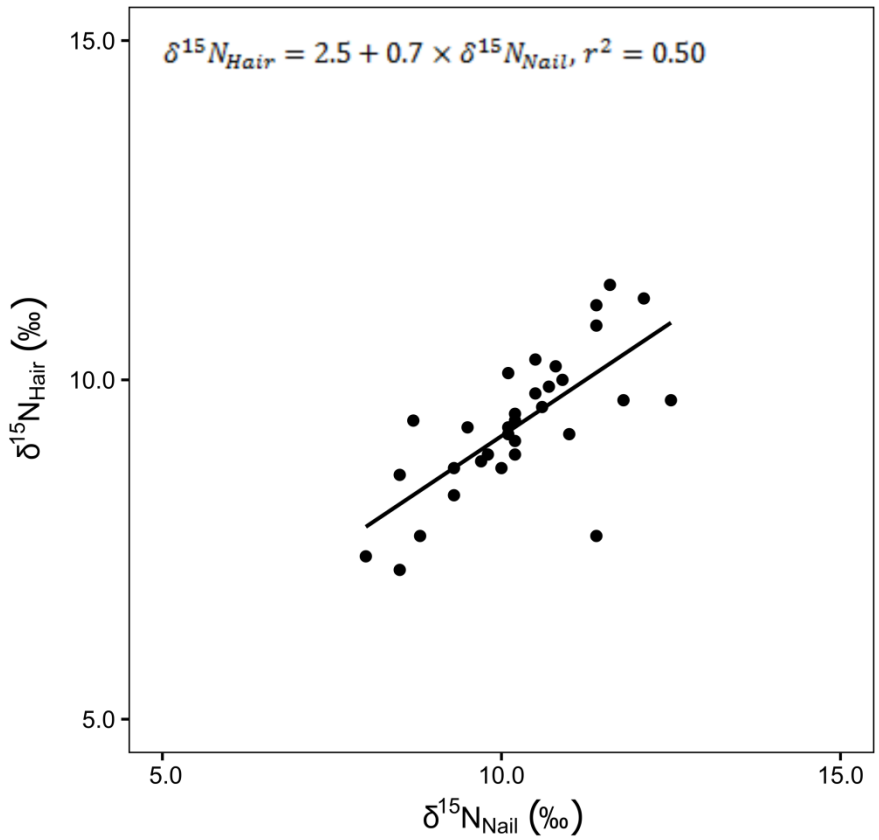


Figure 7.12 Comparison of nitrogen isotopic values of hair and nail. The plot includes the regression line and equation, as well as the coefficient of determination (r^2).

Finally, one must consider the causes of these differences. In the case of carbon, hair is marginally richer in cysteine and serine than nail (28.1% of total amino acids in hair and 21.9% in nail) (O'Connell et al., 2001). These amino acids are enriched in^{13}C, just like glycine, and thus, this difference in composition could explain the slightly higher hair δ^{13}C values (Hare et al., 1991). In contrast, it is difficult to explain the δ^{15}N results through a difference in the isotopic signal of individual amino acids. Instead, O'Connell et al. (2001) proposes that the intracellular turnover of amino acids may differ between the hair follicle and the nail matrix. This explanation is untested but consistent with the stronger effect of metabolism on the composition of nitrogen isotopes (through the de- and transamination of amino acids) (Hülsemann et al., 2009; O'Connell & Hedges, 1999a), and by the lack of constant offset between the δ^{15}N of different tissues and their correspondent pools of free amino acids (Poupin et al., 2014). Alternatively, one must also consider that hair and nail represent slightly different time periods: 1 cm of hair represents one month growth, while 1 cm of nail represents three

months' growth (Chapter 2, p. 42) (Saitoh et al., 1969; Zaias, 1980). Therefore, it is possible that a changing diet associated with seasonality could explain the differences between the two tissues. In particular, discrimination factors between the $\delta^{15}N$ of the diet and that of tissue depend on the efficiency of nitrogen utilization, and usually decrease with increasing trophic level and protein quality (Cantalapiedra-Hijar et al., 2015; Robbins et al., 2005).

7.4.2. *HAIR-BREATH COMPARISON*

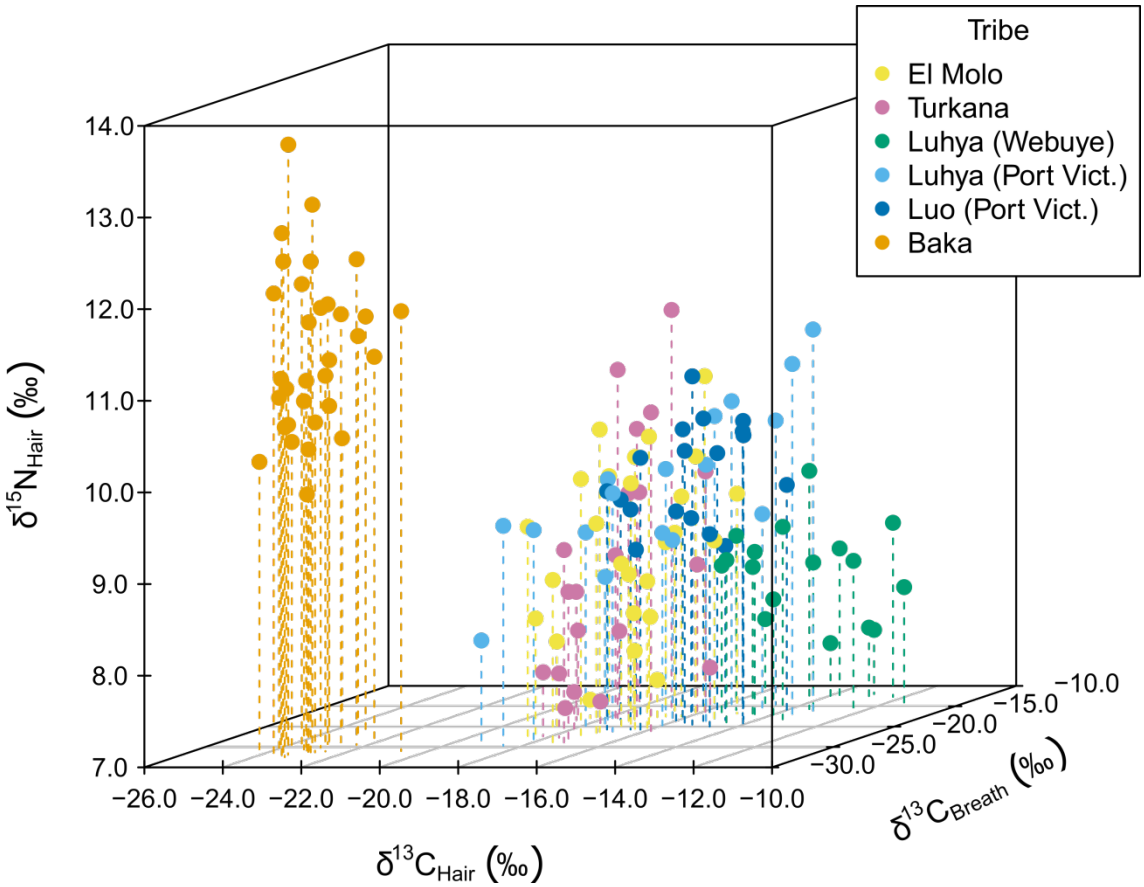


Figure 7.13 Three-dimensional plot of hair carbon isotopes, breath carbon isotopes, and hair nitrogen isotopes, coloured by population.

As mentioned at the start of this section, the relationship between hair and breath carbon is likely to be more complex than the one between hair and nail carbon, since in the former, the two body pools are recording different parts of the diet. In particular, it is likely to depend on the type of protein source (C_3, C_4, or marine) that the populations consume (Froehle et al., 2010, 2012).

Indeed, the residuals of a simple regression break the linearity assumption, despite the seemingly high r^2 of 0.70, therefore suggesting that a simple linear equation cannot explain the variability within the data. Indeed, the Baka probably drive the high r^2 and the lack of linearity, because it is such a distinct group isotopically. Accordingly, if this population is removed, the r^2 plummets to 0.29 (Appendix D, p. 405). However, and as pointed out previously (Chapter 4, p. 93), precise information is not available on the diet of animal foods consumed, and thus, it is impossible to know whether the protein source of these populations is C_3 or C_4. A possible solution is to add $\delta^{15}N$ information to the model, as a measure of trophic level and protein quality (Froehle et al., 2012). Despite this, this multiple linear regression does not improve the model ($r^2 = 0.71$) (Appendix D, p. 406). Figure 7.13 illustrates this point, wherein the Baka remain a distinct group, whereas the other groups do not show any clear pattern, despite the added $\delta^{15}N$ data.

Another option is to use a measure of C_4 intake from the diet assessment as a proxy of C_3 versus C_4 plant protein. The assumption being that C_4 intake varies in the same direction in both animals and humans, i.e. if a group eats more C_4 foods than the other groups, then they will also feed the animals they consume more C_4 foods. In this case, such a model will automatically exclude the data from the Baka, since diet questionnaires were not conducted among this population. Nevertheless, a multiple linear regression between hair $\delta^{13}C$, breath $\delta^{13}C$, and C_4 intake in g/day still has a low r^2 of 0.32. If, instead, one breaks down C_4 intake into tertiles, the correspondent multiple regression still has a low r^2 of 0.32. However, Figure 7.14 shows that the first tertile (i.e. those that consume less C_4 foods) has a lower intercept, but a similar slope to the other two equations, suggesting that the relationship between the hair and breath $\delta^{13}C$ will change depending on the amount of C_4 intake, just as suggested by Froehle (2010). Notwithstanding, this model still explains little of the variance present in the data set.

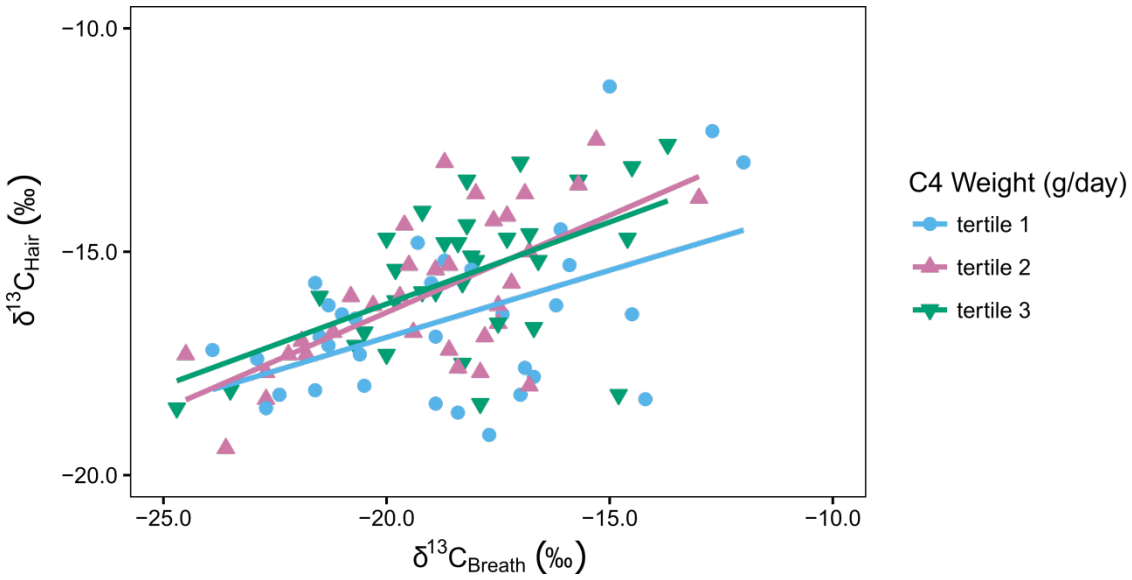


Figure 7.14 Comparison of carbon isotopic values of hair and breath, separated by tertiles of C_4 intake in g/day, with the regression line for each tertile.

Thus, a final attempt to model the difference between the two body pools used a multiple linear regression to predict hair $\delta^{13}C$ based on breath $\delta^{13}C$, and on the ethnic affiliation, which produces a separate equation for each population (Figure 7.15). Within this model, a significant interaction was found between breath $\delta^{13}C$ and ethnic affiliation, $F(5, 118) = 5.25$, $p < 0.001$, which means that the slopes of the equations will vary between the groups. This model produced the best fit to the data by far, where a significant regression was found, $F(12, 118) = 3878$, $p < 0.0001$, with an r^2 of 1.00. In addition, the r^2 is maintained even if the Baka are removed from the dataset. When examining this model in detail, one can observe that most equations have null slopes, as depicted in Figure 7.15 (Appendix D, p. 408-409). In effect, the only slope ($b_1 = 0.4$) that is significantly different from zero is that of the Luhya (Port Victoria), $t(14) = 3.15$, $p < 0.01$. As confirmation, the confidence intervals around this estimate do not include zero, 95% CI [0.1, 0.6], although the interval is wide, suggesting that the estimate is unreliable. Similarly, the intercept ($b_0 = -5.7$) for this population is also significantly different from zero, $t(14) = -3.34$, $p < 0.01$, and the confidence interval around the estimate is also wide, 95% CI [−9.1, −2.3].

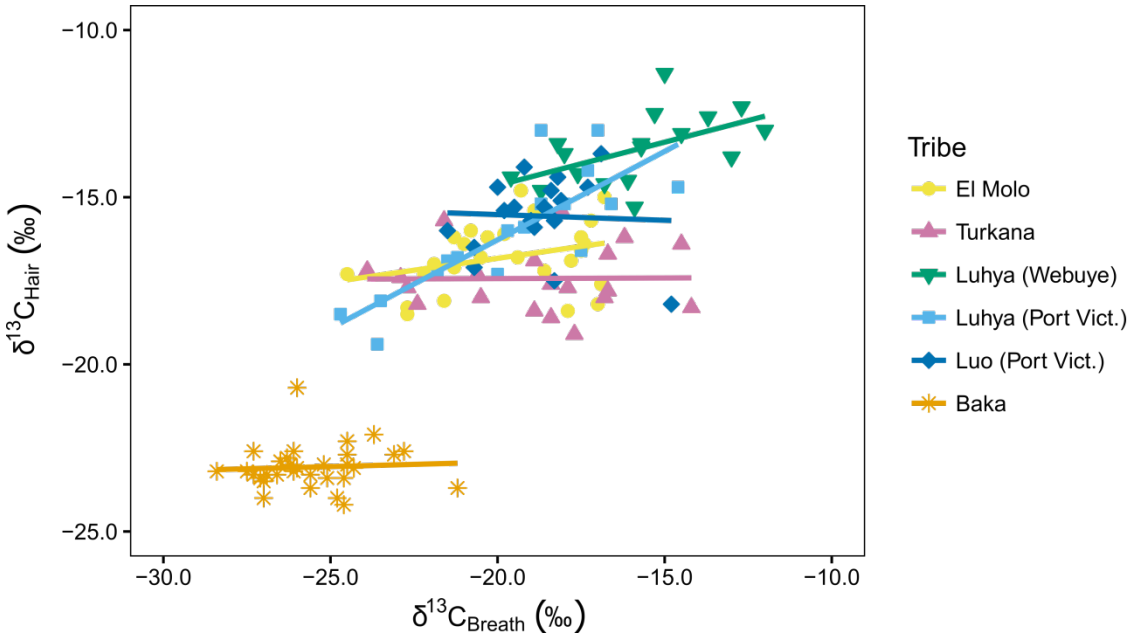


Figure 7.15 Comparison of carbon isotopic values of hair and breath, separated by ethnic affiliation, with the regression line for each group.

Overall, this suggests that the equation found for the Luhya (Port Victoria) is a poor fit to the data, whereas the null slopes for the other groups indicate that when breath $\delta^{13}C$ increases, hair $\delta^{13}C$ values remain the same. Indeed, the hair $\delta^{13}C$ will match the intercepts found for each group, which vary from −22.3‰ for the Baka to −9.5‰ for the Luhya (Webuye). Two main reasons explain these results. First, by adding the variable ethnic affiliation, the number of equations the model had to estimate increased from one to six since the interaction between breath $\delta^{13}C$ and group was significant, which meant that two different estimates (intercept and slope) had to be calculated for each equation. Consequently, the sample size per equation decreased, reducing the statistical power of the test. At the same time, the complex model might lead to an inflated r^2, despite the fact that the adjusted version of this statistic is reported (Field et al., 2012, pp. 263, 273). Second, the slopes for most of the equations may fall close to zero because the variability in breath $\delta^{13}C$ is much higher than that of hair $\delta^{13}C$ (see Table 7.2 and Table 7.4). It is not clear why the Luhya (Port Victoria) differ from the other groups, but this could be a spurious result, given the wide confidence intervals around the estimates. In addition, the high variability in breath $\delta^{13}C$ might also explain why the previous models had such low r^2. In this sense, breath $\delta^{13}C$ varied too much among groups, and so, only when this information was included in the model, could it

explain the variance observed. Therefore, population affiliation might be acting as a measure of diet variation in the model. On the other hand, if one trusts that the null slopes represent a true effect, they would indicate that overall protein sources do not change isotopically, whereas overall diet does. The only exception being the Luhya (Port Victoria), where the protein sources vary isotopically. For this group, hair $\delta^{13}C$ increases 0.4‰ for each 1‰ increase in breath $\delta^{13}C$. In this case, more westernised diets for the Luhya (Port Victoria) could explain this result. This group may acquire some food through commerce, which may originate from locations that imprint a different isotopic signature in the resources exported. However, it is not possible to confirm or discredit this hypothesis since no data was collected on the provenance of each protein source.

7.4.3. NAIL-BREATH COMPARISON

To examine the relationship between the carbon isotopic signal of nail and breath, the same reasoning used for hair was applied. First, a simple linear regression predicted nail $\delta^{13}C$ based on breath $\delta^{13}C$. Since this dataset does not include any samples from the Baka, and as before, the r^2 of 0.27 for this regression is low. Similarly, adding the $\delta^{15}N$ signal of nail to the model also produces a low r^2 of 0.26. Figure 7.16 illustrates this model, where the lack of pattern across isotopic signals is clear. Following this, including C_4 intake in the model, either overall, or broken down in tertiles, also produces a low r^2 of 0.26 for both models. In this case, the lower tertile does not differ considerably from the other two (Figure 7.17). This probably arises from the lower number of carbon isotopic results for nail.

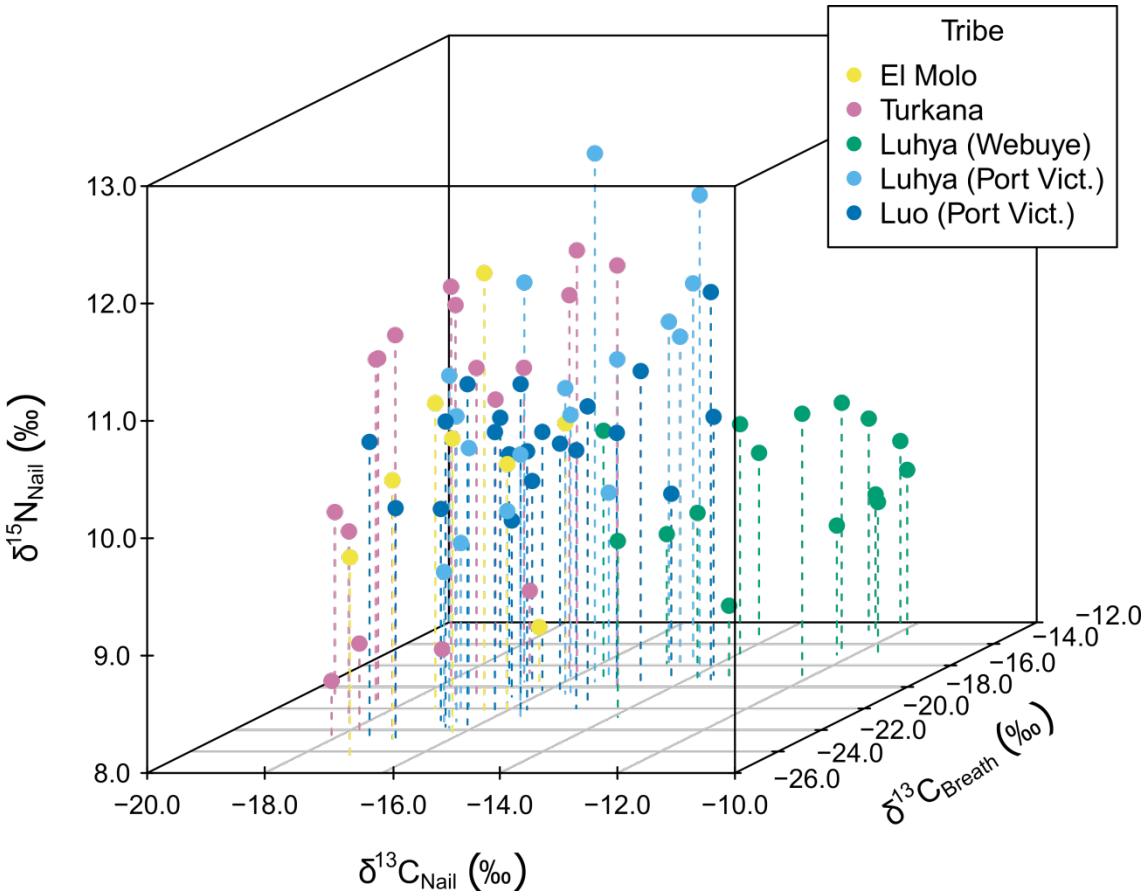


Figure 7.16 Three-dimensional plot of nail carbon isotopes, breath carbon isotopes, and nail nitrogen isotopes, coloured by population.

Finally, just as for hair, the last model produced a separate equation for each population to predict nail $\delta^{13}C$ based on breath $\delta^{13}C$. In this case, the interaction between breath $\delta^{13}C$ and the ethnic group was not significant, $F(4, 68) = 0.92$, $p = 0.46$, indicating that the slope does not differ much between populations, as can be seen in Figure 7.18. Subsequently, and as before, this model produced the best fit to the data, and a significant multiple regression was found, $F\ (6, 72) = 3101$, $p < 0.0001$, with an r^2 of 1.00. In this model, the slope ($b_1 = 0.2$) was significantly different from zero $t\ (72) = 0.16$, $p < 0.01$, but close to it, with a confidence interval that almost includes it 95% CI [0.1, 0.3]. Notably, the slope of the Luhya (Port Victoria) equation still seems to be the one that deviates the most from the other lines, but not considerably so (Figure 7.18). As for hair, this small slope probably results from the inadequate sample size for a more complex model and from the high variability in breath $\delta^{13}C$. Regarding the

intercepts estimated by the model, these vary from −17.1‰ for the Turkana to −9.5‰ for the Luhya (Webuye).

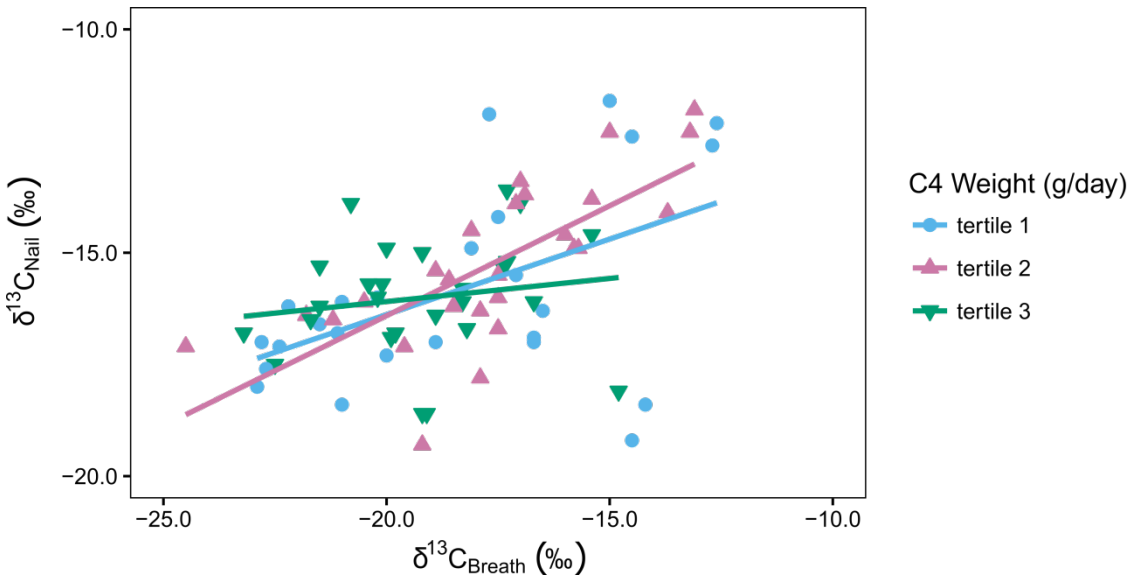


Figure 7.17 Comparison of carbon isotopic values of nail and breath, separated by tertiles of C_4 intake in g/day, with the regression line for each tertile

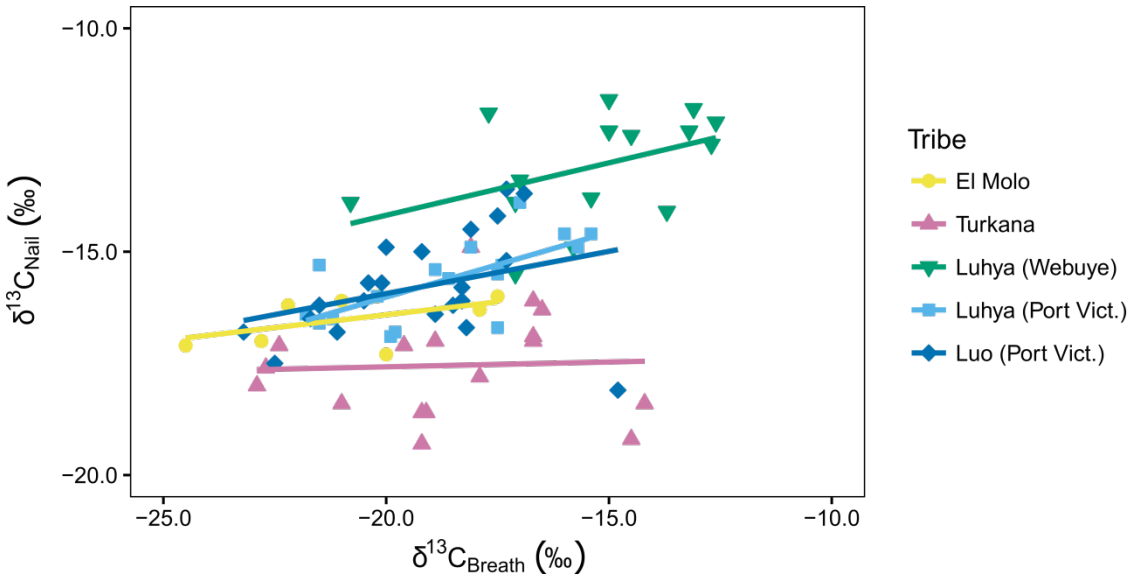


Figure 7.18 Comparison of carbon isotopic values of hair and breath, separated by ethnic affiliation, with the regression line for each group.

Thus, this model presents a similar pattern to that found in the comparison between nail and breath carbon, where nail $\delta^{13}C$ increases only 0.16‰ for each 1‰ increase in breath $\delta^{13}C$, but where the specific offsets vary according to the population. In a way, this finding agrees with Froehle (2010) by suggesting that the offset will differ according to the groups' diet. However, it also suggests that when dealing with mixed

diets and complex systems, a simple model, such as the one proposed by Froehle (2010), might fail to model the relationship between different tissues.

7.5. EFFECT OF DIET ON MODERN HUMAN ISOTOPIC VALUES

At this point, it is pertinent to examine how the human isotopic signatures described so far in this chapter relate to the information gathered through the semi-quantitative food frequency questionnaires. Naturally, these analyses focus on the populations for which dietary information was obtained in the context of this thesis, and thus, exclude the Baka. Nevertheless, the diet of this group is still addressed, whenever possible. In order to relate the two sets of data, this section uses three approaches. The first is simply descriptive, and considers the isotopic results vis-à-vis the diet assessment results. The second looks at the correlations between stable isotopes and dietary intake measurements. And, finally, the third uses exploratory data analysis, or, to be more precise, principal component analysis (PCA), to detect any underlying constructs that might relate different variables (Field et al., 2012, p. 750). Notably, the first approach still looks at the results from a population perspective, whereas the other two take into consideration the variability found in the diet assessment, since they can examine how the data relate at the individual level, rather than at the population level.

7.5.1. A VISUAL COMPARISON

Within the first approach, Figure 7.19 and Figure 7.20 summarise the information obtained from diet assessment and from the carbon and nitrogen isotopic analysis of all body pools.

Figure 7.19 Schematic of the diet assessment results compared to the carbon isotopic results. Population position along the axis does not represent absolute values, but instead the relative relation between the groups. Populations with overlapping positions indicate similar diets (in food questionnaires) or non-significant comparisons. The Baka are not present in the diet assessment or nail section because that data were not collected for that population.

In Figure 7.19, it is clear that the isotopic results corroborate the diet assessment to some extent, but not completely. First, the Baka have the most C_3 based diet. Despite the lack of diet assessment data for this population, this agrees with dietary studies conducted among the Baka (Chapter 5, p. 145) (Sato et al., 2012; Yamauchi et al., 2000). However, this ^{13}C depletion might partly derive from a "canopy effect", wherein plant $\delta^{13}C$ values decrease the closer they are to the ground. This argument is compelling because yams (Baka's main food staple) are a root tuber (Bonafini et al., 2013; Yamauchi et al., 2000).

Second, the Turkana and the El Molo have a mixed C_3 and C_4 plant diet, but have the lowest C_4 component from among the remaining populations, just as predicted by the diet assessment. This pattern is not as clear in breath, which could be the result of (1) the high variability within this body pool, or (2) a difference in the isotopic signal of the protein part of the diet versus the overall diet (Froehle et al., 2010). In particular, this population had lower breath $\delta^{13}C$ values (mean = −19.0‰, SD = 2.5) than hair (mean = −17.5‰, SD = 1.1) and nail $\delta^{13}C$ (mean = −17.5‰, SD = 1.0), suggesting that animal foods eat more C_4 foods than the individuals themselves do. Unfortunately, animal foods were not analysed isotopically, making it difficult to confirm this.

Third, the Luhya (Webuye) do not follow the pattern suggested by the diet assessment, which suggests that the diets of the Luo and the Luhya (Port Victoria) should have a higher C_4 component than the Luhya (Webuye), and consequently higher carbon isotopic values. As for the Turkana, indirect consumption of C_4 foods among the Luhya (Webuye) could partly explain this pattern. Accordingly, the carbon isotopic values of breath (mean = −15.6‰, SD = 2.2) for this group are lower than those of hair (mean = −13.6 ‰, SD = 1.0) and nail (mean = −13.1 ‰, SD = 1.0). Furthermore, the isotopic signal of hair and nail among the Luhya (Webuye) is very close to the values obtained for C_4 foods (Figure 7.1, Table 7.1), suggesting that the protein part of this group's diet is entirely based on C_4 plants. In addition, the regression line for the Luhya (Webuye) in Figure 7.15 and Figure 7.18 is always drawn higher in the plot, and with a corresponding higher intercept, than the regression lines for the other groups. This matches Froehle's (2010) hypothesis that the regression model of a C_4 protein diet has a higher intercept than a C_3 protein diet. Nevertheless, the carbon isotopic signal of the Luhya (Webuye) is always higher than the two groups from Port Victoria, independently of the body pool, suggesting that indeed the diet assessment underestimated the amount of C_4 foods consumed, possibly through the incorrect estimation of food portions. Overall, these findings stress the necessity to analyse different body pools when assessing diet isotopically (Froehle et al., 2010). Finally, the Luhya and Luo, from Port Victoria, always overlap in carbon isotopic values,

confirming that these two groups consume similar amounts of C_3 and C_4 plant foods, just as suggested by the diet questionnaires.

Following this, Figure 7.20 shows that the nitrogen isotopic results validate the diet assessment only marginally. Note that this illustration assumes that intermediate $\delta^{15}N$ values equate to a greater consumption of animal protein, whereas higher $\delta^{15}N$ values equate to fish consumption (Chapter 2, p. 39) (O'Brien, 2015), when in fact, many of the results here suggest that this assumption might not always hold true.

Figure 7.20 Schematic of the diet assessment results compared to the nitrogen isotopic results. Population position along the axis does not represent absolute values, but instead the relative relation between the groups. Populations with overlapping positions indicate similar diets (in food questionnaires) or non-significant comparisons. The Baka are not present in the diet assessment or nail section because that data were not collected for that population.

To start, no clear trophic shift is present between the fishers and the other populations. In this case, the high nitrogen isotopic content of plants and the low nitrogen isotopic content of fish obtained through the analysis of food sources (see Figure 7.1 and Table 7.1) could explain this result. This is a key finding, since it suggests that the exploitation of aquatic resources on these shores is very difficult to detect isotopically.

In addition, the Baka have surprisingly high nitrogen isotopic values, considering that previous studies indicate that they consume low amounts of animal foods (Sato et al., 2012; Yamauchi et al., 2000). Since no food samples were collected from Cameroon, it is impossible to know if plants there have even higher $\delta^{15}N$ values that could explain this result. Notwithstanding, a few primate studies in West African forests also report relatively high $\delta^{15}N$ values (Macho & Lee-Thorp, 2014; Oelze et al., 2011). In detail, Oelze et al. (2011) analysed hair samples from a bonobo population in a tropical forest in the Democratic Republic of Congo ($\delta^{15}N$: mean = +8.4‰, SD = 0.3, $n \geq 21$), whereas Macho & Lee-Thorp (2014) analysed hair samples from a gorilla ($\delta^{15}N$: mean = +8.1‰, SD = 1.6, $n \geq 55$) and chimpanzee population ($\delta^{15}N$: mean = +9.4‰, SD = 0.9, $n \geq 39$) from Central Cameroon. Although Macho et al. (2014) state that the higher $\delta^{15}N$ values of chimpanzees could reflect the consumption of more animal foods for this species, the overall $\delta^{15}N$ values for these two studies are still relatively high, and, indeed Oelze et al. (2011) found little evidence for ^{15}N enrichment resultant from the consumption of animal foods. Furthermore, Oelze et al. (2011, p. 9794) also reports relatively high $\delta^{15}N$ values for other faunal samples, and they attribute this general enrichment to "the complexity of the nitrogen assimilation pathways in the tropical forest".

In parallel, the Luhya (Webuye) have surprisingly low nitrogen isotopic values considering that this population has the highest proportion of animal foods in dietary intake overall, and the second highest in terms of protein intake (Table 5.11, p. 150). Similarly, the Turkana have the lowest nitrogen isotopic values, despite consuming the highest amount of animal protein, although admittedly, they do not have one of the highest consumptions of animal foods in overall dietary intake. Different phenomena could be driving these results. One is simply an inaccuracy in the estimations obtained by the diet assessment, just as suggested previously for the Luhya (Webuye). Otherwise, low nitrogen isotopic values at the baseline might explain the Luhya (Webuye) results, as indicated by the isotopic results of the food samples (see Figure 7.1 and Table 7.1). For the Turkana, the lower mean in nitrogen isotopic results might derive from a higher variability within the population through the inclusion of many

individuals with a carbohydrate-rich diet, and consequent lower consumption of animal foods. Accordingly, the Turkana with the highest nitrogen isotopic values overlap with those with higher protein consumption as recorded in the diet questionnaires (the outliers in Table 5.1.b, p. 113). In detail, the four highest hair nitrogen isotopic values belong to TK35 (+11.4‰), TK23 (+11.1‰), TK03 (+10.8‰), and TK22 (+10.5‰), and the four highest nitrogen isotopic values in nail values belong to TK35 (+11.6‰), TK14 (+11.6‰), TK23 (+11.6‰), and TK03 (+11.4‰), whereas the top four outliers in protein intake are TK23, TK03, TK22, and TK35.

Alternative explanations involve the influence on isotopic signals of factors other than the diet. A possibility is the enrichment of ^{15}N due to nutritional stress (Chapter 2, p. 48) (Fuller et al., 2004a; Reitsema, 2013). This could be true in the case of the El Molo and the Turkana in view of the levels of acute and chronic malnutrition reported by Jeffery (2018). However, despite equivalent data not being available for the other groups, and based on estimated energy intake, one expects the Luhya and Luo (Port Vict.) to have lower levels of malnutrition, and the Luhya (Webuye) to have malnutrition levels similar to the Turkana (Chapter 5, p. 145). In addition, recent work in rhesus macaques revealed that nitrogen isotopes are resistant to several forms of short and long term nutritional stress (Tuross, 2017). Another possibility is the enrichment of ^{15}N due to low precipitation and high temperatures (Amundson et al., 2003; Hartman, 2011; Hülsemann et al., 2015). Once again, this could explain the higher nitrogen isotopic results of the El Molo and of the Turkana in relation to the Luhya (Webuye), but is less applicable to the Luo and the Luhya, since Port Victoria is wetter and colder than the Lake Turkana area, although it is also somewhat drier and hotter than Webuye (Chapter 3, p. 62, p. 73). Furthermore, this effect also cannot explain the high nitrogen isotopic values from the tropical Baka.

Figure 7.20 also reiterates a point explored in the previous section. In particular, hair $\delta^{15}N$ pools the El Molo and the Turkana with the Luhya (Webuye), whereas nail $\delta^{15}N$ pools these two groups with the Luhya and Luo (Port Vict.). In a context of assessment of past diets, this could have critical consequences to the interpretation of results.

Therefore, it is necessary to continue research on the isotopic metabolism if one hopes to interpret correctly the isotopic signals of present and past diets.

Finally, Figure 7.19 and Figure 7.20 highlight the less informative nature of nitrogen isotopes relative to carbon isotopes. In other words, carbon isotopes distinguish between groups better than nitrogen isotopes do. The reasons behind this are twofold. First, the relative abundance of ^{13}C (1.11%) is higher than the abundance of ^{15}N (0.37%) (see Table 2.2, p. 34), and second the relative differences in atomic weight between the two isotopes is higher for carbon than for nitrogen. Taken together, these two factors lead to higher carbon isotopic fractionations. In turn, this means that higher sample sizes are required to detect putative differences in nitrogen isotopic values.

7.5.2. CORRELATION BETWEEN STABLE ISOTOPES AND DIETARY INTAKE

Following the visual comparison of the results of the diet assessment and stable isotope analyses, Kendall's tau (τ) tests were carried out to examine the relationship between the isotopic signals and dietary intake, in a similar approach to that used by Patel et al. (2014) in an epidemiological study.

First, one must decide which diet variables will be analysed: one may look at general measures of overall diet, namely weight and energy, and of macronutrient consumption, namely protein, fat, and carbohydrates. Other important variables include these same measurements of dietary intake, but separated by food categories, i.e. animal, fish, C_3 plant, and C_4 plant. In the end, this work studied similar variables to those analysed by Patel et al. (2014), namely total energy, protein, fat, and carbohydrate, total animal intake, fish intake, C_3 plant intake, and C_4 plant intake, animal protein, fish protein, C_3 plant protein, C_4 plant protein, C_3 plant carbohydrate, and C_4 plant carbohydrate. Table 7.11 reports on the results of these correlations for all tissues and isotopes, and for all dietary variables.

Table 7.11
Correlation analysis between dietary characteristics and isotopic signals.

	Hair (n = 131)				Nail (n = 82)				Breath (n = 151)	
	$\delta^{13}C$		$\delta^{15}N$		$\delta^{13}C$		$\delta^{15}N$		$\delta^{13}C$	
Dietary Characteristics	Kendall's τ	*p- value*	Kendall's τ	*p- value*	Kendall's τ	*p- value*	Kendall's τ	*p- value*	Kendall's τ	*p- value*
Total energy (kcal/day)	− 0.09	0.50	− 0.22	**0.02**	− 0.03	0.83	0.002	0.98	0.07	0.50
Total protein (g/day)	− 0.14	0.13	− 0.24	**0.009**	− 0.11	0.44	− 0.03	0.83	0.01	0.92
Total fat (g/day)	− 0.15	0.10	− 0.19	**0.04**	− 0.14	0.25	− 0.002	0.98	− 0.04	0.74
Total carbohydrate (g/day)	− 0.06	0.67	− 0.20	**0.03**	− 0.002	0.98	− 0.006	0.98	0.06	0.52
Animal intake (g/day)	− 0.002	0.98	0.08	0.52	− 0.08	0.50	0.05	0.74	− 0.03	0.80
Fish intake (g/day)	− 0.04	0.75	0.27	**0.009**	− 0.10	0.51	0.10	0.50	− 0.18	**0.03**
C_3 plant intake (g/day)	0.06	0.61	0.16	0.10	0.01	0.96	0.18	0.10	− 0.02	0.84
C_4 plant intake (g/day)	0.18	0.05	0.16	0.10	0.08	0.59	0.06	0.67	− 0.03	0.75
Animal protein (g/day)	0.005	0.98	0.08	0.50	− 0.10	0.50	0.14	0.26	− 0.02	0.88
Fish protein (g/day)	− 0.04	0.75	0.27	**0.009**	− 0.09	0.51	0.10	0.50	− 0.18	**0.03**
C_3 plant protein (g/day)	− 0.002	0.98	0.04	0.74	0.09	0.50	0.14	0.26	− 0.009	0.97
C_4 plant protein (g/day)	− 0.11	0.35	0.11	0.35	0.04	0.78	0.04	0.79	− 0.05	0.63
C_3 plant carbohydrate (g/day)	− 0.06	0.61	0.05	0.70	− 0.12	0.35	0.12	0.36	− 0.04	0.67
C_4 plant carbohydrate (g/day)	0.15	0.13	0.16	0.10	− 0.08	0.57	0.05	0.74	− 0.05	0.63

Notes. Values in **bold** are significant ($p < 0.05$).

From this table, no significant correlations were found for nail, or for hair $\delta^{13}C$, but hair $\delta^{15}N$ correlated significantly with several dietary variables. The difference between the results of hair and nail $\delta^{15}N$ might be explained by the lower sample size available to the nail tests, but also by the different relationship between the $\delta^{15}N$ of both tissues (see the section Hair-Nail Comparison). Following this, all measures of overall dietary intake have a small ($|\tau| \leq 0.3$) negative correlation with hair $\delta^{15}N$. This result is surprising as it implies that hair $\delta^{15}N$ decreases as the overall intake of food increases. There is no obvious explanation for this result, but it is possible that the individuals eating higher absolute quantities of food are also eating foods with lower ^{15}N content. However, the Luhya and the Luo (Port Victoria), which have the highest dietary intake (see Table 5.2, p. 116), are also the ones whose analysed foods had the highest nitrogen isotopic values (Figure 7.1 and Table 7.1). Alternatively, this is a pattern that might be present at an individual level, but that is obscured by the variability at the population level. In addition, it is also surprising that hair $\delta^{15}N$ correlates with *all* measures of total dietary intake. This could be simply due to the fact that overall intake and macronutrient intake increase together, proportionally. Consistent with this, the diets of all populations had similar proportions of macronutrient intake (Figure 5.2, p. 132), and all measures of total dietary intake correlated highly among themselves (Appendix D, p. 413).

Concurrently, hair $\delta^{15}N$ has a small positive correlation with fish intake and fish protein. This is an easy result to explain, since one expects nitrogen isotopic values to increase with higher amounts of animal and fish intake. This is an encouraging finding, given the lack of trophic shift between fisher and non-fisher populations. On the other hand, the correlations between the variables that measured animal intake and hair $\delta^{15}N$ are not significant, a result that cannot be easily explained, but that could result from an error in diet questionnaires, such as respondent biases.

Finally, breath $\delta^{13}C$ has a small negative correlation with fish intake and fish protein. This is a particularly puzzling result because it is present in breath $\delta^{13}C$, which records the total diet, but not in hair $\delta^{13}C$, which records the protein part of the diet, and thus

should be more biased towards fish consumption. A somewhat convoluted explanation could be that as fish consumption increases the consumption of C_4 foods (particularly by direct consumption) decreases, thus indirectly decreasing the carbon isotopic signal. This is possible considering that most relief food is provided in the form of maize and sorghum (Campbell et al., 1999; Ellis et al., 1987), and individuals with less access to fish may be at higher risk of famine, and thus more prone to resort to relief programmes. However, no correlation is found between C_4 plant intake and any of the isotopic variables.

7.5.3. PRINCIPAL COMPONENT ANALYSIS

As stated at the beginning of this chapter, principal component analysis (PCA) aims to detect any underlying constructs that might be latent in the variables measured. In other words, a PCA is a multivariate technique for identifying the linear components of a set of variables, by extracting the multifactorial effects of the variables measured, and thus explaining multiple sets of multivariate correlation existent in a dataset (Field et al., 2012, p. 750). In the present work, the main goal is to find which diet variables group with which isotopic variables.

Unfortunately, the data set available had a few problems that led to the exclusion of several individuals, and variables, the reasoning for which is now explained. First, one must consider the sample size. To begin, the sample size automatically decreases to 160 because the Baka do not have data available from diet questionnaires. From the remaining, only 31 individuals had isotopic results for all the three body pools, whereas 78 had isotopic results for nail and breath, and 98 had isotopic results for hair and breath. Thus, I opted to exclude the nail isotopic results, and to include the results of hair and breath for the 98 individuals.

At this point, one must decide which *diet* variables will be incorporated in the analysis. As in the previous part, these variables encompass measures of overall and macronutrient intake, but also measures of overall and macronutrient intake separated by food categories. The choice of variables depended on three factors. To start,

multicollinearity should not be present in the dataset. This is problematic because including nearly redundant variables, i.e. variables that correlate too highly ($|r| \geq 0.8$), can cause the PCA to overemphasize the contribution of these variables. Multicollinearity is detected through a determinant below 0.000001. Conversely, very low correlations ($|r| \leq 0.3$) are also undesirable, since they imply that no clusters of variables exist in the data set to be found. To test this, Bartlett's test examines whether the population correlation matrix resembles an identity matrix (a matrix with ones on the main diagonal and zeros elsewhere). Finally, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy indicates whether the pattern of correlations in the data is compact (desirable) or diffuse: a value below is 0.5 is unacceptable, and above 0.7 is good (Field et al., 2012, pp. 769–771).

Accordingly, the correlation matrix of all the measures of food category intake had a determinant below 0.000001, indicating the presence of multicollinearity in the dataset. Therefore, variables were eliminated under the following rationale: (1) as mentioned above, all measures of total dietary intake correlated highly among themselves, therefore only measures of macronutrient intake were considered; (2) animal energy correlated highly with animal protein and fat, whereas animal weight correlated highly with animal carbohydrate, hence animal weight was kept as a measure of overall intake, and animal protein has a measure of macronutrient intake; (3) fat and energy from C_4 plants correlated highly with all other C_4 plant variables except weight, and thus were eliminated; (4) C_3 plant weight correlated highly with C_3 plant carbohydrate, of which the latter was kept, considering that C_3 plant energy is already a measure of overall intake; (5) all fish variables correlated very highly among themselves, and hence, the only variable kept was fish energy, as a measure of overall intake. One may speculate on the reasons behind such high correlations. For measures of total dietary intake, macronutrient and energy may vary proportionally across individuals, just as they do across populations (Figure 5.2, p. 132, Table 5.5, p. 133). For animal and C_3 plant intake, energy correlates highly with the macronutrient(s) on which that type of food is richer. For fish intake, it is possible that the inclusion of three fisher populations

has led to many similar values across individuals, and thus, to high correlations. Finally, the reasons for the pattern in C_4 plant intake are unclear. Nevertheless, one point must be kept in mind when interpreting the PCA results: if a variable that correlated highly with an eliminated variable pools with a third variable in the PCA, then that third variable also groups with the eliminated variable. For instance, if animal protein groups with a variable, then that variable would also group with animal fat and animal energy.

Subsequently, the overall Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy of the correlation matrix was 0.62, which is acceptable, although not ideal, whereas animal protein had a KMO below 0.5. Thus, this variable was also eliminated. This raised the overall KMO to 0.65 and left no variable with KMO below 0.5, although, at 0.51, breath $\delta^{13}C$ was close to this limit. In addition, after eliminating the variables listed above, the determinant of the final correlation matrix was 0.006, which is well above 0.000001. Furthermore, the Bartlett's test was significant, $X^2(66) = 465.8$, $p < 0.0001$, demonstrating that the correlation matrix differs from an identity matrix. Notwithstanding, most correlations were still below an absolute value of 0.3 (Appendix D, p. 413), which warrants caution, since they suggest that the analysis might not be meaningful. To sum up, the twelve variables incorporated into the PCA are: animal weight, C_3 plant energy, C_3 plant protein, C_3 plant fat, C_3 plant carbohydrate, C_4 plant weight, C_4 plant protein, C_4 plant carbohydrate, fish energy, hair $\delta^{13}C$ and $\delta^{15}N$, and breath $\delta^{13}C$.

Having chosen which variables to incorporate in the PCA, an initial analysis was run to obtain eigenvalues for each component in the data. Six components had eigenvalues over Joliffe's criterion of 0.7 and in combination explained 83% of the variance, whereas five factors had eigenvalues over Kaiser's criterion of 1, and in combination explained 77% of the variance. However, if only five factors were extracted, one of the communalities dropped below 0.7, and the fit based upon off diagonal values dropped below 0.95. Conversely, keeping six factors did not present any of these problems, and thus, this was the number of factors extracted (Appendix D, p. 415-416). In addition,

when keeping six factors, the proportion of absolute residuals over 0.05 was below 50%, and the root-mean-square residual was below 0.08, both of which indicate that the there is no need to include more factors in the analysis (Field et al., 2012, pp. 762–787). Finally, rotation may be used to maximize the loading of each variable on one of the extracted factors while minimizing which variables relate to each factor. Yet, when either orthogonal or non-orthogonal rotations are applied to these data, it simply brings together under one factor the variables that belong to the same food category, to the exclusion of all others (Appendix D, p. 416-417). Since the goal of this analysis is to relate diet variables with isotopic variables, it was opted to leave the factors un-rotated (Field et al., 2012, pp. 764–767).

Table 7.12
Summary of exploratory factor analysis results for diet and isotopic variables (n = 98).

	PC1	PC2	PC3	PC4	PC5	PC6
Animal intake (g/day)	0.37	0.03	**−0.51**	0.03	**0.62**	−0.40
Fish energy (kcal/day)	0.45	−0.17	−0.21	**0.65**	−0.22	0.25
C_3 plant energy (kcal/day)	0.49	**0.62**	0.37	−0.14	−0.11	−0.13
C_3 plant protein (g/day)	**0.76**	0.43	0.09	0.00	0.05	−0.03
C_3 plant fat (g/day)	**0.56**	0.41	−0.07	0.38	0.22	0.35
C_3 plant carbohydrate (g/day)	**0.62**	**0.56**	0.32	−0.11	−0.12	−0.11
C_4 plant intake (g/day)	**0.61**	−0.48	0.14	−0.36	−0.12	0.08
C_4 plant protein (g/day)	**0.56**	**−0.59**	0.06	−0.01	0.25	0.18
C_4 plant carbohydrate (g/day)	**0.66**	**−0.60**	0.21	−0.30	0.05	0.10
Hair $\delta^{13}C$	−0.36	−0.11	**0.75**	0.18	0.27	0.03
Hair $\delta^{15}N$	−0.43	0.41	0.12	−0.33	0.35	**0.52**
Breath $\delta^{13}C$	−0.07	−0.24	**0.74**	0.38	0.14	−0.23
Eigenvalues	3.27	2.26	1.75	1.11	0.81	0.75
% of variance	0.27	0.19	0.15	0.09	0.07	0.06

Notes. PC = Principal component; Factors loadings over |0.50| in **bold**.

Table 7.12 shows the factor loadings for each variable included in the PCA. For a sample size around 100, factor loadings above an absolute value of 0.5 are deemed as significant contributors to a given principal component (Field et al., 2012, p. 767). The items that cluster on the same components suggest that the PCA mostly fails to connect diet variables with the isotopic variables. In detail, component 1 represents a joint increase in C_3 and C_4 foods, which indicates that as dietary intake increases, it does so through the inclusion of higher quantities of plant foods, both C_3 and C_4. This finding

might result from the high plant component in all diets (Table 5.6, p. 135), and it could also explain the reduction in hair $\delta^{15}N$ with increased dietary intake (Table 7.11, p. 227), despite the fact that this isotopic signal is not involved in component 1. Conversely, component 2 implies the opposite, i.e. as C_4 food consumption increases, the C_3 component decreases. This effect might act at a lower level, wherein individuals with higher C_3 consumption consume less C_4 foods, and vice versa. This finding could be driven by the Baka that consume a lot of C_3 resources, but no C_4 resources, vis-à-vis the other groups, which usually have the opposite pattern. Next, component 3 is the first that involves the isotopic signals, but it is difficult to interpret. This component indicates that as hair and breath $\delta^{13}C$ increase, animal intake decreases. Just as for the significant correlation between breath $\delta^{13}C$ and fish intake in the previous section, this could mean that people replace animal foods with C_4 plant foods, most likely famine relief foods. Finally, the remaining components cannot be interpreted since only one variable has a high factor loading in each of them: component 4 represents fish intake; component 5 represents animal intake; and component 6 represents hair $\delta^{15}N$.

7.5.4. SUMMARY OF THE EFFECT OF DIET ON MODERN HUMAN ISOTOPIC VALUES

To sum up, this section provided some insight on the effect of diet on modern human isotopic values. The first part: (1) described how diet assessment and stable isotope ratios mostly match for carbon isotopes, but differ for nitrogen isotopes; (2) identified differences in the carbon isotopic signal of the protein part versus the whole diet; (3) determined a ^{15}N depleted baseline as a possible cause for the low $\delta^{15}N$ values of the Luhya (Webuye); (4) pinpointed a lack of trophic shift in the fisher populations studied; (5) explained the high variability in Turkana nitrogen isotopic values with a true variation in the consumption of animal foods; and (6) discussed the possible influence on isotopic signals of factors other than diet, such as metabolic and environmental factors. The second part, found a correlation between increasing fish intake and increasing hair nitrogen isotopic values, despite the lack of a trophic shift. On the other hand, it also detected some puzzling patterns, such as the decrease of hair nitrogen isotopic values with increasing total dietary intake, and the increase in carbon

isotopic values of breath (but not hair or nail carbon) when fish intake decreases, the latter of which could be explained through a replacement of fish foods with C_4 foods. And finally, the most intriguing finding from the third part is a similar pattern to that found in the second part, where individuals replace animal foods with C_4 plant foods, considering that hair and breath $\delta^{13}C$ increase, as the consumption of this staple decreases.

Many aspects might have affected the analysis conducted in this chapter. The most obvious one is an error associated with the diet assessment. For example, in the visual comparison, a possible error in the estimation of (indirect) C_4 consumption might explain why the Luhya (Webuye) have a higher carbon isotopic signal than expected through the semi-quantitative food frequency questionnaires. Other sources of error in this method were discussed in Chapter 5 (p. 114), and one cannot exclude that they might obscure dietary signals that would otherwise explain the isotopic signals. This confirms that semi-quantitative food frequency questionnaires are appropriate to characterise the diet of a group, but fail short when targeting the variability within that group. For such cases, methods with higher precision (e.g. weighed food records) are necessary. Another issue affecting particularly the correlations and PCA is that these analyses relate the absolute values of dietary signal to what is an averaged signal of diet (the isotopes). This is the same reason that led to the analysis of proportions, rather than raw values in Chapter 5. For instance, an individual with a seemingly high consumption of C_4 plants might dilute this signal in considerable amounts of C_3 foods, and the absolute values of food intake cannot reflect this phenomenon. Previous studies (e.g. Patel et al., 2014), also failed to consider this. Most likely, in many similar studies, the dietary signal is strong enough to overcome this issue. However, in complex systems, such as the ones analysed here, this might lead to crucial misinterpretations. Ideally, future work would take this aspect of the analysis into consideration. Nevertheless, in this case, the true ratio of C_3 versus C_4 plant consumption will be hard to determine, since no isotopic analysis of the animal foods were conducted. A final element that might cloud these analyses is simply the

influence of external factors (other than diet) on the isotopic signals. The first part of this section discusses some of these factors, such as precipitation, temperature, and nutritional stress.

7.6. COMPARISON WITH OTHER MODERN HUMAN STUDIES

To complete the analysis in the chapter, this section considers the isotopic results of this work in the context of other studies on modern human hair and nail. Other isotope studies relevant to this work are considered in the discussion. Figure 7.21 illustrates the isotopic results of 112 groups of hair samples, and 28 groups of nail samples, totalling over 2000 individual hair and nail samples, from more than 20 studies, from every geographical area in the world. Appendix D.1 contains the full details of all the studies included. This exercise is not an exhaustive account of all isotopic studies on hair and nail, but aims only to frame the results from the present work. Hülsemann et al. (2015) provide a much more complete report on this subject.

In order to interpret Figure 7.21 correctly, one must keep in mind that each point represents the mean value reported for a given group, as defined in the respective study. For instance, O'Connell and Hedges (1999) report data from three distinct groups, all from the United Kingdom, but each with a different diet, and hence, three data points are present in the figure, from this study alone. Consequently, this figure is biased towards studies that reported data from more groups, and since it shows only mean values, it fails to represent the true variability in isotopic results. In addition, this dataset is weighted towards Europe, simply because most of the published data hails from this region (Hülsemann et al., 2015).

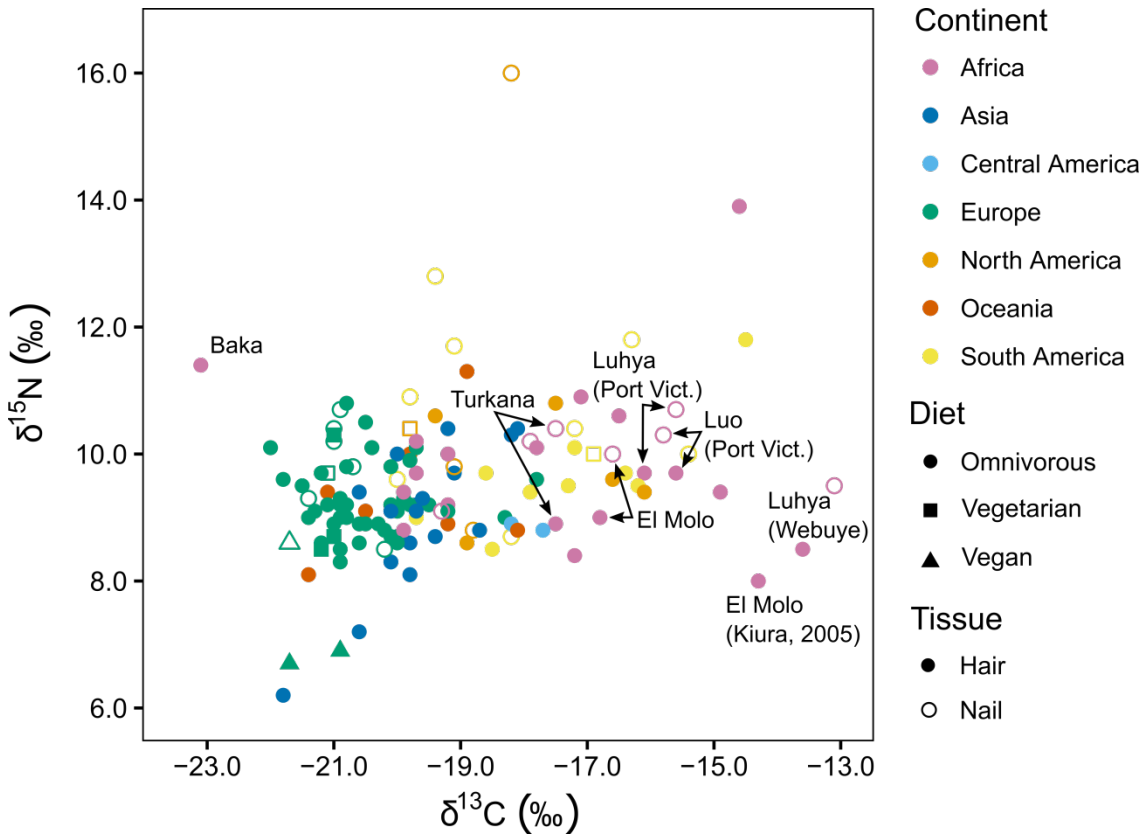


Figure 7.21 Carbon and nitrogen isotope ratios for human hair and nail, by continent and diet. Each data point represents the mean value for a group within a study. The labels indicate the data points belonging to this study, but also the one obtained for the El Molo by Kiura (2005).

Notwithstanding, Figure 7.21 provides important insights into the global distribution of isotopic values, which prove useful to the present work. On carbon, regions such as Africa and South America have higher $\delta^{13}C$ values than Europe or Asia. Latitude might explain this pattern, where $\delta^{13}C$ values decrease with increasing distance from the equator, which has been previously reported (Hülsemann et al., 2015; Valenzuela et al., 2012). These increased values result from the inclusion of higher amounts of C_4 plants in the diet, and, in turn, climatic factors, such as aridity and temperature, explain the distribution of C_4 plants in the ecosystem (Chapter 2, p. 37) (Bender et al., 1973; O'Brien, 2015; O'Leary, 1988; Schoeninger, 2010; Sealy, 2001; Smith & Epstein, 1971). Nevertheless, latitude is not the only relevant variable. For instance, $\delta^{13}C$ values from North America are also high due to the dominant use of C_4 plants as a dietary source for both humans and animals. On the other hand, the Baka, from this study, have the lowest carbon isotopic values from the entire dataset, despite their closeness to the equator. The forest environment where they live, dominated by C_3 plants, easily

clarifies this result, although the "canopy effect" might also contribute to the low carbon isotopic values. In addition, the highest carbon isotopic value in the data set also comes from the present work, more exactly, from the Luhya (Webuye).

Regarding $\delta^{15}N$ values, Figure 7.21 shows that most African data points (including the groups from this study) have a high nitrogen isotopic signal when compared to European values, despite the fact that they come from populations with comparatively low meat and fish consumption (Appendix D.1). Hülsemann et al. (2015) argue that this result relates to the ^{15}N enrichment observed in hot and arid environments (Amundson et al., 2003; Hartman, 2011). As previously discussed, this phenomenon explains the overall pattern, and perhaps even why the Luhya (Webuye), which live in a tropical climate (Chapter 3, p. 68), have lower nitrogen isotopic values than the other participants of this study, despite the fact that they consume relatively high amounts of animal foods (Table 5.11, p. 150). However, it does not explain the high nitrogen isotopic values observed in the Baka, which also inhabit a tropical environment. In fact, this population has nitrogen isotopic values that are above the main cloud of data points, even though they do not consume high amounts of animal foods (Table 5.11, p. 150). As stated by Oelze et al. (2011), this might result from the complex pathways of nitrogen isotopic fixation in tropical forests.

Other remarkable points in the data set include the very ^{15}N depleted vegans (O'Connell & Hedges, 1999a). These two data points are close to two omnivorous diets, one from Nepal, and another from India, which might be explained by low meat and fish intake in these areas (Hülsemann et al., 2015). On the opposite end, the clear outlier from North America represents an Inuit population, known for eating seals, a carnivorous marine mammal that is at the top of a very long food chain (Buchardt, Bunch, & Helin, 2007).

Finally, one must consider the results from the only other similar isotopic study on modern groups from Kenya, conducted by Kiura (2005). This work studied three modern populations in the eastern margin of Lake Turkana, and it conducted diet

interviews, documented food refuse, and analysed the carbon and nitrogen isotopic signals of human hair samples, as well as of the animals consumed by these people, and plants consumed by the animals. The groups studied were the Dassanech (agropastoralists), the Gabra (pastoralists), and the El Molo (fishers), the last of which overlaps with the present work. Compared to the present study, Kiura (2005) conducted a year-long study and collected more food samples, which allowed for a better understanding of diet variation, but conducted simpler diet interviews, which were harder to link with isotopic signals. For nitrogen isotopic results, Kiura (2005) also found that the El Molo had lower nitrogen isotopic values than expected, and that the agropastoralists Dassanech had higher nitrogen isotopic values than the pastoralists Gabra, which she explains by a consumption of both animal and plant products high in $\delta^{15}N$ (e.g. sorghum). For carbon isotopic results, the human hair signals matched the consumption of C_3 and C_4 plants for each group. In Figure 7.21, the data point from Africa with very high $\delta^{15}N$ belongs to this Dassenach population. This finding is consistent with the hypothesis that the low nitrogen isotopic values obtained for the Turkana result from a sedentarisation process that led to significant dietary change. Finally, Kiura (2005, pp. 312-313) report higher $\delta^{13}C$ (mean = −14.3‰, SD = 0.7‰, n = 20) and lower $\delta^{15}N$ (mean = +8.0‰, SD = 0.6‰, n = 20) for the El Molo than the present work. It is unclear why the results are so different. Perhaps the population has gone through a recent change in diet, such as a decrease in the consumption of C_4 foods and an increase in the consumption of animal foods. This would lend credit to the increased intake of animal foods detected in the semi-quantitative food frequency questionnaires, when compared to the ethnographic diet (Chapter 3, p. 59; Chapter 5, p. 149), initially attributed to high milk consumption and interviewer bias. On the other hand, it is also true that Kiura (2005) fails to report on any measure of quality of isotopic data, or, in fact, on the individual isotopic analysis, which renders further comparison difficult.

In brief, by contextualising the results of this work within other studies, it is brought to light the influence of environmental factors, such as latitude, precipitation, and temperature, on the isotopic patterns of modern hair and nail.

7.7. STABLE ISOTOPE RATIO ANALYSIS DISCUSSION

Stable isotope ratio analysis, just like dental microwear analysis, provide direct evidence on the diet of an individual during its lifetime. And in effect, these analyses were successful in distinguishing between some modern populations consuming different diets. Furthermore, this work added important information on the isotopic signals of tropical African populations. This previously identified gap in our knowledge (Hülsemann et al., 2015), is made all the clearer by the fact that one encompasses such a diverse and immense geographical area under one single call for more research.

As stressed throughout this work, this lack of investigation is even more important because so much of human evolution and diversification occurred in an African context, and thus research into these subjects could profit from a better understanding of the isotopic landscapes of this region. Indeed, several findings within this study are relevant to the analysis of archaeological human remains from an East African context, such as those found by the IN-AFRICA project.

First, and foremost, the inclusion of freshwater fish in the diets of this region is difficult to detect through the analysis of nitrogen isotopic results, since there are no signs of a clear trophic shift between fisher and non fisher populations. On one hand, Kiura (2005, p. 380) suggests that the low $\delta^{15}N$ values of Lake Turkana result from its high alkalinity and salinity (Vallerini & Ceccanti, 2011), which stimulate an increase in nitrogen-fixing blue green algae, the presence of which explains why Lake Turkana is sometimes called the Jade Sea. On the other hand, the lack of trophic shift might be magnified by the combined effects of low nitrogen isotopic values in fish with high nitrogen isotopic values in plants (Figure 7.1), and it is possible that the second was

absent in past populations. In addition, some bioaccumulation of ^{15}N was present with increasing amounts of fish consumption. Therefore, one may posit that lake fish consumption in past diets will be difficult to find, but not impossible, in particular if accompanied with isotopic analysis of archaeological fish remains. However, it will require careful data analysis, which might be hampered by an overreliance in a static concept of trophic shift. Nevertheless, analysis of such a potential staple could reveal key aspects of human evolution and adaptation (Erlandson, 2001).

Concurrently, the isotope results of fish and fisher populations from the two lakes (Lake Turkana and Lake Victoria) did not differ substantially, which suggests that the low nitrogen isotopic values are not a phenomenon unique to one lake. In fact, some evidence exists that this pattern extends to other African Great Lakes, such as Lake Tanganyika and Lake Malawi (Genner, Turner, Barker, & Hawkins, 1999; Gu, 2009; Hata et al., 2015).

Regarding the consumption of animal foods, no clear relationship was found between consumption of these foods and the nitrogen isotopic signal. This is puzzling considering that a higher proportion of the diets of all populations were made up of animal foods than of fish foods (Chapter 5, p. 133). In addition, it must be stressed that, with lower levels of ^{15}N in fish, it is probable that the isotopic signal of animals and fish to become entangled in past populations. To this, one must add the nitrogen isotopic enrichment observed in arid and dry areas. Albeit, in order to clarify this latter issue, one may cross isotopic information from different tissues that record different elements of the diet. In particular, the consumption of animal and fish staples will have a bigger effect on proteic tissues, when compared to other tissues, whereas the environmental enrichment will affect all tissues equally.

To further understand the differences in this area between the isotopic signals of an animal versus a fish diet, this study should include a population with distinctively high animal consumption. Initially, the Turkana, ethnographically pastoralists, were chosen as representative of such a population. However, the diet assessment and isotopic analysis revealed that this group has gone through a sedentarisation process

that led to the substitution of animal resources by C_4 plant foods. Notwithstanding, this was the population with the wider variability in nitrogen isotopic values, which indicated that individuals with very different diets were present in this group. This finding emphasizes the need to explore the variability within the data, as opposed to simply characterising a group through its mean and standard deviation.

Furthermore, one must consider previous isotopic work conducted by Stanley Ambrose in this region (Ambrose, 1986; Ambrose & DeNiro, 1986). Ambrose (1986) analysed 97 historic and prehistoric humans, representing 12 populations with different diets from East and South Africa. Combined use of carbon and nitrogen isotope analyses differentiated between pastoralists and farmers, but also between camel pastoralists and capri-bovine farmers, and grain farmers from non-grain farmers. Relevant to the present work, historic Turkana had much higher $\delta^{15}N$ levels (mean = +15.2‰, SD = 1.3‰, n = 7) and lower $\delta^{13}C$ levels (mean = −14.6‰, SD = 2.9‰, n = 7) than those reported here for hair and nail (Ambrose & DeNiro, 1986), which endorses a recent change in diet, involving a decrease in animal food consumption, and an increase in C_4 plant products. The same study on the collagen of historic remains also reports isotopic values for a single Luhya individual, although this affiliation is somewhat dubious. The $\delta^{15}N$ result of this individual is lower (mean = +7.7‰, n = 1) than that of the Luhya (Webuye) hair and nail, whereas the $\delta^{13}C$ is higher (mean = −11.6‰, n = 1), and even more distant than their Port Victoria counterparts. At face value, this could indicate that the Luhya (Webuye) are now consuming slightly higher amounts of animal foods and lower amounts of C_4 foods (possibly through decreased direct consumption). Nevertheless, Ambrose (1986, p. 725) also suggests that the low $\delta^{15}N$ value could result from reduced water stress and high rates of biological nitrogen fixation at the base of the food web. In addition, since this is a single sample, it can hardly be considered as a definite characterization of the isotopic signal of historical Luhya. Finally, one must consider the differing results for the El Molo between the present study and that of Kiura (2005), which indicate a recent increase in the consumption of animal foods, and a decrease in the consumption of C_4 foods.

These changes in diet highlight the variation observed not only horizontally within one population, but also vertically through time, which emphasizes the ability of individuals to adjust to new conditions and environments. Consistent with this, the Luhya from Webuye pursue a different diet from that of their ethnic counterparts in Port Victoria. In other words, people seem to eat what is available, and not necessarily what their ethnic tradition requires. Although this observation cannot be extended to all circumstances, it advises against inferring a cultural tie based on diet alone between populations separated by space or time. Nevertheless, one must consider that many of these recent changes in diet might result from a recent European intervention, which could have led to a weakening of cultural ties between the Luhya groups in Webuye and Port Victoria (Chapter 3, p. 69).

A final point on the isotopic results of this work concerns the low carbon isotopic results of the Baka. To the best of my knowledge, this is the first study to characterise a population with such a C_3 based diet, which might prove relevant to researchers investigating ancient populations that lived in similar conditions. In particular, it has long been suggested that human populations cannot subsist entirely on forest goods because these environments are difficult to navigate and lack adequate carbohydrate and protein resources (Chapter 3, p. 78) (Bailey et al., 1989; Headland & Bailey, 1991). Notwithstanding, Sato et al. (2012) proved that it is theoretically possible for the Baka hunter gatherers to obtain all necessary energy and nutrients from the forest, namely through the consumption of wild yam tubers, whereas Roberts et al. (2015) used carbon and oxygen isotope analyses to show that human foragers in Sri Lanka relied mostly on forest resources going back to the late Pleistocene. The findings from the present work add to this discussion by proving that the Baka still have a diet largely based on ^{13}C depleted C_3 plant foods.

Following this, one must address the difficulty in linking the data from the food frequency questionnaires with the isotopic data. Even though these two sets of data matched in a broad characterisation of diet, it was complicated to establish concrete links between the two. Clearly, this calls for a continued investigation into the isotopic

metabolism of humans, particularly when eating complex diets, which involve the exploitation of many different resources. With this goal, weighed records may be more appropriate to conduct such studies, considering that they record individual variability in diet better than semi-quantitative food frequency questionnaires (Gibson, 1990, p. 89). Furthermore, this work highlights the need to analyse isotopically as many components of the food web to which a targeted individual belongs, as well as different tissues of the same individuals, if one aims to have a comprehensive understanding of that individual's diet.

To end this discussion, one must focus on the potential of isotopic analysis as a measure of ecological niches. One of the most enticing aspects of isotopic analysis is that it can provide information on both the diet and the environment of an individual. In view of the refugia network system proposed by Mirazón Lahr (2013b) to explain the distribution of past populations across East Africa, isotopic analysis offers an exciting approach to explore and delimit these refugia.

7.8. SUMMARY OF CHAPTER 7

Summing up, this chapter delves into the effects of both dietary and environmental effects on isotopic values of the studied populations, and arrives at several important findings:

- The Baka have the lowest carbon isotopic values, very likely due to a diet based on C_3 plants, as well as to the "canopy effect", and the highest nitrogen isotopic values, for which there is no clear explanation, but that agree with other isotopic studies in the area;
- The Luhya (Webuye) have the highest carbon isotopic values, resulting from a diet containing a high proportion of C_4 plants, and lower nitrogen isotopic values than the other groups, perhaps due to the consumption of foods depleted in ^{15}N,

or though a ^{15}N enrichment in the other groups, resulting from low precipitation and high temperatures;

- The El Molo and the Turkana have carbon isotopic values befitting a mixed C_3 and C_4 diet, with only hair $\delta^{13}C$ distinguishing between the two groups, whereas the nitrogen isotopic values vary more widely for the Turkana, indicating that this group included some true pastoralists, but that sedentarisation led to a higher consumption of plant foods;

- The Luo and the Luhya have similar carbon and nitrogen isotopic values, where the carbon isotopic values detect a mixed C_3 and C_4 diet for both groups, but with a higher C_4 component than for the two previous populations, while the nitrogen isotopic values could be indicative of the higher fish consumption.

Overall, these results reveal that:

- Within an East African context, due to the high nitrogen isotopic content of plants and the low nitrogen isotopic content of fish, the consumption of lake fish might be difficult, but not impossible to detect;

- In all the ethnic groups where dietary information is available over the last 50 years (the El Molo, the Turkana, and the Luhya), evidence is found of a relatively recent change in diet; in addition, being close geographically leads to more similar isotopic values than having the same ethnic affiliation; taken together, this suggests that these groups adjust quickly to changing conditions, and that local factors influence isotopic signals;

- In order to achieve a more complete picture of the isotopic signal of a targeted organism, one must (1) sample different types of body pools; (2) explore the variability within the data; and (3) analyse all components of the food web.

CHAPTER 8

CONCLUSION

This study provided quantitative information on the diet of modern populations in tropical Africa. It addressed a gap on the isotopic and microwear signals of populations with well-known diets, a lack that is particularly stark within an African context. This chapter summarises the findings of the work, and discusses its implications for future work, both for diet assessment in the present and in the past.

8.1. ADDRESSING THE RESEARCH QUESTIONS

At the beginning of this document, the Introduction posed several research questions that guided the work conducted. This section considers each of these questions in turn, and examines the achieved answers.

8.1.1. DO MICROWEAR ANALYSES DISTINGUISH BETWEEN MODERN POPULATIONS CONSUMING DIFFERENT DIETS?

Unclear. Unfortunately, applying DMTA to living participants is more difficult than applying the same technique to archaeological or fossil material. In particular, the removal of bacterial biofilm proved to be challenging. This biofilm is the result of bacterial activity in the mouth and it adheres to the teeth of living organisms, obscuring dental microwear. This issue had already been identified in previous microwear research (Teaford & Oyen, 1989b), and the present work used a moulding method that specifically targeted biofilm removal (Livengood, 2015, personal communication). Nevertheless, many casts still had biofilm, decreasing the final sample size, and consequently, statistical analysis failed to detect differences between the groups. Notwithstanding, a visual analysis of the results suggests that microwear features vary between populations. In particular, the El Molo present microwear features that indicate a harder and tougher diet than the other groups, which could result from the inclusion of sand particles in fish during sun drying.

8.1.2. DO CARBON AND NITROGEN ISOTOPE ANALYSES DISTINGUISH BETWEEN MODERN POPULATIONS CONSUMING DIFFERENT DIETS?

Yes. Carbon and nitrogen isotope analysis distinguish between the groups of this study. In detail, the Baka hunter gatherers have low carbon isotopic values and relatively high nitrogen isotopic values, while the Luhya (Webuye) farmers have high carbon isotopic values, and relatively low nitrogen isotopic values. The remaining groups' isotopic values overlap to varying degrees, but it is possible to discern that the Turkana and the El Molo have slightly lower carbon and nitrogen isotopic values,

whereas the Luhya and Luo, from Port Victoria, have slightly higher carbon and nitrogen isotopic values. A diet based on C_3 plants, with an important contribution of meat and fish, explains the isotopic results of the Baka, whereas a C_4 based diet, with a smaller meat component, explains the isotopic results of the Luhya (Webuye). In the case of the Turkana, no clear trophic shift is present in nitrogen isotopic values in relation to other groups, which is likely to result from the sedentarisation of the Turkana, and the consequent shift towards a more plant-based diet. However, this group presents a wide variability in isotopic values, which suggests that the individuals included in this group have very different diets, perhaps due to the ongoing sedentarisation. Finally, the three fisher populations, the El Molo, and the Luo and the Luhya, from Port Victoria, overlap considerably in isotopic values, but this overlap cannot be entirely explained by fish consumption, since there is no trophic shift in nitrogen isotopic values. Instead, it is more likely due to a reliance on food staples with similar isotopic values.

8.1.3. USING THE COMBINED TECHNIQUES, CAN ONE ANSWER THE FOLLOWING?

Having discussed how the different groups separate using each technique individually, one can look at how the combined use informs on the diets of the studied people. To achieve this, this section is driven by research questions that target the diet itself, instead of focusing on the technique *per se*.

Do pastoralists consume significantly more animal products than other populations?

Unclear. Due to the sedentarisation of the Turkana, it is impossible to answer this question with definite certainty, but previous research suggests that the higher consumption of animal products can be detected among pastoralists through both isotopic and microwear analysis (Kiura, 2005; Schmidt et al., 2016). In the present work, individuals who are outliers in protein and fat consumption among the Turkana had the highest nitrogen isotopic values, confirming both that higher consumption of animal products leads to ^{15}N enrichment and that the low nitrogen isotopic values of the Turkana may result from the process of sedentarisation the population is going

through. On the other hand, the recent shift in diet towards famine relief foods is worth of note in itself. This pattern is present in most of the groups of this study, and it suggests a dietary change that differs from that associated with the spread of the Western diet, which is usually characterised by a decrease in the intake of plant foods, and an increase in the intake of animal foods. Instead, the Luhya (Port Victoria) might be a population whose diet is becoming more westernised, which would explain the increased nitrogen isotopic values. Although not necessarily relevant to the study of diet in the past, understanding this current shift might benefit research into the nutrition and health of these populations.

Do hunter gatherers rely mostly on C_3 forest produce?

Yes. In fact, comparing to previous research, the hunter gatherer Baka had the lowest carbon isotopic values, indicating a diet based entirely on C_3 plants, although this result could be compounded by the "canopy effect", which also contributes to low carbon isotopic values within a forest environment. This finding will be useful when analysing archaeological remains that could follow a similar diet, and could help frame the controversy on whether past foragers could subsist while relying mostly on forest produce (Bailey et al., 1989; Roberts et al., 2015). Considering that the Baka did not have dental moulds taken in the context of this study, this research question focused on the consumption of C_3 plants, which are detected through carbon isotopic analysis. Nevertheless, previous research suggests that this group has a particularly soft diet that leaves few microwear features on teeth, probably due to the processing and cooking of food (Romero et al., 2007).

Can one distinguish lake fishers from other populations?

Partially. The Luhya and the Luo, from Port Victoria were distinct from all other groups and very similar to each other, although often close to the El Molo, whereas the El Molo often overlapped with the Turkana, to whom they are closer to in space. This research question rested on the evidence that fish contributes more importantly to the protein part of the diet, and that ^{15}N enrichment is expected along the trophic chain. And accordingly, carbon isotopic values of breath cannot distinguish these groups

from the Turkana because breath records the averaged diet and not the protein part of the diet. In the same manner, the carbon isotopic values of hair distinguish the fishers from the others, but nail finds an overlap between the El Molo and the Turkana. In contrast, the nitrogen isotopic values of hair separate the two groups from Port Victoria from the El Molo and the Turkana, whereas the values of nail pool all four groups. Finally, a modest correlation is found between increasing fish consumption and increasing $\delta^{15}N$. These results have three important implications: (1) the baseline of the food chain has closer carbon and nitrogen isotopic values in areas closer to each other in space than in areas that are far apart; (2) no clear trophic shift is observable among these fisher groups, possibly due to the combined effects of a ^{15}N enrichment in plant foods, and ^{15}N depletion in the two lakes, Turkana and Victoria; and (3) different tissues can lead to different conclusions, and thus, one must bear in mind the metabolic pathways that are responsible for isotopic signals of each body pool.

Are fishers from different lakes distinguishable?

Partially. For carbon isotopic results, the El Molo never separate significantly from the two Port Victoria groups, except for hair results, where the El Molo separate from the Luo, but not from the Luhya. Incidentally, hair has a bigger sample size than nail, which could explain the higher number of significant comparisons. In turn, for nitrogen isotopic results, the El Molo separate from the Port Victoria groups in hair results, but pool with them in nail results. These results have the same three implications outlined above. In addition, they suggest that a clear trophic shift, commonly associated with fish consumption, is absent in at least two lakes in East Africa, which could hinder the detection of lake fish consumption in the past. Notwithstanding, considering that individual $\delta^{15}N$ values increase with higher fish consumption, and that $\delta^{15}N$ values of plants were probably not as high in the past, lake fish consumption may still be detected in such past contexts, but it will demand careful analysis of isotopic results.

In addition, the El Molo differ from all other groups in their dental microwear signal. In fact, they have the hardest and toughest signal of all groups, possibly as a result of

the inclusion of abrasives in the diet, through food processing habits. Thus, the combination of the techniques permits the separation of fisher from other groups, and between fishers from different lakes.

Is ethnic affiliation more important than diet or location?

No. In the present case, the Luhya from Port Victoria have a different diet from Luhya from Webuye, but a similar diet to the Luo from Port Victoria, which is detected through the diet questionnaires and stable isotope ratio analysis. Note that the dental microwear texture analysis also suggest a similar diet, but that, due to the low sample size, one cannot be sure of a true negative. This similarity in diet according to region, but not ethnic group, suggests that the geographical location of a population, and consequent food availability, are more important determinants of diet than ethnic affiliation. Admittedly, colonialism and the resulting weakening of cultural ties between the two Luhya groups could be driving this divergence in diet. A reasoning that is supported by the recent development and urbanization of the area, as well as the spread of Christianity (*First Bungoma County Integrated Development Plan 2013-2017*, 2013). Nevertheless, this finding serves as a cautionary tale against inferring similar diets based on cultural ties, and vice versa, against inferring separate cultural affiliations based on distinct diets alone.

8.1.4. ARE ISOTOPE AND MICROWEAR VALUES ASSOCIATED WITH THE RESULTS FROM DIET QUESTIONNAIRES?

Partially. Although it was possible to connect the results from the diet questionnaires with the isotopic results, both at a population level, and in some instances, at an individual level, these relationships were often less straightforward than desired. On the other hand, given the low sample size, the association between dental microwear and diet questionnaire results was only interrogated at a broad level, and even though some ties were found, the dental microwear results did not possess enough resolution to explore the relationship more fully.

As already mentioned, and as expected, increasing hair nitrogen isotopic values were associated with increasing fish intake. On the contrary, it was counterintuitive that increasing fish intake, as well as increasing animal intake, was associated with decreasing carbon isotopic values, given the measured fish samples. A possible explanation is that people in these regions replace meat and fish foods with C_4 foods, the latter obtained through famine relief programmes. In parallel, it is also surprising that no clear link was present between increasing animal intake and increasing nitrogen isotopic values across all participants, nor between increasing C_4 plant intake and increasing carbon isotopic values. These outcomes indicate that, in complex systems, as is the case of human diets, many factors play a role in the final isotopic signal of an individual. As another example, the decrease in plant $\delta^{15}N$ values (Figure 7.1) with increasing precipitation, and decreasing temperature could explain the low nitrogen isotopic values of the Luhya (Webuye) in relation to the other groups, when the diet questionnaires detect a higher intake of animals.

To conclude, the lack of certainty in linking a measured diet and consequent isotopic and microwear signals calls for more research within the context of human diets, as opposed to blindly applying principles gleaned from animal studies to archaeological human groups.

8.1.5. IS COMBINED ISOTOPIC AND MICROWEAR DATA FROM MODERN HUMAN POPULATIONS USEFUL FOR DIET ASSESSMENT IN THE PAST?

Yes. Due to methodological issues, the present work obtained better results from the stable isotope ratio analysis, than from the dental microwear texture analysis. Notwithstanding, information on microwear features suggests that the diet of the El Molo has distinct material properties from the diets of the other groups. Incidentally, this group overlapped with others in isotopic signals, namely the Turkana, and the Luo and Luhya, from Port Victoria. Thus, in a past context, the combination of the two techniques would allow the Turkana to be distinguished from the other groups. In addition, both techniques inform on different aspects of diet and

environment. Therefore, each technique can provide complementary data that contribute to a more inclusive view of diet, and associated behaviours. This point is developed further in Contributions.

8.2. FURTHER CONSIDERATIONS

After addressing the achieved responses to the research questions that guided this study, one may go beyond these questions and discuss other important points that arose throughout the work. These points encompass all aspects that revealed themselves relevant to the successful use of the methodologies of this study, i.e. semi−quantitative food frequency questionnaires, dental microwear texture analysis, and stable isotope ratio analysis.

First, even though semi−quantitative food frequency questionnaires are adequate to characterise the diet of a group, they are less appropriate when trying to describe the variability within the group, and in particular, when relating the individual diet to other individual measurements, such as isotopic or microwear results. Instead, weighed food records present a better alternative, although they are more time consuming, and more intrusive of the daily life of the participants (Gibson, 1990). Second, dental moulding in living participants for the purpose of dental microwear texture analysis would benefit from better and more thorough research into the cleaning of bacterial biofilm from teeth. Indeed, this technique is complex enough to require training in person, through, for instance, workshops. Unfortunately, it is also possible that complete biofilm removal cannot be attained (Teaford & Oyen, 1989b), especially in remote areas, without access to proper dental equipment, and thus that dental microwear texture analysis will remain unsuitable for living participants. Finally, for stable isotope ratio analysis, this research confirmed that, to improve diet assessment, one must consider the variability present in the data, instead of reducing the isotopic signal of a group to its mean, and that one must strive to analyse all the components of the food web, as failing to do so will produce an incomplete image of

the diet. Even more critical, is the finding that the isotopic signals of distinct tissues lead to different evaluations of diet, even when dealing with tissues that record the same proteic part of diet, such as hair and nail. This indicates that for an inclusive view of the isotopic diet, one must measure different tissues (Froehle et al., 2010), but also that isotopic research will benefit from a greater understanding of the metabolic pathways in the human body.

8.3. CONTRIBUTIONS

Having discussed the main findings of the present study, one may consider how these contribute to diet assessment, both in the present and in the past.

To begin, this work filled a gap in the isotopic and the dental microwear landscape of modern diets, by providing new data from tropical African origin. This is particularly relevant for carbon isotopic values, where two of the populations expanded the range of known isotopic values for modern diets: the Luhya (Webuye) had very high carbon isotopic values, whereas the Baka had very low carbon isotopic values, indicating diets that are largely based on C_4 and C_3 plants, respectively. In addition, this work confirmed that dental microwear texture analyses are useful in detecting food processing techniques. However, this research could not fully understand how dental microwear texture analysis and stable isotope ratio analysis related to each other, due to the low success rate of the first. Furthermore, dental microwear analyses may only be useful under certain circumstances. In the present work, the El Molo presented distinct microwear surfaces possibly because of the inclusion of grit in food during sundrying. This interpretation depended on the knowledge of the environmental context and dietary behaviours of the population. Similar situations are likely to be true to archaeological groups, where DMTA can only be appropriately interpreted when other lines of evidence indicate which were the conditions in which the group operated. This remark may be particularly true in the case of modern humans, as opposed to other primates, where diets are more likely to overlap between

populations, and where food processing techniques will alter the abrasiveness of the diet.

Nonetheless, it was clear that combining the two methods improved the resolution of the diet assessment, and that both stable isotope ratio analysis and dental microwear analysis inform on two important aspects of the ecological niche of an individual: the resources an individual uses, and the environment that it inhabits. In detail, stable isotope analyses informed on the plants and animals consumed, as well some aspects of the environment (e.g. canopy effect among the Baka), whereas dental microwear analyses informed on the some of the food processing techniques, and on the presence of windblown dust in the environment. In other words, research will often prove more fruitful when grounded in transdisciplinarity, where several lines of analysis contribute to a common picture.

Following this, these methods would be useful to the understanding of past populations, and of their relationships with their habitat, such as those found in the context of the IN−AFRICA Project. For instance, the focus on C_3 plants observed in hunter gatherers could make it possible to track populations with this subsistence strategy as they spread across East Africa at the onset of the Holocene, and left behind the refugia they inhabited during the Middle and Late Pleistocene. Furthermore, dental microwear texture analysis could detect whether these hunter gatherers shared food processing techniques, which would be plausible, under the proposal by Mirazón Lahr (2016) that these groups formed an extensive population or trading network across Africa. Conversely, past fishers of East African lakes should fail to present a clear trophic shift in relation to other groups, just as modern fishers do. Thus, the exploitation of lake fish cannot be detected or excluded based on stable isotopic ratio analysis alone, and instead, one must rely on other sources of evidence, such as the bone harpoons found by the IN-AFRICA Project (Mirazón Lahr & Foley, 2012, 2014). In addition, stable isotope ratio analysis and dental microwear analysis could also detect the arrival of pastoralists in East Africa around 4.5 Ka, through increased nitrogen isotopic values and lower microwear texture values (Kiura, 2005; Lane, 2013; Schmidt

et al., 2016). Finally, these two techniques could also reveal the fluctuations in subsistence patterns throughout the Holocene, or further back in time, as groups contracted into refugia during periods of shortage, and then spread across the landscape in times of abundance. Indeed, this work successfully detected changes in diet in historical times, as modern groups settle, abandon traditional diets, and start consuming higher amounts of famine relief food.

8.4. FUTURE WORK

To follow up on this work, two main avenues of research are left open, one at a macro, population level, and another at a micro, individual level. On one hand, it is necessary to collect more data on the isotopic and microwear variation of modern diets across the world. As Western diets, or even famine relief foods, expand their reach, a lot of the diversity of human diet is being lost, and with it, the opportunity to understand the significance of such subsistence strategies in a past context. Thus, as often stated throughout this study, research in diet assessment in the past would benefit from bigger data sets on current diets. On the other hand, it is necessary to understand better the processes that determine isotope and microwear signals. In other words, the study of diet assessment would also benefit from controlled studies within the individual on isotope metabolism and on the mechanics of chewing.

REFERENCES

Alemseged, Z., & Bobe, R. (2009). Diet in Early Hominin Species: A Paleoenvironmental Perspective. In J. Hublin & M. Richards (Eds.), *The Evolution of Hominin Diets - Integrating Approaches to the Study of Palaeolithic Subsistence* (pp. 181–188). Dordrecht, Netherlands: Springer.

Ambrose, S. (1986). Stable Carbon and Nitrogen Isotope Analysis of Human and Animal Diet in Africa. *Journal of Human Evolution*, *15*, 707–731.

Ambrose, S. (1990). Preparation and characterization of bone and tooth collagen for isotopic analysis. *Journal of Archaeological Science*, *17*(4), 431–451. https://doi.org/10.1016/0305-4403(90)90007-R

Ambrose, S., Butler, B., Hanson, D., Hunter-Anderson, R., & Krueger, H. (1997). Stable isotopic analysis of human diet in the Marianas Archipelago, western Pacific. *American Journal of Physical Anthropology*, *104*(3), 343–361.

Ambrose, S., & DeNiro, M. (1986). Reconstruction of African human diet using bone collagen carbon and nitrogen isotope ratios. *Nature*, *319*, 321–324. https://doi.org/10.1038/319321a0

Ambrose, S., & Norr, L. (1993). Experimental Evidence for the Relationship of the Carbon Isotope Ratios of Whole Diet and Dietary Protein to Those of Bone Collagen and Carbonate. In J. Lambert & G. Grupe (Eds.), *Prehistoric Human Bone - Archaeology at the Molecular Level* (pp. 1–37). Berlin, Germany: Springer-Verlag.

Amundson, R., Austin, A., Schuur, E., Yoo, K., Matzek, V., Kendall, C., Uebersax, A., Brenner, D., & Baisden, W. (2003). Global patterns of the isotopic composition of soil and plant nitrogen. *Global Biogeochemical Cycles*, *17*(1). https://doi.org/10.1029/2002GB001903

Andika, S., Mwangangi, S., & Kitaka, D. (2011). *Bunyala District Crops, Livestock and Fisheries Assessment Report for Medium and High Rainfall Areas*, Government of Kenya.

Aoba, T., & Fejerskov, O. (2002). Dental fluorosis: Chemistry and Biology. *Critical Reviews in Oral Biology & Medicine, 13*(2), 155–170.

Arman, S., Ungar, P., Brown, C., DeSantis, L., Schmidt, C., & Prideaux, G. (2016). Minimizing inter-microscope variability in dental microwear texture analysis. *Surface Topography: Metrology and Properties, 4*. https://doi.org/doi:10.1088/2051-672X/4/2/024007

Awaïss, A., Azeroual, A., Getahun, A., Hanssens, M., Lalèyè, P., Marshall, B., Moelants, T., Ntakimazi, G., & Tweddle, D. (2010). *Tilapia rendalli*. Retrieved from http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T60690A12387069.en

Ayliffe, L., Cerling, T., Robinson, T., West, A., Sponheimer, M., Passey, B., Hammer, J., Roeder, B., Dearing, M. D., & Ehleringer, J. (2004). Turnover of carbon isotopes in tail hair and breath CO_2 of horses fed an isotopically varied diet. *Oecologia, 139*(1), 11–22. https://doi.org/10.1007/s00442-003-1479-x

Azeroual, A., Entsua-Mensah, M., Getahun, A., Lalèyè, P., Moelants, T., & Ntakimazi, G. (2010). *Lates niloticus*. Retrieved from http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T181839A7746967.en

Bahuchet, S. (1993). History of the Inhabitans of the Central African rain Forest: Perspectives from Comparative Linguistics. In C. Hladik, A. Hladik, F. Linares, H. Pagezy, A. Semple, & M. Hadley (Eds.), *Tropical Forests, People and Food: Buocultural Interactions and Applications to Development* (pp. 37–54). Paris, France: UNESCO/The Parthenon Publishing Group.

Bahuchet, S. (2014). Cultural Diversity of African Pygmies. In B. Hewlett (Ed.), *Hunter-Gatherers of the Congo Basin: Cultures, Histories and Biology of African Pygmies* (pp. 1–29). New Jersey, USA: Transaction Publishers.

Bailey, R., Head, G., Jenike, M., Owen, B., Rechtman, R., & Zechenter, E. (1989). Hunting and gathering in tropical rain forest: Is it possible? *American Anthropologist, 91*(1), 59–82. https://doi.org/10.1525/aa.1989.91.1.02a00040

Barham, L., & Mitchell, P. (2008). *The First Africans - African Archaeology from the Earliest Toolmakers to most recent Foragers*. Cambridge, UK: Cambridge University Press.

Basell, L. (2008). Middle Stone Age (MSA) site distributions in eastern Africa and their relationship to Quaternary environmental change, refugia and the evolution of Homo sapiens. *Quaternary Science Reviews, 27*(27–28), 2484–2498. https://doi.org/10.1016/j.quascirev.2008.09.010

Batini, C., Lopes, J., Behar, D., Calafell, F., Jorde, L., Van Der Veen, L., Quintana-Murci, L., Spedini, G., Destro-Bisol, G., & Comas, D. (2011). Insights into the demographic history of African pygmies from complete mitochondrial genomes. *Molecular Biology and Evolution, 28*(2), 1099–1110. https://doi.org/10.1093/molbev/msq294

Bender, M., Rouhani, I., Vines, H., & Black, C. (1973). $^{13}C/^{12}C$ Ratio Changes in Crassulacean Acid Metabolism Plants. *Plant Physiology, 52*(5), 427–430. https://doi.org/10.1104/pp.52.5.427

Bersaglieri, T., Sabeti, P., Patterson, N., Vanderploeg, T., Schaffner, S., Drake, J., Rhodes, M., Reich, D., & Hirschhorn, J. (2004). Genetic signatures of strong recent positive selection at the lactase gene. *American Journal of Human Genetics, 74*(6), 1111–20. https://doi.org/10.1086/421051

Berthaume, M., Dumont, E., Godfrey, L., & Grosse, I. (2013). How does tooth cusp radius of curvature affect brittle food item processing? *Journal of The Royal Society Interface, 10*(84), 20130240. https://doi.org/10.1098/rsif.2013.0240

Betti, L., Balloux, F., Hanihara, T., & Manica, A. (2010). The relative role of drift and selection in shaping the human skull. *American Journal of Physical Anthropology, 141*(1), 76–82. https://doi.org/10.1002/ajpa.21115

Blumenschine, R., & Pobiner, B. (2007). Zooarchaeology and the Ecology of Oldowan Hominin Carnivory. In P. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (First Edition, pp. 167–189). Oxford, UK: Oxford University Press, Inc.

Boattini, A., Castrì, L., Sarno, S., Useli, A., Cioffi, M., Sazzini, M., Garagnani, P., De Fanti, S., Pettener, D., & Luiselli, D. (2013). MtDNA variation in East Africa unravels the history of afro-asiatic groups. *American Journal of Physical Anthropology, 150*(3), 375–385. https://doi.org/10.1002/ajpa.22212

Bocherens, H., & Drucker, D. (2003). Trophic level isotopic enrichment of carbon and nitrogen in bone collagen: Case studies from recent and ancient terrestrial ecosystems. *International Journal of Osteoarchaeology, 13*(1–2), 46–53. https://doi.org/10.1002/oa.662

Bonafini, M., Pellegrini, M., Ditchfield, P., & Pollard, A. (2013). Investigation of the "canopy effect" in the isotope ecology of temperate woodlands. *Journal of Archaeological Science, 40*(11), 3926–3935. https://doi.org/10.1016/j.jas.2013.03.028

Bonnefille, R., Roeland, J., & Guiot, J. (1990). Temperature and rainfall estimates for the past 40 000 years in equatorial Africa. *Nature, 346*, 347–349. https://doi.org/10.1038/346347a0

Brainard, J. (1991). *Health and Development in a Rural Kenyan Community*. New York, USA: Peter Lang.

Bringing food, animals or plants into the UK. (2014). Retrieved January 29, 2015, from https://www.gov.uk/bringing-food-animals-plants-into-uk/food

Buchardt, B., Bunch, V., & Helin, P. (2007). Fingernails and diet: Stable isotope signatures of a marine hunting community from modern Uummannaq, North Greenland. *Chemical Geology, 244*, 316–329. https://doi.org/10.1016/j.chemgeo.2007.06.022

Campbell, B., Leslie, P., Little, M., Brainard, J., & Deluca, M. (1999). Settled Turkana. In M. Little & P. Leslie (Eds.), *Turkana Herders of the Dry Savanna - Ecology and Behavioural response of Nomads to an Uncertain Environment* (pp. 333–352). Oxford, UK: Oxford University Press, Inc.

Campbell, L., Hecky, R., & Wandera, S. (2003). Stable Isotope Analyses of Food Web Structure and Fish Diet in Napoleon and Winam Gulfs, Lake Victoria, East Africa. *Journal of Great Lakes Research, 29*, 243–257. https://doi.org/10.1016/S0380-1330(03)70552-8

Cantalapiedra-Hijar, G., Ortigues-Marty, I., Sepchat, B., Agabriel, J., Huneau, J., & Fouillet, H. (2015). Diet–animal fractionation of nitrogen stable isotopes reflects the efficiency of nitrogen assimilation in ruminants. *British Journal of Nutrition, 113*, 1158–1169. https://doi.org/10.1017/S0007114514004449

Castrì, L., Garagnani, P., Useli, A., Pettener, D., & Luiselli, D. (2008). Kenyan crossroads: Migration and gene flow in six ethnic groups from Eastern Africa. *Journal of Anthropological Sciences, 86*, 189–192.

Caut, S., Angulo, E., & Courchamp, F. (2009). Variation in discrimination factors ($\Delta^{15}N$ and $\Delta^{13}C$): The effect of diet isotopic values and applications for diet reconstruction. *Journal of Applied Ecology, 46*, 443–453. https://doi.org/10.1111/j.1365-2664.2009.01620.x

Cerling, T., Manthi, F., Mbua, E., Leakey, L., Leakey, M., Leakey, R., Brown, F., Grine, F., Hart, J., Kaleme, P., Roche, H., Uno, K., & Wood, B. (2013). Stable isotope-based diet reconstructions of Turkana Basin hominins. *Proceedings of the National Academy of Sciences of the United States of America, 110*(26), 10501–10506. https://doi.org/10.1073/pnas.1222568110

Cernusak, L. A., Ubierna, N., Winter, K., Holtum, J. A. M., Marshall, J. D., & Farquhar, G. D. (2013). Environmental and physiological determinants of carbon isotope discrimination in terrestrial plants. *The New Phytologist, 200*(4), 950–65. https://doi.org/10.1111/nph.12423

Chisholm, B., Nelson, D., & Schwarcz, H. (1982). Stable-carbon isotope ratios as a measure of marine versus terrestrial protein in ancient diets. *Science, 216*(4550), 1131–1132. https://doi.org/10.1126/science.216.4550.1131

Climate-data.org. (2017a). Busia. Retrieved February 4, 2017, from https://en.climate-data.org/location/11168/

Climate-data.org. (2017b). Lodwar. Retrieved January 22, 2017, from https://en.climate-data.org/location/11143/

Climate-data.org. (2017c). Lomié. Retrieved February 5, 2017, from https://en.climate-data.org/location/897231/

Climate-data.org. (2017d). Webuye. Retrieved January 25, 2017, from https://en.climate-data.org/location/11118/

Clist, B. (1987). A Critical Reappraisal of the Chronological Framework of the Early Urewe Iron Age Industry. *Muntu, 6*, 35–62.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (Second Edition). Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates.

Compton, J. (2011). Pleistocene sea-level fluctuations and human evolution on the southern coastal plain of South Africa. *Quaternary Science Reviews*, *30*(5–6), 506–527. https://doi.org/10.1016/j.quascirev.2010.12.012

Copeland, S. (2009). Potential hominin plant foods in northern Tanzania: semi-arid savannas versus savanna chimpanzee sites. *Journal of Human Evolution, 57*(4), 365–78. https://doi.org/10.1016/j.jhevol.2009.06.007

Copeland, S., Sponheimer, M., Roux, P., Grimes, V., Lee-Thorp, J., Ruiter, D., & Richards, M. (2008). Strontium isotope ratios ($^{87}Sr/^{86}Sr$) of tooth enamel: a comparison of solution and laser ablation multicollector inductively coupled plasma mass spectrometry methods. *Rapid Communications in Mass Spectrometry, 22*, 3187–3194. https://doi.org/10.1002/rcm.3717

Crowley, B., Carter, M., Karpanty, S., Zihlman, A., Koch, P., & Dominy, N. (2010). Stable carbon and nitrogen isotope enrichment in primate tissues. *Oecologia*, *164*, 611–626. https://doi.org/10.1007/s00442-010-1701-6

Daegling, D. (2010). Understanding Skull Function from a Mechanobiological Perspective. In C. S. Larsen (Ed.), *A Companion to Biological Anthropology* (pp. 501–515). Oxford, UK: Blackwell Publishing Ltd.

Darwin, C. (2009). *The descent of Man, and Selection in Relation to Sex*. New York, USA: Cambridge University Press. (Original work published in 1871)

De Nardo, R., Chiappe, V., Gómez, M., Romanelli, H., & Slots, J. (2012). Effects of 0.05% sodium hypochlorite oral rinse on supragingival biofilm and gingival inflammation. *International Dental Journal, 62*, 208–212. https://doi.org/10.1111/j.1875-595X.2011.00111.x

DeNiro, M. (1985). Postmortem preservation and alteration of in vivo bone collagen isotope ratios in relation to palaeodietary reconstruction. *Nature, 317*(6040), 806–809. https://doi.org/10.1038/317806a0

DeNiro, M., & Epstein, S. (1977). Mechanism of carbon isotope fractionation associated with lipid synthesis. *Science, 197*, 261–263. https://doi.org/10.1126/science.327543

DeNiro, M., & Epstein, S. (1978). Influence of diet on the distribution of carbon isotopes in animals. *Geochimica et Cosmochimica Acta, 42*, 495–506. https://doi.org/10.1016/0016-7037(78)90199-0

DeNiro, M., & Epstein, S. (1981). Influence of diet on the distribution of nitrogen isotopes in animals. *Geochimica et Cosmochimica Acta, 45*(3), 341–351. https://doi.org/10.1016/0016-7037(81)90244-1

Dennis, J., Ungar, P., Teaford, M., & Glander, K. (2004). Dental topography and molar wear in Alouatta palliata from Costa Rica. *American Journal of Physical Anthropology, 125*(2), 152–61. https://doi.org/10.1002/ajpa.10379

DeSantis, L., Scott, J., Schubert, B., Donohue, S., McCray, B., Van Stolk, C., Winburn, A., Greshko, M., & O'Hara, M. (2013). Direct Comparisons of 2D and 3D Dental Microwear Proxies in Extant Herbivorous and Carnivorous Mammals. *PLoS ONE, 8*(8), 1–11. https://doi.org/10.1371/journal.pone.0071428

DuPré, C. (1968). *The Luo of Kenya: An Annotated Bibliography*. Washington D.C, USA: Institute for Cross Cultural Research.

Durnin, J., Edholm, O., Miller, D., & Waterlow, J. (1973). How Much Food Does Man Require? *Nature, 242*, 418.

Dyble, M., Thompson, J., Smith, D., Salali, G., Chaudhary, N., Page, A., Vinicius, L., Mace, R., & Migliano, A. (2016). Networks of Food Sharing Reveal the Functional Significance of Multilevel Sociality in Two Hunter-Gatherer Groups. *Current Biology, 26*(15), 2017–2021. https://doi.org/10.1016/j.cub.2016.05.064

Dyson-Hudson, N., & Dyson-Hudson, R. (1999). The social organization of resource exploitation. In M. Little & P. Leslie (Eds.), *Turkana Herders of the Dry Savanna - Ecology and Behavioural response of Nomads to an Uncertain Environment* (pp. 69–86). Oxford, UK: Oxford University Press, Inc.

Dyson-Hudson, R. (1999). Turkana in time perspective. In M. Little & P. Leslie (Eds.), *Turkana Herders of the Dry Savanna - Ecology and Behavioural response of Nomads to an Uncertain Environment* (pp. 25–40). Oxford, UK: Oxford University Press, Inc.

Ehret, C. (1998). *An African Classical Age: Eastern and Southern Africa in World History, 1000B.C. to A.D.400*. Charlottesville, USA: The University Press of Virginia.

Ehret, C. (2002). *The Civilizations of Africa: A History to 1800*. Oxford, UK: James Currey Publishers.

El Zaatari, S. (2010). Occlusal microwear texture analysis and the diets of historical/prehistoric hunter-gatherers. *International Journal of Osteoarchaeology, 20*, 67–87. https://doi.org/10.1002/oa.1027

Ellis, J., Galvin, K., McCabe, J. T., & Swift, D. (1987). *Pastoralism and drought in Turkana District, Kenya. Report to the Norwegian Agency for International Development*. Bellvue, Colorado.

Enattah, N., Jensen, T., Nielsen, M., Lewinski, R., Kuokkanen, M., Rasinpera, H., El-Shanti, H., Seo, J. K., Alifrangis, M., Khalil, I., Natah, A., Ali, A., Natah, S., Comas, D., Mehdi, S., Groop, L., Vestergaard, E., Imtiaz, F., Rashed, M., Meyer, B., Troelsen, J., & Peltonen, L. (2008). Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *American Journal of Human Genetics, 82*(1), 57–72. https://doi.org/10.1016/j.ajhg.2007.09.012

Erhardt, E., & Bedrick, E. (2012). A Bayesian framework for stable isotope mixing models. *Environmental and Ecological Statistics, 20*(3), 377–397. https://doi.org/10.1007/s10651-012-0224-1

Erlandson, J. (2001). The Archaeology of Aquatic Adaptations: Paradigm for a New Millennium. *Journal of Archaeological Research, 9*(4), 287–350. https://doi.org/10.1023/A:1013062712695

Fakhouri, T., Jahren, A. H., Appel, L., Chen, L., Alavi, R., & Anderson, C. (2014). Serum carbon isotope values change in adults in response to changes in sugar-sweetened beverage intake. *The Journal of Nutrition, 144*(6), 902–5. https://doi.org/10.3945/jn.113.186213

FAO. (2003). *Food energy – Methods of Analysis and Conversion Factors*. Geneva, Switzerland. https://doi.org/ISSN 0254-4725

FAO. (2010). *Fats and fatty acids in human nutrition - Report of a FAO Expert Consultation* (Vol. 91). Geneva, Switzerland. https://doi.org/978-92-5-106733-8

FAO. (2016). *FAO/INFOODS Analytical Food Composition Database Version 1.1 - AnFooD1.1*. Rome, Italy. Retrieved from http://www.fao.org/infoods/infoods/tables-and-databases/faoinfoods-databases/en/

FAO/WHO. (1998). *Carbohydrates in Human Nutrition - Report of a Joint FAO/WHO Expert Consultation*. Rome, Italy. Retrieved from http://www.who.int/nutrition/publications/nutrientrequirements/9251041148/en/

Farquhar, G., Ehleringer, J., & Hubick, K. (1989). Carbon Isotope Discrimination and Photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, *40*, 503–537. https://doi.org/1040-2519/89/0601-503

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fernandes, R., Millard, A., Brabec, M., Nadeau, M.-J., & Grootes, P. (2014). Food reconstruction using isotopic transferred signals (FRUITS): a Bayesian model for diet reconstruction. *PloS ONE, 9*(2), e87436. https://doi.org/10.1371/journal.pone.0087436

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics using R*. London, United Kingdom: SAGE Publications Ltd.

Fiorenza, L., Benazzi, S., & Kullmer, O. (2009). Do it yourself: Morphology, wear and 3D digital surface models: Materials and techniques to create high-resolution replicas of teeth. *Journal of Anthropological Sciences, 87*, 211–218.

Fiorenza, L., Benazzi, S., Tausch, J., Kullmer, O., Bromage, T., & Schrenk, F. (2011). Molar macrowear reveals Neanderthal eco-geographic dietary variation. *PloS ONE, 6*(3), e14769. https://doi.org/10.1371/journal.pone.0014769

First Bungoma County Integrated Development Plan 2013-2017. (2013). The County Government of Kenya

FishBase team RMCA & Geelhand, D. (2016). *Rastrineobola argentea*. https://doi.org/http://dx.doi.org/10.2305/IUCN.UK.2016- 3.RLTS.T61257A47242399.en

Fleagle, J. (2013). *Primate adaptation and evolution*. Amsterdam; Boston: Elsevier/Academic.

Fogel, M., & Tuross, N. (2003). Extending the limits of paleodietary studies of humans with compound specific carbon isotope analysis of amino acids. *Journal of Archaeological Science, 30*, 535–545. https://doi.org/10.1016/S0305-4403(02)00199-1

Foley, R. (2013). Hominin evolution as the context for African prehistory. In P. Mitchell & P. Lane (Eds.), *The Oxford Handbook of African Archaeology* (First Edition, pp. 269–287). Oxford, UK: Oxford University Press, Inc.

Ford, V. (1955). *The Trade of Lake Victoria: A Geographical Study*. Kampala, Uganda: East African Institute of Social Research.

Froehle, A., Kellner, C., & Schoeninger, M. (2010). FOCUS: Effect of diet and protein source on carbon stable isotope ratios in collagen: follow up to Warinner and Tuross (2009). *Journal of Archaeological Science, 37*(10), 2662–2670. https://doi.org/10.1016/j.jas.2010.06.003

Froehle, A., Kellner, C., & Schoeninger, M. (2012). Multivariate carbon and nitrogen stable isotope model for the reconstruction of prehistoric human diet. *American Journal of Physical Anthropology, 147*(3), 352–369. https://doi.org/10.1002/ajpa.21651

Fuller, B., Fuller, J., Harris, D., & Hedges, R. (2006). Detection of breastfeeding and weaning in modern human infants with carbon and nitrogen stable isotope ratios. *American Journal of Physical Anthropology, 129*(2), 279–293. https://doi.org/10.1002/ajpa.20249

Fuller, B., Fuller, J., Sage, N., Harris, D., O'Connell, T., & Hedges, R. (2004a). Nitrogen balance and $\delta^{15}N$: Why you're not what you eat during nutritional stress. *Rapid Communications in Mass Spectrometry, 18*(23), 2889–2896. https://doi.org/10.1002/rcm.1708

Fuller, B., Fuller, J., Sage, N., Harris, D., O'Connell, T., & Hedges, R. (2004b). Nitrogen balance and $\delta^{15}N$: Why you're not what you eat during pregnancy. *Rapid Communications in Mass Spectrometry, 18*(23), 2889–2896. https://doi.org/10.1002/rcm.1708

Gaciri, S., & Davies, T. (1993). The occurrence and geochemistry of fluoride in some natural waters of Kenya. *Journal of Hydrology, 143*(3–4), 395–412. https://doi.org/10.1016/0022-1694(93)90201-J

Galbany, J., Estebaranz, F., Martínez, L., Romero, A., De Juan, J., Turbón, D., & Pérez-Pérez, A. (2006). Comparative analysis of dental enamel polyvinylsiloxane impression and polyurethane casting methods for SEM research. *Microscopy Research and Technique, 69*, 246–252. https://doi.org/10.1002/jemt.20296

Galbany, J., Martínez, L., & Perez-Perez, A. (2004). Tooth Replication Techniques, Sem Imaging and Microwear Analysis in Primates: Methodological Obstacles. *Anthropologie, XLII*(1), 5–12.

Galvin, K. (1985). *Food Procurement, Diet, Activities and Nutrition of Ngisonyoka, Turkana Pastoralists in and Ecological and Social Context*. PhD Thesis. State University of New York.

Galvin, K., & Little, M. (1999). Dietary intake and nutritional status. In M. Little & P. Leslie (Eds.), *Turkana Herders of the Dry Savanna - Ecology and Behavioural response of Nomads to an Uncertain Environment* (pp. 126–145). Oxford, UK: Oxford University Press, Inc.

Genner, M., Turner, G., Barker, S., & Hawkins, S. (1999). Niche segregation among Lake Malawi cichlid fishes? Evidence from stable isotope signatures. *Ecology Letters, 2*, 185–190. https://doi.org/10.1046/j.1461-0248.1999.00068.x

Gibson, R. (1990). *Principles of Nutritional Assessment*. Oxford, UK: Oxford University Press, Inc.

Gibson, R. (1993). *Nutritional Assessment - A Laboratory Manual*. Oxford, UK: Oxford University Press, Inc.

González-José, R., Ramírez-Rozzi, F., Sardi, M., Martínez-Abadías, N., Hernández, M., & Pucciarelli, H. (2005). Functional-cranial approach to the influence of economic strategy on skull morphology. *American Journal of Physical Anthropology, 128*(4), 757–71. https://doi.org/10.1002/ajpa.20161

Goodall, R., Darras, L., & Purnell, M. (2015). Accuracy and precision of silicon based impression media for quantitative areal texture analysis. *Scientific Reports*, *5*, 10800. https://doi.org/10.1038/srep10800

Gordon, K. D. (1988). A review of methodology and quantification in dental microwear analysis. *Scanning Microscopy*, *2*(202), 1139–1147.

Gownaris, N., Pikitch, E., Ojwang, W., Michener, R., & Kaufman, L. (2015). Predicting Species' Vulnerability in a Massively Perturbed System: The Fishes of Lake Turkana, Kenya. *PLoS ONE*, *10*(5), 1–24. https://doi.org/10.1371/journal.pone.0127027

Grine, F., Sponheimer, M., Ungar, P., Lee-Thorp, J., & Teaford, M. (2012). Dental microwear and stable isotopes inform the paleoecology of extinct hominins. *American Journal of Physical Anthropology*, *148*(2), 285–317. https://doi.org/10.1002/ajpa.22086

Grine, F., Ungar, P., & Teaford, M. (2002). Error Rates in Dental Microwear Quantification Using Scanning Electron Microscopy. *Scanning*, *24*, 144–153. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15712754

Gu, B. (2009). Variations and controls of nitrogen stable isotopes in particulate organic matter of lakes. *Oecologia*, *160*(3), 421–431. https://doi.org/10.1007/s00442-009-1323-z

Hare, P. E., Fogel, M., Stafford Jr., T., Mitchell, A., & Hoering, T. (1991). The isotopic composition of carbon and nitrogen in individual amino acids isolated from modern and fossil proteins. *Journal of Archaeological Science*, *18*, 277–292. https://doi.org/10.1016/0305-4403(91)90066-X

Hartman, G. (2011). Are elevated $\delta^{15}N$ values in herbivores in hot and arid environments caused by diet or animal physiology? *Functional Ecology*, *25*(1), 122–131. https://doi.org/10.1111/j.1365-2435.2010.01782.x

Harvati, K., & Weaver, T. (2006). Human cranial anatomy and the differential preservation of population history and climate signatures. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology, 288*, 1225–1233. https://doi.org/10.1002/ar.a.20395

Hata, H., Shibata, J., Omori, K., Kohda, M., Hori, M., Bunkyo, M., & Kagamiyama, H.-H. (2015). Depth segregation and diet disparity revealed by stable isotope analyses in sympatric herbivorous cichlids in Lake Tanganyika. *Zoological Letters, 1*(1), 15. https://doi.org/10.1186/s40851-015-0016-1

Hayward, D. (1984). *The Arbore Language: A First Investigation*. Hamburg, Germany: Helmut Buske Verlag.

Headland, T., & Bailey, R. (1991). Introduction: Have hunter-gatherers ever lived in tropical rain forest independently of agriculture? *Human Ecology*, *19*(2), 115–122. https://doi.org/10.1007/BF00888742

Heaton, T. (1987). The $^{15}N/^{14}N$ ratios of plants in South Africa and Namibia: relationship to climate and coastal/saline environments. *Oecologia, 74*, 236–246.

Hedges, R., Clement, J., Thomas, D., & O'Connell, T. (2007). Collagen Turnover in the Adult Femoral Mid-Shaft: Modeled From Anthropogenic Radiocarbon Tracer Measurements. *American Journal of Physical Anthropology, 133*, 808–816. https://doi.org/10.1002/ajpa.20598

Hedges, R., & Reynard, L. (2007). Nitrogen isotopes and the trophic level of humans in archaeology. *Journal of Archaeological Science, 34*(8), 1240–1251. https://doi.org/10.1016/j.jas.2006.10.015

Hedges, R., Rush, E., & Aalbersberg, W. (2009). Correspondence between human diet, body composition and stable isotopic composition of hair and breath in Fijian villagers. *Isotopes in Environmental and Health Studies, 45*(1), 1–17. https://doi.org/10.1080/10256010802522010

Hedges, R., & van Klinken, G. (2000). "Consider a Spherical Cow..."—on Modeling and Diet. In S. Ambrose & A. Katzenberg (Eds.), *Biogeochemical Approaches to Paleodietary Analysis* (pp. 211–241). New York, USA: Kluwer Academic/Plenum Publishers.

Heine, B. (1980). The non-Bantu Languages of Kenya. In B. Heine & W. Möhlig (Eds.), *Language and Dialect Atlas of Kenya, Volume II* (pp. 173–218). Berlin, Germany: Dietrich Reimer Verlag.

Henry, A. (2012). Recovering Dietary Information from Extant and Extinct Primates Using Plant Microremains. *International Journal of Primatology, 33*(3), 702–715. https://doi.org/10.1007/s10764-011-9556-1

Henry, A., Brooks, A., & Piperno, D. (2011). Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). *Proceedings of the National Academy of Sciences of the United States of America, 108*(2), 486–91. https://doi.org/10.1073/pnas.1016868108

Hewlett, B. (1996). Cultural diversity among African Pygmies. In S. Kent (Ed.), *Cultural diversity among twentieth-century foragers* (pp. 215–244). Cambridge, UK: Cambridge University Press.

Hoefs, J. (2004). *Stable Isotope Geochemistry*. Berlin, Germany: Springer-Verlag.

Hoppe, K., Koch, P., & Furutani, T. (2003). Assessing the preservation of biogenic strontium in fossil bones and tooth enamel. *International Journal of Osteoarchaeology*, 13(1–2), 20–28. https://doi.org/10.1002/oa.663

Houpt, T. R. (1963). Urea utilization by rabbits fed a low protein ration. *American Journal of Physiology, 205*, 1144–1150.

Houpt, T. R., & Houpt, K. (1971). Nitrogen conservation by ponies fed a low -protein ration. *American Journal of Veterinary Research, 32*(4), 579–588.

Hua, L.-C., Brandt, E., Meullenet, J.-F., Zhou, Z.-R., & Ungar, P. (2015). Technical note: An in vitro study of dental microwear formation using the BITE Master II chewing machine. *American Journal of Physical Anthropology, 158*(4), 769–775. https://doi.org/10.1002/ajpa.22823

Hülsemann, F., Flenker, U., Koehler, K., & Schaenzer, W. (2009). Effect of a controlled dietary change on carbon and nitrogen stable isotope ratios of human hair. *Rapid Communications in Mass Spectrometry, 23*, 2448–2454. https://doi.org/10.1002/rcm.4039

Hülsemann, F., Lehn, C., Schneider, S., Jackson, G., Hill, S., Rossmann, A., Scheid, N., Dunn, P. J. H., Flenker, U., & Schänzer, W. (2015). Global spatial distributions of nitrogen and carbon stable isotope ratios of modern human hair. *Rapid Communications in Mass Spectrometry, 29*(22), 2111–21. https://doi.org/10.1002/rcm.7370

Humphrey, L., De Groote, I., Morales, J., Barton, N., Collcutt, S., Ramsey, C., & Bouzouggar, A. (2014). Earliest evidence for caries and exploitation of starchy plant foods in Pleistocene hunter-gatherers from Morocco. *Proceedings of the National Academy of Sciences of the United States of America, 111*(3), 954–9. https://doi.org/10.1073/pnas.1318176111

Hunter, D., Sampson, L., Stampfer, M., Colditz, G., Rosner, B., & Willett, W. (1988). Variability in portion sizes of commonly consumed foods among a population of women in the United States. *American Journal of Epidemiology, 127*(6), 1240–1249.

Ingram, C., Mulcare, C., Itan, Y., Thomas, M., & Swallow, D. (2009). Lactose digestion and the evolutionary genetics of lactase persistence. *Human Genetics, 124*, 579–591. https://doi.org/10.1007/s00439-008-0593-6

Jeffery, J. (2018). *How does residential mobility affect body size? A comparison of growth and development across three northern Kenyan livelihoods*. PhD Thesis. University of Cambridge.

Jim, S., Ambrose, S., & Evershed, R. (2004). Stable carbon isotopic evidence for differences in the dietary origin of bone cholesterol, collagen and apatite: Implications for their use in palaeodietary reconstruction. *Geochimica et Cosmochimica Acta, 68*(1), 61–72. https://doi.org/10.1016/S0016-7037(03)00216-3

Jim, S., Jones, V., Ambrose, S., & Evershed, R. (2006). Quantifying dietary macronutrient sources of carbon for bone collagen biosynthesis using natural abundance stable carbon isotope analysis. *British Journal of Nutrition, 95*(6), 1055–1062. https://doi.org/doi:10.1079/BJN20051685

Johnson, A. (2013). Exploring Adaptive Variation among Hunter-gatherers with Binford's Frames of Reference. *Journal of Archaeological Research, 22*(1), 1–42. https://doi.org/10.1007/s10814-013-9068-y

Joiris, D. (1996). A comparative approach to hunting rituals among Baka Pygmies (southeastern Cameroon). In S. Kent (Ed.), *Cultural diversity among twentieth-century foragers* (pp. 245–275). Cambridge, UK: Cambridge University Press.

Kanyoro, R. (1983). *Unity in diversity: A linguistic survey of the Abaluyia of Western Kenya*. Vienna, Austria: Afro-Pub.

Kay, R. (1975). Allometry and Early Hominids (comment). *Science, 189*, 61–64.

Kay, R. (1977). The evolution of molar occlusion in the Cercopithecidae and early Catarrhines. *American Journal of Physical Anthropology, 46*, 327–352. https://doi.org/10.1002/ajpa.1330460213

Kay, R., & Hiiemae, K. (1974). Jaw movement and tooth use in recent and fossil primates. *American Journal of Physical Anthropology, 40*(2), 227–256. https://doi.org/10.1002/ajpa.1330400210

Keeley, J., & Sandquist, D. (1992). Carbon: freshwater plants. *Plant Cell and Environment, 15*(9), 1021–1035. https://doi.org/10.1111/j.1365-3040.1992.tb01653.x

Kellner, C., & Schoeninger, M. (2007). A Simple Carbon Isotope Model for Reconstructing Prehistoric Human Diet. *American Journal of Physical Anthropology, 133*, 1112–1127. https://doi.org/10.1002/ajpa

Kelly, R. (2013). *The Lifeways of Hunter-Gatherers - The Foraging Spectrum* (2nd ed.). Cambridge, UK: Cambridge University Press.

Kiage, L., & Liu, K. (2006). Late Quaternary paleoenvironmental changes in East Africa: a review of multiproxy evidence from palynology, lake sediments, and associated records. *Progress in Physical Geography, 30*(5), 633–658. https://doi.org/10.1177/0309133306071146

King, C., Tayles, N., & Gordon, K. (2011). Re-examining the chemical evaluation of diagenesis in human bone apatite. *Journal of Archaeological Science, 38*(9), 2222–2230. https://doi.org/10.1016/j.jas.2011.03.023

Kiura, P. (2005). *An Ethnoarchaeology and Stable Isotope Study on the Diets of Three Modern Groups of People in Northern Kenya*. PhD Thesis. The State University of New Jersey.

Klaus, B. (2013). Effect Size Estimation and Misclassification Rate Based Variable Selection in Linear Discriminant Analysis. *Journal of Data Science, 11*, 537–558.

Knösel, M., Bojes, M., Jung, K., & Ziebolz, D. (2012). Increased susceptibility for white spot lesions by surplus orthodontic etching exceeding bracket base area. *American Journal of Orthodontics and Dentofacial Orthopedics, 141*(5), 574–582. https://doi.org/10.1016/j.ajodo.2011.11.017

Krishnamachari, K. (1986). Skeletal fluorosis in humans: a review of recent progress in the understanding of the disease. *Progress in Food & Nutrition Science, 10*(3–4), 279–314.

Krueger, K., Scott, J., Kay, R., & Ungar, P. (2008). Technical note: Dental microwear textures of "Phase I" and "Phase II" facets. *American Journal of Physical Anthropology, 137*(May), 485–490. https://doi.org/10.1002/ajpa.20928

Kusimba, S. (2013). Hunter-Gatherer-Fishers of Eastern and South-Central Africa Since 20,000 Years Ago. In P. Mitchell & P. J. Lane (Eds.), *The Oxford Handbook of African Archaeology* (First Edition, pp. 1–10). Oxford University Press, Inc. https://doi.org/10.1093/oxfordhb/9780199569885.013.0032

Lambert, J., & Rothman, J. (2015). Fallback Foods, Optimal Diets, and Nutritional Targets: Primate Responses to Varying Food Availability and Quality. *Annual Review of Anthropology*, *44*(1), 493–512. https://doi.org/10.1146/annurev-anthro-102313-025928

Lamphear, J. (1988). The People of the Grey Bull: The Origin and Expansion of the Turkana. *The Journal of African History*, *29*(1), 27–39.

Lane, P. (2013). The archaeology of pastoralism and stock-keeping in East Africa. In P. Mitchell & P. Lane (Eds.), *The Oxford Handbook of African Archaeology* (First Edition, pp. 585–601). Oxford, UK: Oxford University Press, Inc.

Lee-Thorp, J., & Sponheimer, M. (2006). Contributions of Biogeochemistry to Understanding Hominin Dietary Ecology. *Yearbook of Physical Anthropology*, *49*, 131–148. https://doi.org/10.1002/ajpa.20519

Lewis, J. (2014). Egalitarian Social Organization: The case of the Mbendjele BaYaka. In B. Hewlett (Ed.), *Hunter-Gatherers of the Congo Basin: Cultures, Histories and Biology of African Pygmies* (pp. 219–243). New Jersey, USA: Transaction Publishers.

Lieberman, D. (2008). Speculations about the selective basis for modern human craniofacial form. *Evolutionary Anthropology*, *17*(1), 55–68. https://doi.org/10.1002/evan.20154

Lieverse, A. (1999). Diet and the Aetiology of Dental Calculus. *International Journal of Osteoarchaeology*, *232*, 219–232.

Lightfoot, E., Przelomska, N., Craven, M., O'Connell, T., He, L., Hunt, H., & Jones, M. (2016). Intraspecific carbon and nitrogen isotopic variability in foxtail millet (Setaria italica). *Rapid Communications in Mass Spectrometry, 30*, 1475–1487. https://doi.org/10.1002/rcm.7583

Loftus, E., & Sealy, J. (2012). Technical note: interpreting stable carbon isotopes in human tooth enamel: an examination of tissue spacings from South Africa. *American Journal of Physical Anthropology, 147*(3), 499–507. https://doi.org/10.1002/ajpa.22012

Luca, F., Perry, G., & Di Rienzo, A. (2010). Evolutionary Adaptations to Dietary Changes. *Annual Review of Nutrition, 30*, 291–314. https://doi.org/10.1146/annurev-nutr-080508

MacArthur, J. (2012). The Making and Unmaking of African Languages: Oral Communities and Competitive Linguistic Work in Western Kenya. *Journal of African History, 53*, 151–172. https://doi.org/10.1017/s0021853712000229

MacArthur, J. (2013). When did the Luyia (or any other group) become a tribe? *Canadian Journal of African Studies, 47*(3), 351–363.

Macho, G., & Lee-Thorp, J. (2014). Niche partitioning in sympatric Gorilla and Pan from Cameroon: Implications for life history strategies and for reconstructing the evolution of hominin life history. *PLoS ONE*, *9*(7), e102794. https://doi.org/10.1371/journal.pone.0102794

Mair, P., & Wilcox, R. (2017). *Robust Statistical Methods in R Using the WRS2 Package*. Retrieved from https://cran.r-project.org/web/packages/WRS2/vignettes/WRS2.pdf

Makarewicz, C., & Sealy, J. (2015). Dietary reconstruction, mobility, and the analysis of ancient skeletal tissues: Expanding the prospects of stable isotope research in archaeology. *Journal of Archaeological Science, 56*, 146–158. https://doi.org/10.1016/j.jas.2015.02.035

Margvelashvili, A., Zollikofer, C., Lordkipanidze, D., Peltomäki, T., & Ponce de León, M. (2013). Tooth wear and dentoalveolar remodeling are key factors of morphological variation in the Dmanisi mandibles. P*roceedings of the National Academy of Sciences of the United States of America, 110*(43), 17278–17283.

McCabe, J. T. (1990). Success and Failure: The Breakdown of Traditional Drought Coping Institutions Among the Pastoral Turkana of Kenya. *Journal of African and Asian Studies, 25*(3), 146–160. https://doi.org/10.1163/156852190X00021

McCue, M., & Pollock, E. (2008). Stable isotopes may provide evidence for starvation in reptiles. *Rapid Communications in Mass Spectrometry, 22*, 2307–2314. https://doi.org/10.1002/rcm.3615 Stable

McCutchan, J., Lewis, W., Kendall, C., & McGrath, C. (2003). Variation in trophic shift for stable isotope ratios of carbon, nitrogen, and sulfur. *Oikos, 102*, 378–390. https://doi.org/10.1034/j.1600-0706.2003.12098.x

McKee, K., Feller, I., Popp, M., & Wanek, W. (2002). Mangrove isotopic ($\delta^{15}N$ and $\delta^{13}C$) fractionation across a nitrogen vs. phosphorus limitation gradient. *Ecology, 83*(4), 1065–1075.

McPherron, S., Alemseged, Z., Marean, C., Wynn, J., Reed, D., Geraads, D., Bobe, R., & Béarat, H. (2010). Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature, 466*(7308), 857–60. https://doi.org/10.1038/nature09248

Mirazón Lahr, M. (2012). The evolution of modern human diversity in Africa. Retrieved March 20, 2014, from www.in-africa.org/origins-2/after-the-origin/

Mirazón Lahr, M. (2013a). Genetic and fossil evidence for modern human origins. In P. Mitchell & P. J. Lane (Eds.), *The Oxford Handbook of African Archaeology* (First Edition, pp. 325–340). Oxford, UK: Oxford University Press, Inc.

Mirazón Lahr, M. (2013b). Modern human evolution in East Africa. Retrieved March 20, 2014, from www.inafrica.org/origins-2/east-african-perspectives/

Mirazón Lahr, M. (2016). The shaping of human diversity: filters, boundaries and transitions. *Philosophical Transactions of the Royal Society of London - Biological Sciences, 371*, 20150241. https://doi.org/10.1098/rstb.2015.0241

Mirazón Lahr, M., & Foley, R. (1994). Multiple Dispersals and Modern Human Origins. *Evolutionary Anthropology, 3*(2), 48–60.

Mirazón Lahr, M., & Foley, R. (2012). *Report on the 2012 Work of the Project "IN AFRICA: The role of East Africa in the evolution of human diversity."* Leverhulme Centre for Human Evolutionary Studies, University of Cambridge.

Mirazón Lahr, M., & Foley, R. (2014). *Report on the 2012-2014 work of the Project "IN AFRICA: The role of East Africa in the evolution of human diversity."* Leverhulme Centre for Human Evolutionary Studies, University of Cambridge.

Mirazón Lahr, M., & Foley, R. (2016). Human Evolution in Late Quaternary Eastern Africa. In S. Jones & B. Stewart (Eds.), *Africa from MIS 6-2: Population Dynamics and Paleoenvironments* (pp. 215–231). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-017-7520-5

Mirazón Lahr, M., Foley, R., & Mbua, E. (2009). *Report on 2011 field season - Project: "Later Quaternary archaeology and human palaeontology of the Kerio River and its environs."* Leverhulme Centre for Human Evolutionary Studies, University of Cambridge.

Mirazón Lahr, M., Rivera, F., Power, R., Mounier, A., Copsey, B., Crivellaro, F., Edung, J., Fernandez, J., Kiarie, C., Lawrence, J., Leakey, A., Mbua, E., Miller, H., Muigai, A., Mukhongo, D., Van Baelen, A., Wood, R., Schwenninger, J.-L., Grün, R., Achyuthan, H., Wilshaw, A., & Foley, R. (2016). Inter-group violence among early Holocene hunter-gatherers of West Turkana, Kenya. *Nature, 529*(7586), 394–398. https://doi.org/10.1038/nature16477

Miyawaki, Y. (2008). Purity and Heterogeneity: Warfare, Ethnic Identity, and Resistance Against the State Among the Hor of Southwestern Ethiopia. *Nilo-Ethiopian Studies, 12*, 61–74.

Mook, W., Bommerson, J., & Staverman, W. (1974). Carbon isotope fractionation between dissolved bicarbonate and gaseous carbon dioxide. *Earth and Planetary Science Letters, 22*(2), 169–176. https://doi.org/10.1016/0012-821X(74)90078-8

Muccio, Z., & Jackson, G. (2009). Isotope Ratio Mass Spectrometry. *Analyst, 134*, 213–222. https://doi.org/10.1039/b808232d

Murphy, S., Gewa, C., Grillenberger, M., & Neumann, C. (2004). Adapting an international food composition table for use in rural Kenya. *Journal of Food Composition and Analysis, 17*(3–4), 523–530. https://doi.org/10.1016/j.jfca.2004.03.019

Murphy, S., Weinberg-Andersson, S., Neumann, C., Mulligan, K., & Calloway, D. (1991). Development of Research Nutrient Data Bases: An Example Using Foods Consumed in Rural Kenya. *Journal of Food Composition and Analysis, 4*, 2–17.

Nash, S., Kristal, A., Boyer, B., King, I., Metzgar, J., & O'Brien, D. (2009). Relation between stable isotope ratios in human red blood cells and hair: implications for using the nitrogen isotope ratio of hair as a biomarker of eicosapentaenoic. *The American Journal of Clinical Nutrition*, *90*, 1642–1647. https://doi.org/10.3945/ajcn.2009.28482

Nash, S., Kristal, A., Hopkins, S., Boyer, B., & O'Brien, D. (2014). Stable Isotope Models of Sugar Intake Using Hair, Red Blood Cells, and Plasma, but Not Fasting Plasma Glucose, Predict Sugar Intake in a Yup'ik Study Population. *The Journal of Nutrition, 144*, 75–80. https://doi.org/10.3945/jn.113.182113

Newsome, S., Martinez del Rio, C., Bearhop, S., & Phillips, D. (2007). A niche for isotopic ecology. *Frontiers in Ecology and the Environment, 5*, 429–436. https://doi.org/10.1890/060150.01

Nicholson, E., & Harvati, K. (2006). Quantitative Analysis of Human Mandibular Shape Using Three-Dimensional Geometric Morphometrics. *American Journal of Physical Anthropology, 383*, 368–383. https://doi.org/10.1002/ajpa

Noback, M., & Harvati, K. (2015). Covariation in the Human Masticatory Apparatus. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology, 84*, 64–84. https://doi.org/10.1002/ar.23067

O'Brien, D. (2015). Stable Isotope Ratios as Biomarkers of Diet for Health Research. *Annual Review of Nutrition, 35*(1), 565–594. https://doi.org/10.1146/annurev-nutr-071714-034511

O'Connell, T., & Hedges, R. (1999a). Investigations into the effect of diet on modern human hair isotopic values. *American Journal of Physical Anthropology, 108*(4), 409–425.

O'Connell, T., & Hedges, R. (1999b). Isotopic Comparison of Hair and Bone: Archaeological Analyses. *Journal of Archaeological Science, 26*(6), 661–665. https://doi.org/10.1006/jasc.1998.0383

O'Connell, T., & Hedges, R. (2017). Chicken and Egg: Testing the Carbon Isotopic Effects of Carnivory and Herbivory. *Archaeometry, 59*(2), 302-315. https://doi.org/10.1111/arcm.12253

O'Connell, T., Hedges, R., Healey, M., & Simpson, A. (2001). Isotopic Comparison of Hair, Nail and Bone: Modern Analyses. *Journal of Archaeological Science, 28*, 1247–1255. https://doi.org/10.1006/jasc.2001.0698

O'Connell, T., Kneale, C., Tasevska, N., & Kuhnle, G. (2012). The diet-body offset in human nitrogen isotopic values: A controlled dietary study. *American Journal of Physical Anthropology, 149*(3), 426–434. https://doi.org/10.1002/ajpa.22140

O'Leary, M. (1988). Carbon Dynamics in Plants. *BioScience, 38*(5), 328–336.

Oelze, V., Fuller, B., Richards, M., Fruth, B., Surbeck, M., Hublin, J.-J., & Hohmann, G. (2011). Exploring the contribution and significance of animal protein in the diet of bonobos by stable isotope ratio analysis of hair. *Proceedings of the National Academy of Sciences of the United States of America, 108*(24), 9792–9797. https://doi.org/10.1073/pnas.1018502108

Ogot, B. (1967). *History of the Southern Luo*. Nairobi, Kenya: East African Publishing House.

Olang, M. (1984). Vegetation Cover Assessment in Turkana District, Kenya. In *Proceedings of the Workshop on Land Evaluation for Extensive Grazing* (pp. 183–194). Wageningen, ILRI.

Opiyo, F., Nyangito, M., Wasonga, O., & Omondi, P. (2014). Trend Analysis of Rainfall and Temperature Variability in Arid Environment of Turkana, Kenya. *Environmental Research Journal, 8*(2), 30–43.

Osmond, C., Vallane, N., Haslam, S., Uotila, P., & Roksandic, Z. (1981). Comparisons of $\delta^{13}C$ Values in Leaves of Aquatic Macrophytes from Different Habitats in Britain and Finland; some Implications for Photosynthetic Processes in Aquatic Plants. *Oecologia, 50*, 117–124.

Pampush, J., Duque, A., Burrows, B., Daegling, D., Kenney, W., & McGraw, W. S. (2013). Homoplasy and thick enamel in primates. *Journal of Human Evolution, 64*(3), 216–24. https://doi.org/10.1016/j.jhevol.2013.01.009

Pante, M., Blumenschine, R., Capaldo, S., & Scott, R. (2012). Validation of bone surface modification models for inferring fossil hominin and carnivore feeding interactions, with reapplication to FLK 22, Olduvai Gorge, Tanzania. *Journal of Human Evolution, 63*(2), 395–407. https://doi.org/10.1016/j.jhevol.2011.09.002

Passey, B., Robinson, T., Ayliffe, L., Cerling, T., Sponheimer, M., Dearing, M. D., Roeder, B., & Ehleringer, J. (2005). Carbon isotope fractionation between diet, breath

CO_2, and bioapatite in different mammals. *Journal of Archaeological Science, 32*(10), 1459–1470. https://doi.org/10.1016/j.jas.2005.03.015

Pate, F. D. (1994). Bone chemistry and paleodiet. *Journal of Archaeological Method and Theory, 1*(2), 161–209. https://doi.org/10.1007/BF02231415

Patel, P., Cooper, A., O'Connell, T., Kuhnle, G., Kneale, C., Mulligan, A., Luben, R., Brage, S., Khaw, K.-T., Wareham, N., & Forouhi, N. (2014). Serum carbon and nitrogen stable isotopes as potential biomarkers of dietary intake and their relation with incident type 2 diabetes: the EPIC-Norfolk study. *The American Journal of Clinical Nutrition, 100*(2), 708–718. https://doi.org/10.3945/ajcn.113.068577.Am

Patin, E., Barreiro, L., Sabeti, P., Austerlitz, F., Luca, F., Sajantila, A., Behar, D., Semino, O., Sakuntabhai, A., Guiso, N., Gicquel, B., McElreavey, K., Harding, R., Heyer, E., & Quintana-Murci, L. (2006). Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes. *American Journal of Human Genetics, 78*(3), 423–36. https://doi.org/10.1086/500614

Patin, E., Laval, G., Barreiro, L., Salas, A., Semino, O., Santachiara-Benerecetti, S., Kidd, K., Kidd, J., Van Der Veen, L., Hombert, J.-M., Gessain, A., Froment, A., Bahuchet, S., Heyer, E., & Quintana-Murci, L. (2009). Inferring the demographic history of African farmers and Pygmy hunter-gatherers using a multilocus resequencing data set. *PLoS Genetics, 5*(4). https://doi.org/10.1371/journal.pgen.1000448

Pearson, J., & Grove, M. (2013). Counting sheep: sample size and statistical inference in stable isotope analysis and palaeodietary reconstruction. *World Archaeology, 45*(3), 373–387. https://doi.org/10.1080/00438243.2013.820646

Perry, G., Dominy, N., Claw, K., Lee, A., Fiegler, H., Redon, R., Werner, J., Villanea, F., Mountain, J., Misra, R., Carter, N., Lee, C., & Stone, A. (2007). Diet and the evolution of human amylase gene copy number variation. *Nature, 39*(10), 1256–60. https://doi.org/10.1038/ng2123

Peters, C. (2007). Theoretical and Actualistic Ecobotanical Perspectives on Early Hominin Diets and Paleoecology. In P. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (pp. 233–261). Oxford, UK: Oxford University Press, Inc.

Phillips, D. (2012). Converting isotope values to diet composition: the use of mixing models. *Journal of Mammalogy, 93*(2), 342–352. https://doi.org/10.1644/11-MAMM-S-158.1

Poupin, N., Mariotti, F., Huneau, J.-F., Hermier, D., & Fouillet, H. (2014). Natural Isotopic Signatures of Variations in Body Nitrogen Fluxes: A Compartmental Model Analysis. *PLoS Computational Biology, 10*. https://doi.org/10.1371/journal.pcbi.1003865

Pryor, A., Stevens, R., O'Connell, T., & Lister, J. (2014). Quantification and propagation of errors when converting vertebrate biomineral oxygen isotope data to temperature for palaeoclimate reconstruction. *Palaeogeography, Palaeoclimatology, Palaeoecology, 412*, 99–107. https://doi.org/10.1016/j.palaeo.2014.07.003

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Raymond, P., & Bauer, J. (2001). Use of ^{14}C and ^{13}C natural abundances for evaluating riverine, estuarine, and coastal DOC and POC sources and cycling: A review and synthesis. *Organic Geochemistry, 32*(4), 469–485. https://doi.org/10.1016/S0146-6380(00)00190-X

Reddy, K., Patrick, W., & Lindau, C. (1989). Nitrification-denitrification at the plant root-sediment in interface wetlands. *Limnology and Oceanography, 34*(6), 1004–1013.

Reed, K., & Rector, A. (2007). African Pliocene Paleoecology - Hominin Habitats, Resources, Diets. In P. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (First Edition, pp. 262–288). Oxford, UK: Oxford University Press, Inc.

Reinhard, K., & Danielson, D. (2005). Pervasiveness of phytoliths in prehistoric southwestern diet and implications for regional and temporal trends for dental microwear. *Journal of Archaeological Science, 32*(7), 981–988. https://doi.org/10.1016/j.jas.2005.01.014

Reitsema, L. (2013). Beyond Diet Reconstruction: Stable Isotope Applications to Human Physiology, Health, and Nutrition. *American Journal of Human Biology, 25*(4), 445–56. https://doi.org/10.1002/ajhb.22398

Richards, M., & Hedges, R. (1999). Stable Isotope Evidence for Similarities in the Types of Marine Foods Used by Late Mesolithic Humans at Sites Along the Atlantic Coast of Europe. *Journal of Archaeological Science, 26*(6), 717–722. https://doi.org/10.1006/jasc.1998.0387

Robbins, C., Felicetti, L., & Sponheimer, M. (2005). The effect of dietary protein quality on nitrogen isotope discrimination in mammals and birds. *Oecologia, 144*(4), 534–540. https://doi.org/10.1007/s00442-005-0021-8

Robbins, L. (2006). Lake Turkana archaeology: The Holocene. *Ethnohistory, 1*, 71–93. Retrieved from http://ethnohistory.dukejournals.org/content/53/1/71.short

Roberts, P., Perera, N., Wedage, O., Deraniyagala, S., Perera, J., Eregama, S., Gledhill, A., Petraglia, M., & Lee-Thorp, J. (2015). Direct evidence for human reliance on rainforest resources in late Pleistocene Sri Lanka. *Science*, *347*(6227), 1246–1249. https://doi.org/10.1126/science.aaa1230

Robinson, D. (2001). $\delta^{15}N$ as an integrator of the nitrogen cycle. *Trends in Ecology & Evolution, 16*(3), 153–162. https://doi.org/10.1016/s0169-5347(00)02098-x

Romero, A., Galbany, J., De Juan, J., & Pérez-Pérez, A. (2012). Brief communication: Short- and long-term in vivo human buccal-dental microwear turnover. A*merican Journal of Physical Anthropology, 148*, 467–472. https://doi.org/10.1002/ajpa.22054

Romero, A., Galbany, J., Perez-Perez, A., & De Juan, J. (2007). Microwear Formation Rates in Human Buccal Tooth Enamel Surfaces: An Experimental *in vivo* Analysis under Induced-Diet. In É. Bodzsár & A. Zsákai (Eds.), *New Perspectives and Problems in Anthropology* (pp. 135–143). Newcastle, UK: Cambridge Scholars Publishing.

Romero, A., Ramírez-Rozzi, F., De Juan, J., & Pérez-Pérez, A. (2013). Diet-related buccal dental microwear patterns in Central African Pygmy foragers and Bantu-speaking farmer and pastoralist populations. *PloS ONE, 8*(12), e84804. https://doi.org/10.1371/journal.pone.0084804

Roseman, C., & Weaver, T. (2004). Multivariate apportionment of global human craniometric diversity. *American Journal of Physical Anthropology, 125*(3), 257–263. https://doi.org/10.1002/ajpa.10424

Saitoh, M., Uzuka, M., Sakamoto, M., & Kobori, T. (1969). Rate of Hair Growth. In W. Montagna & R. Dobson (Eds.), *Hair Growth* (pp. 183–201). Oxford, UK: Pergamon Press.

Sandford, M. (1993). Understanding the Biogenic-Diagenetic Continuum: Interpreting Elemental Concentrations of Archaeological Bone. In M. Sandford (Ed.), *Investigations of Ancient Human Tissue - Chemical Analyses in Anthropology* (pp. 3–57). Langhorne, Pennsylvania, US: Gordon and Breach Science Publishers.

Sanson, G., Kerr, S., & Gross, K. (2007). Do silica phytoliths really wear mammalian teeth? *Journal of Archaeological Science, 34*(4), 526–531. https://doi.org/10.1016/j.jas.2006.06.009

Sato, H. (2014). Foraging Lifestyle in the African Tropical Rainforest. In B. Hewlett (Ed.), *Hunter-Gatherers of the Congo Basin: Cultures, Histories and Biology of African Pygmies* (pp. 165–185). New Jersey, USA: Transaction Publishers.

Sato, H., Kawamura, K., Hayashi, K., Inai, H., & Yamauchi, T. (2012). Addressing the wild yam question: how Baka hunter-gatherers acted and lived during two controlled foraging trips in the tropical rainforest of southeastern Cameroon. *Anthropological Science, 120*(2), 129–149. https://doi.org/10.1537/ase.110913

Scherrer, J. (1978). *Fisherfold of the Desert: An Ethnography of the ElMolo of Kenya*. University of Virginia.

Schlebusch, C., Malmström, H., Jakobsson, M., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A., Vicente, M., Steyn, M., Soodyall, H., Lombard, M., & Jakobsson, M. (2017). Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science*. https://doi.org/10.1126/science.aao6266

Schmidt, C. (2010). On the relationship of dental microwear to dental macrowear. *American Journal of Physical Anthropology*, *142*(1), 67–73. https://doi.org/10.1002/ajpa.21197

Schmidt, C., Beach, J., McKinley, J., & Eng, J. (2016). Distinguishing dietary indicators of pastoralists and agriculturists via dental microwear texture analysis. S*urface Topography: Metrology and Properties, 4*. https://doi.org/10.1088/2051-672X/4/1/014008

Schoeller, D. (1999). Isotope fractionation: why aren't we what we eat? *Journal of Archaeological Science*, *26*(6), 667–673. https://doi.org/10.1006/jasc.1998.0391

Schoeller, D., Brown, C., Nakamura, K., Nakagawa, A., Mazzeo, R., Brooks, A., & Budinger, T. (1984). Influence of Metabolic Fuel on the $^{13}C/^{12}C$ Ratio of Breath CO_2. *Biomedical Mass Spectrometry*, *11*(11), 557–561.

Schoeninger, M. (1995). Stable isotope studies in human evolution. *Evolutionary Anthropology*, *4*(3), 83–98. https://doi.org/10.1002/evan.1360040305

Schoeninger, M. (2010). Diet Reconstruction and Ecology Using Stable Isotope Ratios. In *A Companion to Biological Anthropology* (pp. 445–464).

Schoeninger, M., & DeNiro, M. (1984). Nitrogen and carbon isotopic composition of bone collagen from marine and terrestrial animals. *Geochimica et Cosmochimica Acta*, *48*(4), 625–639. https://doi.org/10.1016/0016-7037(84)90091-7

Scott, G. R., & Poulson, S. (2012). Stable carbon and nitrogen isotopes of human dental calculus: a potentially new non-destructive proxy for paleodietary analysis. *Journal of Archaeological Science, 39*(5), 1388–1393. https://doi.org/10.1016/j.jas.2011.09.029

Scott, R., Teaford, M., & Ungar, P. (2012). Dental microwear texture and anthropoid diets. *American Journal of Physical Anthropology, 147*(4), 551–79. https://doi.org/10.1002/ajpa.22007

Scott, R., Ungar, P., Bergstrom, T., Brown, C., Childs, B., Teaford, M., & Walker, A. (2006). Dental microwear texture analysis: technical considerations. *Journal of Human Evolution, 51*(4), 339–49. https://doi.org/10.1016/j.jhevol.2006.04.006

Scott, R., Ungar, P., Bergstrom, T., Brown, C., Grine, F., Teaford, M., & Walker, A. (2005). Dental microwear texture analysis shows within-species diet variability in fossil hominins. *Nature, 436*(7051), 693–5. https://doi.org/10.1038/nature03822

Sealy, J. (2001). Body Tissue Chemistry and Palaeodiet. In D. Brothwell & A. Pollard (Eds.), *Handbook of Archaeological Sciences* (pp. 267–279). New York, USA: John Wiley & Sons, Inc.

Sealy, J. (2006). Diet, Mobility, and Settlement Pattern among Holocene Hunter-Gatherers in Southernmost Africa. *Current Anthropology*, *47*(4), 569–595. https://doi.org/10.1086/504163

Sealy, J. (2010). Isotopic evidence for the antiquity of cattle-based pastoralism in southernmost Africa. *Journal of African Archaeology*, *8*(1), 65–81. https://doi.org/10.3213/1612-1651-10160

Semaw, S. (2000). The World's Oldest Stone Artefacts from Gona, Ethiopia: Their Implications for Understanding Stone Technology and Patterns of Human Evolution Between 2·6–1·5 Million Years Ago. *Journal of Archaeological Science, 27*(12), 1197–1214. https://doi.org/10.1006/jasc.1999.0592

Sept, J. (2007). Modeling the Significance of Paleoenvironmental Context for Early Hominin Diets. In P. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (First Edition, pp. 289–307). Oxford, UK: Oxford University Press, Inc.

Sharp, Z. (2007). *Principles of Stable Isotope Geochemistry*. New Jersey, USA: Pearson Education Press.

Sillen, A., Sealy, J., & van der Merwe, N. (1989). Chemistry and Paleodietary Research: No More Easy Answers. *American Antiquity, 54*(3), 504–512.

Slater, C., Preston, T., & Weaver, L. T. (2004). Is there an advantage in normalising the results of the Helicobacter pylori ^{13}C urea breath test for CO_2 production rate in children? *Isotopes in Environmental and Health Studies, 40*(1), 89–98. https://doi.org/10.1080/10256010310001621164

Smith, B., & Epstein, S. (1971). Two categories of $^{13}C/^{12}C$ ratios for higher plants. *Plant Physiology, 47*(3), 380–384. https://doi.org/10.1104/pp.47.3.380

Smith, H. (2009). Which cranial regions reflect molecular distances reliably in humans? Evidence from three-dimensional morphology. *American Journal of Human Biology, 21*(1), 36–47. https://doi.org/10.1002/ajhb.20805

Smith, T., Olejniczak, A., Zermeno, J., Tafforeau, P., Skinner, M., Hoffmann, A., Radovčić, J., Toussaint, M., Kruszynski, R., Menter, C., Moggi-Cecchi, J., Glasmacher, U., Kullmer, O., Schrenk, F., Stringer, C., & Hublin, J.-J. (2012). Variation in enamel thickness within the genus Homo. *Journal of Human Evolution, 62*(3), 395–411. https://doi.org/10.1016/j.jhevol.2011.12.004

Sobania, N. (1978). The Problem of Origins: Liguistics Hypotheses and Oral tradition. *Abbay, 9*, 87–100.

Sponheimer, M., & Dufour, D. (2009). Increased Dietary Breadth in Early Hominin Evolution: Revisiting Arguments and Evidence with a Focus on Biogeochemical Contributions. In J.-J. Hublin & M. P. Richards (Eds.), *The Evolution of Hominin Diets - Integrating Approaches to the Study of Palaeolithic Subsistence* (First Edition, pp. 229–240). Dordrecht, Netherlands: Springer.

Sponheimer, M., & Lee-Thorp, J. (2006). Enamel diagenesis at South African Australopith sites: Implications for paleoecological reconstruction with trace elements. *Geochimica et Cosmochimica Acta, 70*(7), 1644–1654. https://doi.org/10.1016/j.gca.2005.12.022

Steele, T., & Klein, R. (2009). Late Pleistocene Subsistence Strategies and Resource Intensification in Africa. In J. Hublin & M. Richards (Eds.), *The Evolution of Hominin Diets - Integrating Approaches to the Study of Palaeolithic Subsistence* (First Edition, pp. 113–126). Dordrecht, Netherlands: Springer.

Stenhouse, M., & Baxter, M. (1979). The uptake of bomb ^{14}C in humans. In R. Berger & H. Suess (Eds.), *Radiocarbon Dating* (pp. 324–341). Berkeley, California: University of California Press, Ltd.

Teaford, M. (2007a). What Do We Know and Not Know about Dental Microwear and Diet? In P. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (First Edition, pp. 106–131). Oxford, UK: Oxford University Press, Inc.

Teaford, M. (2007b). What Do We Know and Not Know about Diet and Enamel Structure? In P. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (First Edition, pp. 56–76). Oxford, UK: Oxford University Press, Inc.

Teaford, M., & Glander, K. (1991). Dental microwear in live, wild-trapped Alouatta palliata from Costa Rica. *American Journal of Physical Anthropology, 85*(3), 313–319. https://doi.org/10.1002/ajpa.1330850310

Teaford, M., & Lytle, J. (1996). Brief communication: Diet-induced changes in rates of human tooth microwear: A case study involving stone-ground maize. *American Journal of Physical Anthropology, 100*(1), 143–147.

Teaford, M., & Oyen, O. (1989a). *In vivo* and *in vitro* turnover in dental microwear. *American Journal of Physical Anthropology*, *80*, 447–460. https://doi.org/10.1002/ajpa.1330800405

Teaford, M., & Oyen, O. (1989b). Live Primates and Dental Replication : New Problems and New Techniques. *American Journal of Physical Anthropology*, *80*, 73–81.

Teaford, M., & Tylenda, C. (1991). A new approach to the study of tooth wear. *Journal of Dental Research, 70*, 204–207. https://doi.org/10.1177/00220345910700030901

Thompson, L., Mosley-Thompson, E., Davis, M., Henderson, K., Brecher, H., Zagorodnov, V., Mashiotta, T., Lin, P.-N., Mikhalenko, V., Hardy, D., & Beer, J. (2002). Kilimanjaro Ice Core Records: Evidence of Holocene Climate Change in Tropical Africa. *Science*, *298*, 589–593. https://doi.org/10.1126/science.1073198

Tieszen, L., & Fagre, T. (1993). Effect of Diet Quality and Composition on the Isotopic Composition of Respiratory CO_2, Bone Collagen, Bioapatite, and Soft Tissues. In J. Lambert & G. Grupe (Eds.), *Prehistoric Human Bone - Archaeology at the Molecular Leve*l (pp. 121–155). Berlin, Germany: Springer-Verlag.

Timm, O., Köhler, P., Timmermann, A., & Menviel, L. (2010). Mechanisms for the Onset of the African Humid Period and Sahara Greening 14.5–11 ka BP*. *Journal of Climate*, *23*, 2612–33.

Tipple, B., Chau, T., Chesson, L., Fernandez, D., & Ehleringer, J. (2013). Isolation of strontium pools and isotope ratios in modern human hair. *Analytica Chimica Acta, 798*, 64–73. https://doi.org/10.1016/j.aca.2013.08.054

Tishkoff, S., Gonder, M., Henn, B., Mortensen, H., Knight, A., Gignoux, C., Fernandopulle, N., Lema, G., Nyambo, T., Ramakrishnan, U., Reed, F., & Mountain, J. (2007). History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Molecular Biology and Evolution, 24*(10), 2180–2195. https://doi.org/10.1093/molbev/msm155

Tishkoff, S., Reed, F., Friedlaender, F., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J., Awomoyi, A., Bodo, J.-M., Doumbo, O., Ibrahim, M., Juma, A., Kotze, M., Lema, G., Moore, J., Mortensen, H., Nyambo, T., Omar, S., Powell, K., Pretorius, G., Smith, M., Thera, M., Wambebe, C., Weber, J., & Williams, S. (2009). The genetic structure and history of Africans and African Americans. *Science, 324*(5930), 1035–44. https://doi.org/10.1126/science.1172257

Tishkoff, S., Reed, F., Ranciaro, A., Voight, B., Babbitt, C., Silverman, J., Powell, K., Mortensen, H., Hirbo, J., Osman, M., Ibrahim, M., Omar, S., Lema, G., Nyambo, T., Ghori, J., Bumpstead, S., Pritchard, J., Wray, G., & Deloukas, P. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature, 39*(1), 31–40. https://doi.org/10.1038/ng1946

Todorov, V., & Filzmoser, P. (2009). An Object Oriented Framework for Robust Multivariate Analysis. *Journal of Statistical Software, 32*, 1–47. https://doi.org/http://dx.doi.org/10.18637/jss.v032.i03

Trotter, M., Hixon, B., & MacDonald, B. (1977). Development and size of the teeth of Macaca mulatta. *American Journal of Anatomy, 150*(1), 109–127. https://doi.org/10.1002/aja.1001500108

Tsutaya, T., & Yoneda, M. (2015). Reconstruction of Breastfeeding and Weaning Practices Using Stable Isotope and Trace Element Analyses: A Review. *Yearbook of Physical Anthropology, 156*, 2–21. https://doi.org/10.1002/ajpa.22657

Tuross, N. (2017, April). The Impact of Caloric Restriction on Tissue Isotopic ($\delta^{15}N$, $\delta^{13}C$ and $\delta^{18}O$) Values. Poster presented at the Meeting of the American Association of Physical Anthropologists, New Orleans, USA.

Ueno, O., Samejima, M., Muto, S., & Miyachi, S. (1988). Photosynthetic characteristics of an amphibious plant, Eleocharis vivipara: Expression of C_4 and C_3 modes in contrasting environments. *Proceedings of the National Academy of Sciences of the United States of America, 85*(18), 6733–6737. https://doi.org/10.1073/pnas.85.18.6733

Ungar, P. (2004). Dental topography and diets of Australopithecus afarensis and early Homo. *Journal of Human Evolution*, *46*(5), 605–22. https://doi.org/10.1016/j.jhevol.2004.03.004

Ungar, P. (2007). Limits to Knowledge on the Evolution of Hominin Diet. In P. S. Ungar (Ed.), *Evolution of the Human Diet - the Known, the Unknown and the Unknowable* (First Edition, pp. 395–407). Oxford, UK: Oxford University Press, Inc.

Ungar, P. (2012). Dental Evidence for the Reconstruction of Diet in African Early Homo. *Current Anthropology*, *53*(S6), S318–S329. https://doi.org/10.1086/666700

Ungar, P. (2014). Dental Allometry in Mammals: A Retrospective. *Annales Zoologici Fennici*, *51*, 177–187. https://doi.org/10.5735/086.051.0218

Ungar, P., & Lucas, P. (2010). Tooth Form and Function in Biological Anthropology. In C. S. Larsen (Ed.), *A Companion to Biological Anthropology* (pp. 516–529). Oxford, UK: Blackwell Publishing Ltd.

Ungar, P., Scott, R., Scott, J., & Teaford, M. (2008). Dental microwear analysis: historical perspectives and new approaches. In J. D. Irish & G. C. Nelson (Eds.), *Technique and Application in Dental Anthropology* (pp. 389–425). Cambridge, UK: Cambridge University Press.

Ungar, P., & Sponheimer, M. (2011). The diets of early hominins. *Science*, *334*(6053), 190–3. https://doi.org/10.1126/science.1207701

Ungar, P., & Sponheimer, M. (2013). Hominin diets. In D. R. Begun (Ed.), *A Companion to Paleoanthropology* (pp. 165–182). Chichester, UK: Blackwell Publishing Ltd.

Ungar, P., Teaford, M., Glander, K., & Pastor, R. (1995). Dust accumulation in the canopy: A potential cause of dental microwear in primates. *American Journal of Physical Anthropology*, *97*(2), 93–99. https://doi.org/10.1002/ajpa.1330970202

USDA. (2015). *USDA National Nutrient Database for Standard Reference, Release 28*. United States Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory. Retrieved from https://ndb.nal.usda.gov/ndb/search/list

Valenzuela, L., Chesson, L., Bowen, G., Cerling, T., & Ehleringer, J. (2012). Dietary Heterogeneity among Western Industrialized Countries Reflected in the Stable Isotope Ratios of Human Hair. *PLoS ONE, 7*(3), e34234. https://doi.org/10.1371/journal.pone.0034234

Vallerini, L., & Ceccanti, G. (2011). Oasis of Loiyangalani Kenya Combat Desertification - Water Resources Management. *AQUA Mundi, Am03025*, 15–26. https://doi.org/10.4409/Am-025-11-0025

Verdu, P. (2014). Population Genetics of Central African Pygmies and Non-Pygmies. In B. Hewlett (Ed.), *Hunter-Gatherers of the Congo Basin: Cultures, Histories and Biology of African Pygmies* (pp. 31–58). New Jersey, USA: Transaction Publishers.

Vogel, J., & van der Merwe, N. (1977). Isotopic Evidence for Early Maize Cultivation in New York State. *American Antiquity, 42*(2), 238–242.

von Cramon-Taubadel, N. (2009). Congruence of individual cranial bone morphology and neutral molecular affinity patterns in modern humans. *American Journal of Physical Anthropology, 140*, 205–215. https://doi.org/10.1002/ajpa.21041

von Cramon-Taubadel, N. (2011). Global human mandibular variation reflects differences in agricultural and hunter-gatherer subsistence strategies. *Proceedings of the National Academy of Sciences of the United States of America, 108*(49), 19546–51. https://doi.org/10.1073/pnas.1113050108

von Cramon-Taubadel, N. (2014a). Evolutionary insights into global patterns of human cranial diversity: population history, climatic and dietary effects. *Journal of Anthropological Sciences, 92*, 43–77. https://doi.org/10.4436/jass.91010

von Cramon-Taubadel, N. (2014b). The microevolution of modern human cranial variation: implications for hominin and primate evolution. *Annals of Human Biology, 41*(4), 323–35. https://doi.org/10.3109/03014460.2014.911350

von Höhnel, L. (1894). *Discovery of Lakes Rudolf and Stephanie: A Narrative of Count Samuel Teleki's Exploring and Hunting Expeditions in Eastern Equatorial Africa in 1887 and 1888*. London, United Kingdom: Longman, Green and Co.

von Vossen, R. (1982). *The Eastern Nilotes: Linguistic and Historical Reconstructions*. Berlin, Germany: Dietrich Reimer Verlag.

Wada, E., Kadonaga, T., & Matsuo, S. (1975). ^{15}N abundance in nitrogen of naturally occurring substances and global assessment of denitrification from isotopic viewpoint. *Geochemical Journal, 9*(3), 139–148. https://doi.org/10.2343/geochemj.9.139

Wagah, M., Bader, E., Deligia, C., & Dop, M. (2005). Kenya Nutrition Profile. *Food and Nutrition Division, FAO*.

Wagner, G. (1949). *The Bantu of Western Kenya*. Oxford, UK: Oxford University Press, Inc.

Waterlow, J. (1986). Metabolic Adaptation to Low Intakes of Energy and Protein. *Annual Review of Nutrition, 6*, 495–526.

Were, G. (1967). *A History of the Abaluyia of Western Kenya*. Nairobi, Kenya: East African Publishing House.

Whigham, L., Butz, D., Johnson, L., Schoeller, D., Abbott, D., Porter, W., & Cook, M. (2014). Breath carbon stable isotope ratios identify changes in energy balance and substrate utilization in humans. *International Journal of Obesity, 38*, 1–3. https://doi.org/10.1038/ijo.2014.7

White, T., Asfaw, B., Beyene, Y., Haile-Selassie, Y., Lovejoy, C. O., Suwa, G., & WoldeGabriel, G. (2009). *Ardipithecus ramidus* and the Paleobiology of Early Hominids. *Science, 326*(5949), 64–64, 75–86. https://doi.org/10.1126/science.1175802

WHO/FAO. (2003). *Diet, nutrition and the prevention of chronic diseases - Report of a Joint WHO/FAO Expert Consultation*. Geneva, Switzerland. https://doi.org/ISBN 92 4 120916 X ISSN 0512-3054 (NLM classification: QU 145)

WHO/FAO/UNU. (2004). *Human energy requirements: Report of a Joint FAO/WHO/UNU Expert Consultation*. Rome, Italy. https://doi.org/92 5 105212 3

WHO/FAO/UNU. (2007). *Protein and amino acid requirements in human nutrition - Report of a Joint WHO/FAO/UNU Expert Consultation*. Geneva, Switzerland. https://doi.org/ISBN 92 4 120935 6

Wilcox, R. (2012). *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press, Elsevier Inc.

Wood, B., & Schroer, K. (2012). Reconstructing the Diet of an Extinct Hominin Taxon: The Role of Extant Primate Models. *International Journal of Primatology, 33*(3), 716–742. https://doi.org/10.1007/s10764-012-9602-7

Woodward, A. (1921). A New Cave Man from Rhodesia, South Africa. *Nature, 108*(2716), 371–372.

Wynn, T., & Coolidge, F. (2010). How Levallois Reduction is Similar to, and Not Similar to, Playing Chess. In A. Nowell & I. Davidson (Eds.), *Stone Tools and the Evolution of Human Cognition* (pp. 83–103). Boulder: University Press of Colorado.

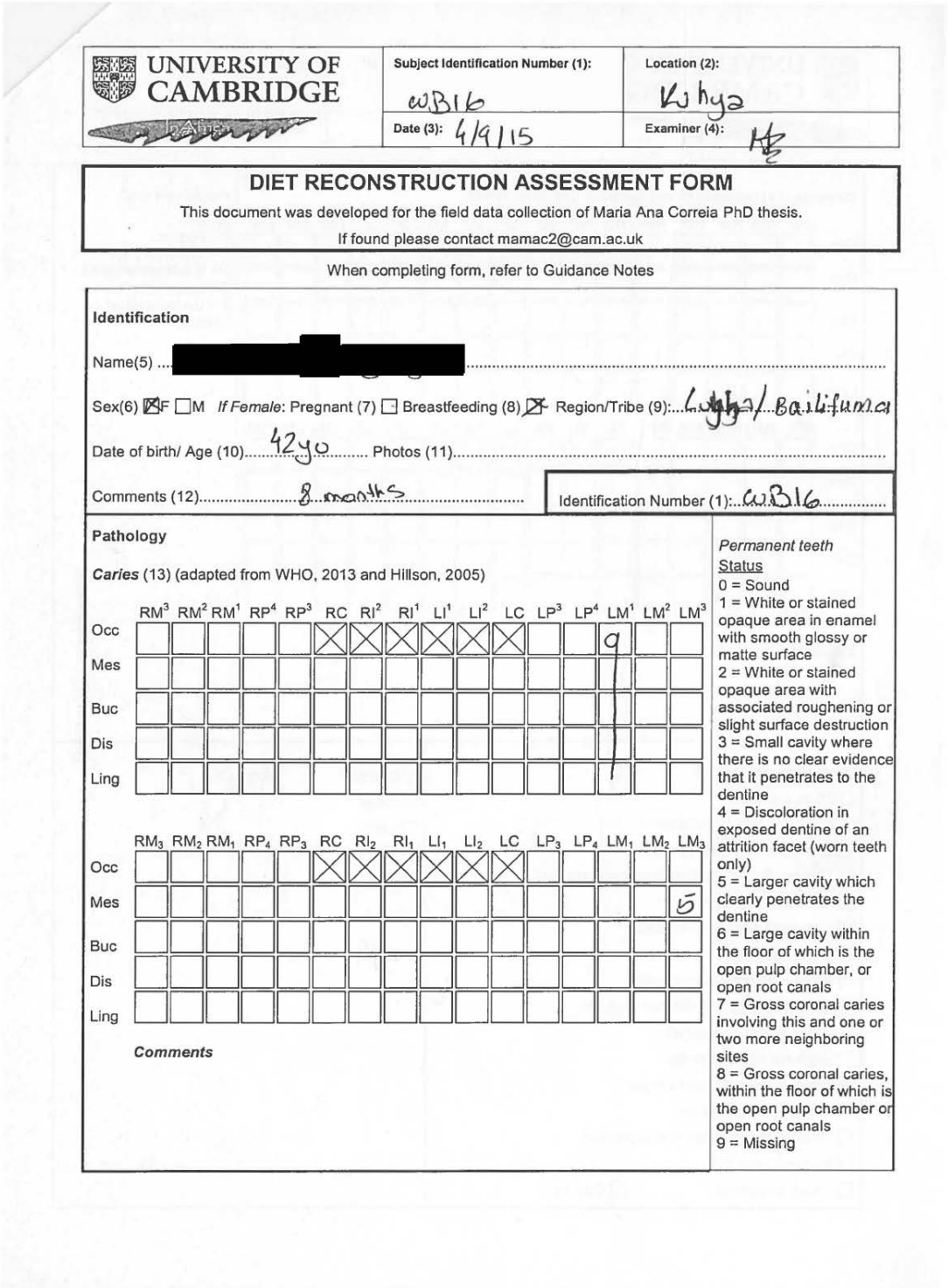
Yamauchi, T., Sato, H., & Kawamura, K. (2000). Nutritional Status, Activity Pattern, and Dietary Intake among the Baka Hunter-gatherers in the Village Camps in Cameroon. *African Study Monographs, 21*(2), 67–82. Retrieved from http://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/68193

Zaias, N. (1980). *The Nail in Health and Disease*. Norwalk, Connecticut: Appleton & Lange.

APPENDIX A

MATERIALS & METHODS

A.1 DATA COLLECTION



UNIVERSITY OF CAMBRIDGE

Subject Identification Number (1):	Location (2):
cuB16	Kuhya
Date (3): 4/9/15	Examiner (4):

DIET RECONSTRUCTION ASSESSMENT FORM

This document was developed for the field data collection of Maria Ana Correia PhD thesis.

If found please contact mamac2@cam.ac.uk

When completing form, refer to Guidance Notes

Identification

Name(5) ....[redacted]....

Sex(6) ☒F ☐M *If Female*: Pregnant (7) ☐ Breastfeeding (8) ☒ Region/Tribe (9): Luhya/Bailituma

Date of birth/ Age (10) 42yo Photos (11)....

Comments (12) 8 months Identification Number (1): cuB16

Pathology

Caries (13) (adapted from WHO, 2013 and Hillson, 2005)

	RM^3	RM^2	RM^1	RP^4	RP^3	RC	RI^2	RI^1	LI^1	LI^2	LC	LP^3	LP^4	LM^1	LM^2	LM^3
Occ						X	X	X	X	X	X			9		
Mes																
Buc																
Dis																
Ling																

	RM_3	RM_2	RM_1	RP_4	RP_3	RC	RI_2	RI_1	LI_1	LI_2	LC	LP_3	LP_4	LM_1	LM_2	LM_3
Occ						X	X	X	X	X	X					
Mes																5
Buc																
Dis																
Ling																

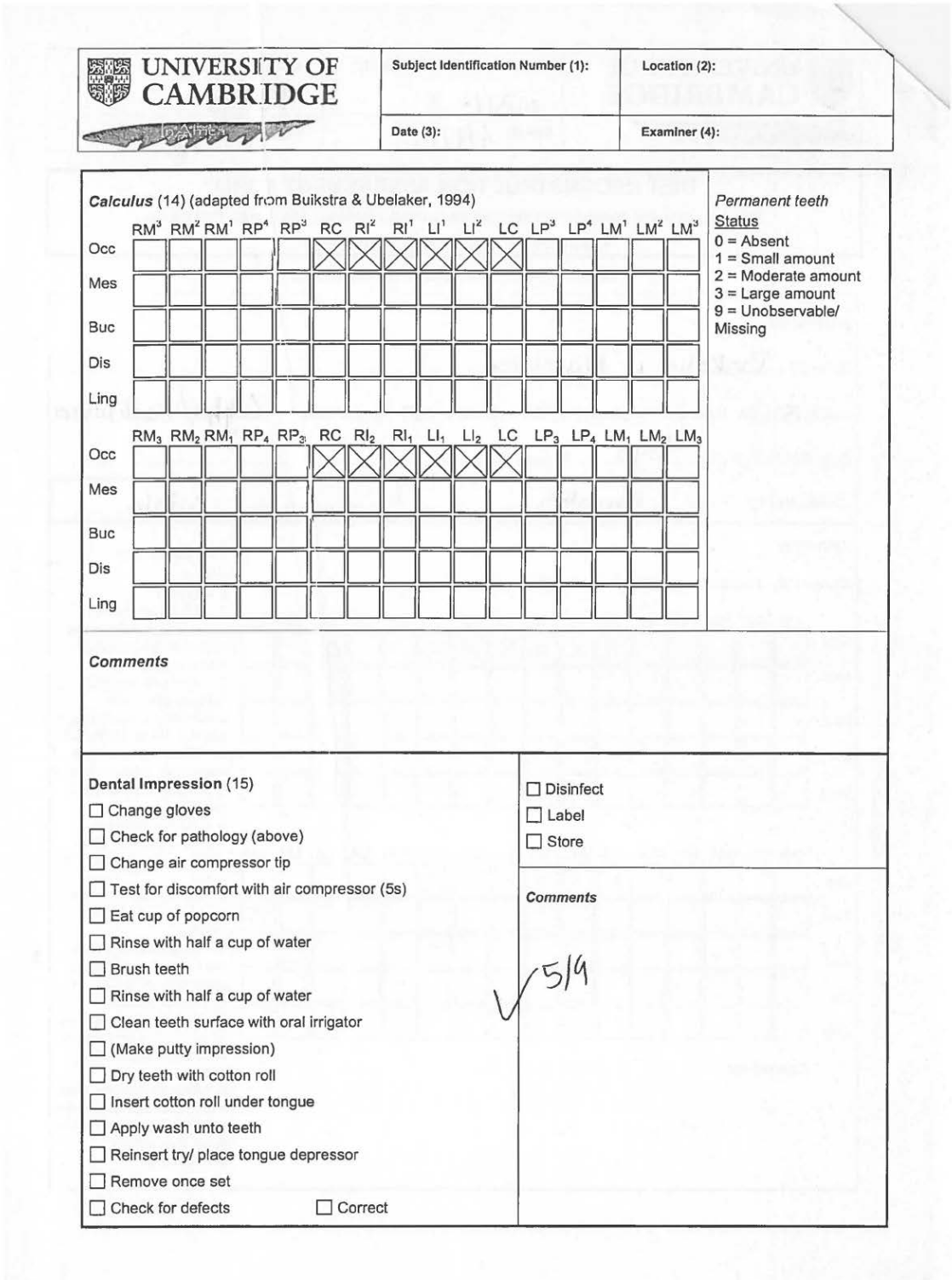
Comments

Permanent teeth

Status

0 = Sound
1 = White or stained opaque area in enamel with smooth glossy or matte surface
2 = White or stained opaque area with associated roughening or slight surface destruction
3 = Small cavity where there is no clear evidence that it penetrates to the dentine
4 = Discoloration in exposed dentine of an attrition facet (worn teeth only)
5 = Larger cavity which clearly penetrates the dentine
6 = Large cavity within the floor of which is the open pulp chamber, or open root canals
7 = Gross coronal caries involving this and one or two more neighboring sites
8 = Gross coronal caries, within the floor of which is the open pulp chamber or open root canals
9 = Missing

Figure A.1 Example of a data collection form (page 1/4).



UNIVERSITY OF CAMBRIDGE

Subject Identification Number (1):	Location (2):
Date (3):	Examiner (4):

Calculus (14) (adapted from Buikstra & Ubelaker, 1994)

	RM³	RM²	RM¹	RP⁴	RP³	RC	RI²	RI¹	LI¹	LI²	LC	LP³	LP⁴	LM¹	LM²	LM³
Occ						X	X	X	X	X	X					
Mes																
Buc																
Dis																
Ling																

	RM₃	RM₂	RM₁	RP₄	RP₃	RC	RI₂	RI₁	LI₁	LI₂	LC	LP₃	LP₄	LM₁	LM₂	LM₃
Occ						X	X	X	X	X	X					
Mes																
Buc																
Dis																
Ling																

Permanent teeth
Status
0 = Absent
1 = Small amount
2 = Moderate amount
3 = Large amount
9 = Unobservable/ Missing

Comments

Dental Impression (15)

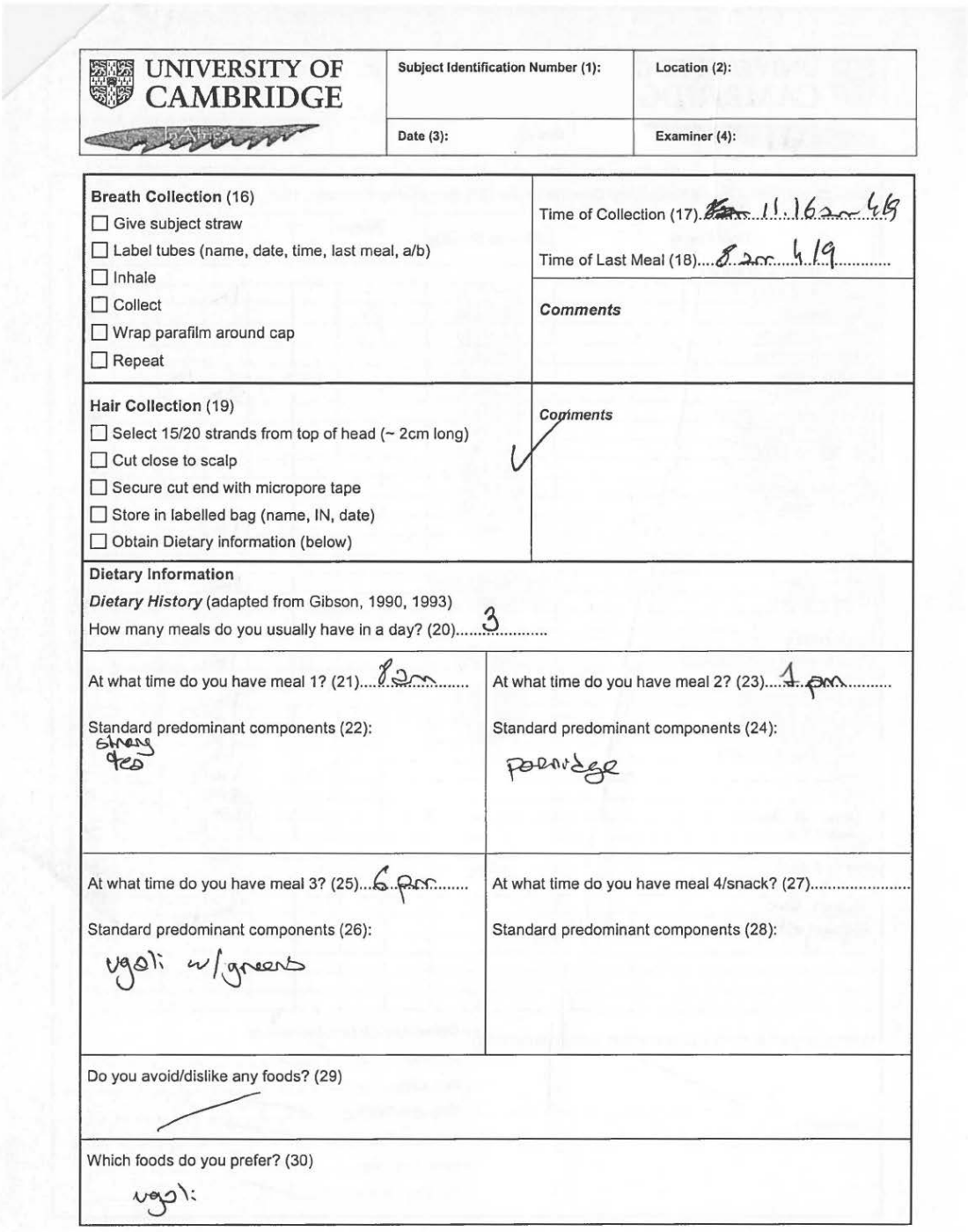
- ☐ Change gloves
- ☐ Check for pathology (above)
- ☐ Change air compressor tip
- ☐ Test for discomfort with air compressor (5s)
- ☐ Eat cup of popcorn
- ☐ Rinse with half a cup of water
- ☐ Brush teeth
- ☐ Rinse with half a cup of water
- ☐ Clean teeth surface with oral irrigator
- ☐ (Make putty impression)
- ☐ Dry teeth with cotton roll
- ☐ Insert cotton roll under tongue
- ☐ Apply wash unto teeth
- ☐ Reinsert try/ place tongue depressor
- ☐ Remove once set
- ☐ Check for defects ☐ Correct

- ☐ Disinfect
- ☐ Label
- ☐ Store

Comments

✓5/9

Figure A.2 Example of a data collection form (page 2/4).



UNIVERSITY OF CAMBRIDGE

Subject Identification Number (1):	Location (2):
Date (3):	Examiner (4):

Breath Collection (16)
- ☐ Give subject straw
- ☐ Label tubes (name, date, time, last meal, a/b)
- ☐ Inhale
- ☐ Collect
- ☐ Wrap parafilm around cap
- ☐ Repeat

Time of Collection (17) 11.16am 4/9

Time of Last Meal (18) 8am 4/9

Comments

Hair Collection (19)
- ☐ Select 15/20 strands from top of head (~ 2cm long)
- ☐ Cut close to scalp
- ☐ Secure cut end with micropore tape
- ☐ Store in labelled bag (name, IN, date)
- ☐ Obtain Dietary information (below)

Comments ✓

Dietary Information

Dietary History (adapted from Gibson, 1990, 1993)

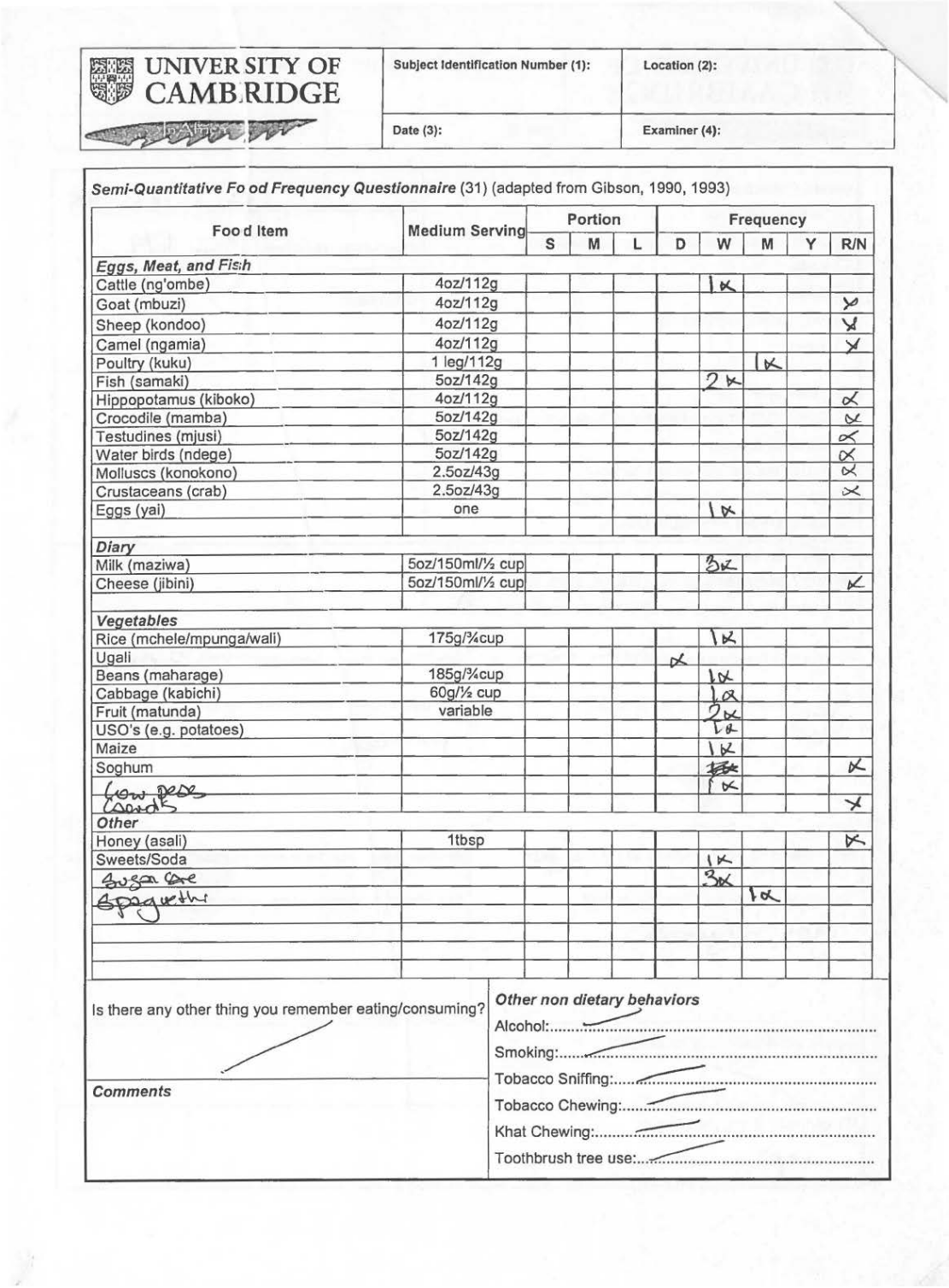
How many meals do you usually have in a day? (20) 3

At what time do you have meal 1? (21) 8am	At what time do you have meal 2? (23) 1pm
Standard predominant components (22): strong tea	Standard predominant components (24): porridge
At what time do you have meal 3? (25) 6pm	At what time do you have meal 4/snack? (27)
Standard predominant components (26): vgbli w/greens	Standard predominant components (28):

Do you avoid/dislike any foods? (29)

Which foods do you prefer? (30) vgbl:

Figure A.3 Example of a data collection form (page 3/4).



UNIVERSITY OF CAMBRIDGE

Subject Identification Number (1):	Location (2):
Date (3):	Examiner (4):

Semi-Quantitative Food Frequency Questionnaire (31) (adapted from Gibson, 1990, 1993)

Food Item	Medium Serving	Portion			Frequency				
		S	M	L	D	W	M	Y	R/N
Eggs, Meat, and Fish									
Cattle (ng'ombe)	4oz/112g					1x			
Goat (mbuzi)	4oz/112g								x
Sheep (kondoo)	4oz/112g								x
Camel (ngamia)	4oz/112g								x
Poultry (kuku)	1 leg/112g						1x		
Fish (samaki)	5oz/142g					2x			
Hippopotamus (kiboko)	4oz/112g								x
Crocodile (mamba)	5oz/142g								x
Testudines (mjusi)	5oz/142g								x
Water birds (ndege)	5oz/142g								x
Molluscs (konokono)	2.5oz/43g								x
Crustaceans (crab)	2.5oz/43g								x
Eggs (yai)	one					1x			
Diary									
Milk (maziwa)	5oz/150ml/½ cup					3x			
Cheese (jibini)	5oz/150ml/½ cup								x
Vegetables									
Rice (mchele/mpunga/wali)	175g/¾cup					1x			
Ugali					x				
Beans (maharage)	185g/¾cup					1x			
Cabbage (kabichi)	60g/½ cup					1x			
Fruit (matunda)	variable					2x			
USO's (e.g. potatoes)						1x			
Maize						1x			
Soghum						~~1x~~			x
Green peas Cowpeas						1x			x
Other									
Honey (asali)	1tbsp								x
Sweets/Soda						1x			
Sugar cane						3x			
Spaghetti							1x		

Is there any other thing you remember eating/consuming?

Comments

Other non dietary behaviors

Alcohol:.....

Smoking:.....

Tobacco Sniffing:.....

Tobacco Chewing:.....

Khat Chewing:.....

Toothbrush tree use:.....

Figure A.4 Example of a data collection form (page 4/4).

Table A.1

Metadata on El Molo participants (data collected marked by X). If female participants were pregnant or breastfeeding, time in months included in brackets (months). If moulds were collected, side in brackets.

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
EM01	Origalgito	30	M	NA	NA	X	X	X	X	X (R)	
EM02	Marle	20	M	NA	NA	X	X		X	X (R)	
EM03	Marle	16	M	NA	NA	X	X	X	X	X (R)	
EM04	Origalgito	21	M	NA	NA	X	X		X	X (R)	
EM05	Origalgito	62	M	NA	NA	X	X		X	X (L)	
EM06	Massula	35	M	NA	NA	X		X	X	X (R)	
EM07	Massula	39	M	NA	NA	X	X		X	X (R)	
EM08	Origalgito	28	F	N	N	X	X		X	X (R)	
EM09	Origalgite	27	F	N	N	X	X		X	X (R)	
EM10	Unknown	22	F	N	N	X		X	X	X (R)	
EM11	Origalgite	23	F	N	Y (12)	X	X		X	X (R)	
EM12	Longeli	25	F	N	N	X	X		X	X (R)	
EM13	Origalgito	28	F	N	N	X	X		X	X (R)	
EM14	Orisole	25	F	N	N	X	X			X (R)	
EM15	Origalgito	28	F	N	Y (6)	X	X		X	X (R)	
EM16	Unknown	25	F	N	NA	X	X		X	X (R)	
EM17	Orisayo	22	M	NA	NA	X		X	X	X (L)	
EM18	Marle	23	F	Y (4)	N	X	X	X	X	X (L)	
EM19	Orisole	19	F	N	N	X	X		X	X (L)	
EM20	Massula	23	F	Y (8)	N	X	X		X	X (L)	
EM21	Marle	28	F	N	Y (25)	X		X	X	X (L)	
EM22	Massula	26	F	N	Y (3)	X	X		X	X (L)	
EM23	Origaya	26	F	N	Y (2)	X	X	X	X	X (L)	
EM24	Marle	20	F	N	N	X	X		X	X (R)	
EM25	Marle	18	M	NA	NA	X	X		X	X (L)	
EM26	Origaya	20	M	NA	NA	X	X	X	X	X (R)	
EM27	Massula	17	M	NA	NA	X	X		X	X (R)	
EM29	Orikara	15	M	NA	NA	X	X		X	X (R)	
EM30	Massula	28	M	NA	NA	X	X		X	X (L)	
EM31	Massula	20	F	N	Y (8)	X	X		X	X (L)	
EM32	Origalgito	24	M	NA	NA	X	X		X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

Table A.2

Metadata on Turkana participants (data collected marked by X). If female participants were pregnant or breastfeeding, time in months included in brackets (months).

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
TK01	Ngipuchio	25	M	NA	NA	X	X	X	X	X (R)	thumb print
TK02	Ngumacharkwata	30	M	NA	NA	X	X	X	X	X (R)	thumb print
TK03	Ngimosokoko	25	M	NA	NA	X	X	X	X	X (R)	
TK04	Ngimacharkwata	35	M	NA	NA	X		X	X	X (L)	
TK05	Ngiduya	35	F	N	N	X			X	X (R)	thumb print
TK06	Ngisiger	40	M	NA	NA	X	X	X	X		
TK07	Nging'oleroto	19	F	N	N	X	X		X	X (L)	
TK08	Unknown	23	F	N	Y (11)	X	X		X	X (R)	thumb print
TK09	Ngimeturana	39	F	N	N	X	X		X	X (R)	thumb print
TK10	Ngimeturana	35	F	N	N	X	X		X	X (L)	
TK11	Unknown	38	F	N	Y (11)	X	X		X	X (L)	thumb print
TK12	Ngisiger	34	F	N	N	X	X				thumb print
TK13	Ngimeturana	20	F	N	Y (5)	X	X		X	X (L)	
TK14	Ngimeturana	24	F	N	Y (4)	X		X	X		thumb print
TK15	Ngimeturana	43	F	N	N	X		X	X	X (R)	thumb print
TK16	Ngimeturana	28	F	N	Y (30)	X	X		X	X (R)	thumb print
TK17	Ngimosuke	29	F	N	Y (8)	X	X		X	X (L)	thumb print
TK18	Ngimeturana	54	F	N	N	X		X	X	X (L)	thumb print
TK19	Unknown	35	F	Y (6)	N	X	X		X	X (L)	thumb print
TK20	Ng'ing'olereto	15	M	NA	NA	X	X	X	X	X (R)	
TK21	Ebilait	18	M	NA	NA	X	X		X	X (L)	thumb print
TK22	Ng'imosoroko	35	M	NA	NA	X	X		X	X (R)	thumb print
TK23	Ng'imachar	30	M	NA	NA	X	X	X			thumb print
TK24	Unknown	23	M	NA	NA	X	X		X		
TK25	Unknown	19	F	N	N	X	X		X	X (L)	thumb print; epilepsy
TK26	Unknown	28	F	N	Y (36)	X		X			thumb print
TK27	Unknown	32	F	N	N	X	X				thumb print
TK28	Unknown	24	F	N	Y (12)	X	X		X	X (R)	thumb print
TK29	Unknown	17	M	NA	NA	X	X	X	X	X (L)	
TK30	Unknown	18	M	NA	NA	X	X	X	X	X (L)	
TK31	Unknown	19	M	NA	NA	X	X	X	X	X (L)	
TK32	Unknown	25	F	N	Y(8)	X		X	X	X (L)	thumb print
TK33	Ekateit	25	M	NA	NA	X		X	X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

(continues in next page)

Table A.2 Continued

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
TK34	Ngimeturana	30	M	NA	NA	X		X	X	X (L)	thumb print
TK35	Emosorokoit	26	M	NA	NA	X	X	X	X	X (L)	thumb print
TK36	Ekateit	26	M	NA	NA	X		X	X	X (L)	
TK37	Ebilait	21	M	NA	NA	X		X	X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

Table A.3

Metadata on Luhya participants in Webuye (data collected marked by X). If female participants were pregnant or breastfeeding, time in months included in brackets (months).

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
WB01	Banyangari	24	F	N	N	X		X	X	X (L)	
WB02	Kobolo	38	M	NA	NA	X	X		X	X (L)	
WB03	Bakobulo	42	M	NA	NA	X	X		X	X (R)	
WB04	Lukulu	42	M	NA	NA	X	X	X	X		
WB05	Baemka	45	M	NA	NA	X		X	X	X (R)	
WB06	Mutsotso	39	F	N	N	X	X		X	X (L)	
WB07	Bakhwami	29	M	NA	NA	X		X	X	X (R)	
WB08	Baala	31	M	NA	NA	X		X	X	X (L)	
WB09	Bakwangwachi	28	M	NA	NA	X		X	X	X (R)	
WB10	Unknown	25	F	N	Y (4)	X	X		X		
WB11	Bamutilu	41	F	N	N	X	X		X	X (R)	
WB12	Balunda	40	M	NA	NA	X		X	X	X (L)	
WB13	Basaya	48	F	N	N	X	X		X	X (R)	
WB14	Vakovoloi	18	F	N	N	X		X	X	X (R)	
WB15	Balunda	27	M	NA	NA	X		X	X	X (L)	
WB16	Bailifuma	42	F	N	Y(8)	X	X		X	X (R)	
WB17	Unknown	20	F	N	N	X	X		X	X (L)	
WB18	Basaya	33	M	NA	NA	X		X	X	X (L)	
WB19	Baengele	21	M	NA	NA	X		X	X	X (R)	
WB20	Basenya	20	F	N	N	X	X		X	X (R)	
WB21	Balukula	21	F	N	N	X	X		X	X (L)	Field Assistant
WB22	Basebacha	40	M	NA	NA	X		X	X	X (R)	Pastor
WB23	Omutecho	43	M	NA	NA	X	X		X	X (L)	Motorbike driver
WB24	Omukitanga	40	F	N	N	X	X		X	X (R)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

(continues in next page)

Table A.3 Continued

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
WB25	Omubuya	46	F	N	N	X	X		X	X (L)	thumb print
WB26	Omukiyabi	71	F	N	N	X	X		X	X (R)	
WB27	Omulako	30	M	NA	NA	X		X	X	X (L)	
WB28	Baala	34	F	N	N	X	X		X	X (R)	
WB29	Baengele	29	F	N	N	X	X		X	X (R)	
WB30	Bakuta	40	F	N	N	X	X		X	X (L)	
WB31	Bakhwami	55	M	NA	NA	X	X	X	X	X (R)	
WB32	Bauma	25	M	NA	NA	X		X	X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

Table A.4

Metadata on Luhya participants in Port Victoria (data collected marked by X). If female participants were pregnant or breastfeeding, time in months included in brackets (months).

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
PT02	Bunyala	55	F	N	N	X	X		X	X (R)	thumb print
PT03	Bunyala	38	F	N	Y(12)	X	X		X	X (R)	
PT05	Bunyala	35	M	NA	NA	X			X		
PT06	Bukoma	45	M	NA	NA	X		X	X	X (L)	
PT07	Balwani	38	F	N	N	X	X		X	X (R)	
PT11	Samia	32	F	N	N	X	X		X	X (L)	
PT12	Unknown	43	M	NA	NA	X		X	X	X (R)	
PT13	Unknown	35	F	N	Y (60)	X	X		X	X (R)	
PT14	Mudembi	40	M	NA	NA	X		X	X	X (R)	
PT15	Bunyala	22	M	NA	NA	X	X	X	X	X (R)	toe nail plus fingernail
PT16	Bunyala	32	M	NA	NA	X		X	X	X (R)	
PT17	Magombe	48	M	NA	NA	X		X	X	X (L)	
PT19	Ganga	52	F	N	N	X	X				
PT21	Namalo	22	F	N	Y (19)	X	X	X	X	X (R)	toe nail plus fingernail
PT22	Bukusi	33	M	NA	NA	X		X	X	X (R)	thumb print
PT23	Budula	53	F	NA	NA	X	X		X	X (R)	thumb print
PT24	Samia	35	F	NA	NA	X	X		X	X (R)	
PT26	Rwambwa	41	M	NA	NA	X		X	X	X (L)	
PT38	Bukoma	58	M	NA	NA	X		X	X	X (L)	thumb print

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

(continues in next page)

Table A.4 Continued

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
PT39	Marenga	34	M	NA	NA	X		X	X	X (L)	
PT40	Bunyala	40	M	NA	NA	X		X	X	X (R)	
PT45	Mudembi	31	M	NA	NA	X	X	X	X	X (R)	
PT46	Unknown	30	F	N	N	X	X		X	X (R)	thumb print
PT47	Samia	46	F	N	N	X	X	X	X	X (R)	
PT48	Bukoma	45	F	N	N	X	X		X	X (L)	
PT49	Samia	24	F	Y (5)	N	X	X	X	X	X (R)	
PT51	Mudembi	30	F	N	N	X	X	X	X	X (R)	
PT52	Bunyala	32	F	N	N	X	X		X	X (R)	
PT58	Bunyala	17	F	N	N	X					
PT63	Unknown	23	M	NA	NA	X		X	X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

Table A.5

Metadata on Luo participants in Port Victoria (data collected marked by X). If female participants were pregnant or breastfeeding, time in months included in brackets (months).

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
PT01	Alego Usongo	39	F	Y (6.5)	N	X	X		X	X (R)	
PT08	Jaligenma Karadolo	41	M	NA	NA	X	X	X	X	X (L)	
PT10	Napunyi	24	M	NA	NA	X		X	X	X (L)	
PT18	Hakati	36	M	NA	NA	X	X	X	X	X (R)	
PT20	Rwambwa	35	M	NA	NA	X		X	X	X (L)	
PT25	Samia	21	F	NA	NA	X	X		X	X (L)	
PT27	Kakamega	38	F	N	Y (48)	X	X		X	X (L)	
PT30	Samia	54	M	NA	NA	X		X	X	X (R)	
PT31	Siginga	19	F	N	Y (12)	X	X	X	X	X (R)	
PT32	Port	19	M	NA	NA	X		X	X	X (R)	
PT33	Rwambwa	22	F	N	Y(12)	X	X	X	X	X (R)	
PT35	Siaya	35	M	NA	NA	X		X	X	X (R)	thumb print
PT36	Ugenya	35	M	NA	NA	X		X	X	X (L)	
PT37	Suba	25	F	N	N	X	X	X	X	X (R)	
PT41	Bunuku	40	M	NA	NA	X		X	X	X (R)	
PT42	Namalo	40	M	NA	NA	X	X	X	X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

(continues in next page)

Table A.5 Continued

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
PT43	Bukhoba	34	F	N	N	X	X		X	X (R)	
PT44	Usonga	29	M	NA	NA	X	X	X	X	X (L)	
PT50	Osieko	28	F	N	N	X	X	X	X	X (R)	
PT54	Usonga	40	F	N	N	X		X	X	X (R)	
PT56	Samia	25	F	N	N	X	X		X	X (L)	
PT59	Banyekera	27	M	NA	NA	X		X	X	X (L)	
PT60	Namberigre	24	M	NA	NA	X	X	X	X	X (L)	
PT61	Osieko	38	F	N	N	X	X	X	X	X (L)	
PT62	Siaya	32	M	NA	NA	X	X	X	X	X (L)	
PT64	Alego Usongo	19	F	N	N	X		X	X	X (L)	
PT65	Gem	38	F	N	N	X	X		X	X (R)	
PT66	Port	28	F	N	Y (19)	X	X		X	X (R)	
PT67	Usenge	21	F	N	N	X	X		X	X (L)	

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

Table A.6

Metadata on Baka participants in Cameroon (data collected marked by X). If female participants were pregnant or breastfeeding, time in months included in brackets (months).

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
BK01	Unknown	33	F	N	Y (18)	X	X		X		
BK02	Unknown	13	F	N	N	X	X		X		
BK03	Unknown	24	F	N	N	X	X		X		
BK04	Unknown	22	F	Y(4)	N	X	X		X		
BK05	Unknown	25	F	N	N	X	X		X		
BK06	Unknown	35	F	N	N	X	X		X		
BK07	Unknown	51	F	N	N	X	X		X		
BK08	Unknown	25	M	NA	NA	X	X		X		
BK09	Unknown	60	F	N	N	X	X		X		
BK10	Unknown	30	M	NA	NA	X	X		X		
BK11	Unknown	19	F	N	N	X	X		X		
BK12	Unknown	25	F	N	N	X	X		X		
BK13	Unknown	35	F	N	Y (10)	X	X		X		
BK14	Unknown	30	F	N	Y (13)	X	X		X		
BK15	Unknown	30	F	N	Y (8)	X	X		X		
BK16	Unknown	40	F	N	Y (10)	X	X		X		

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

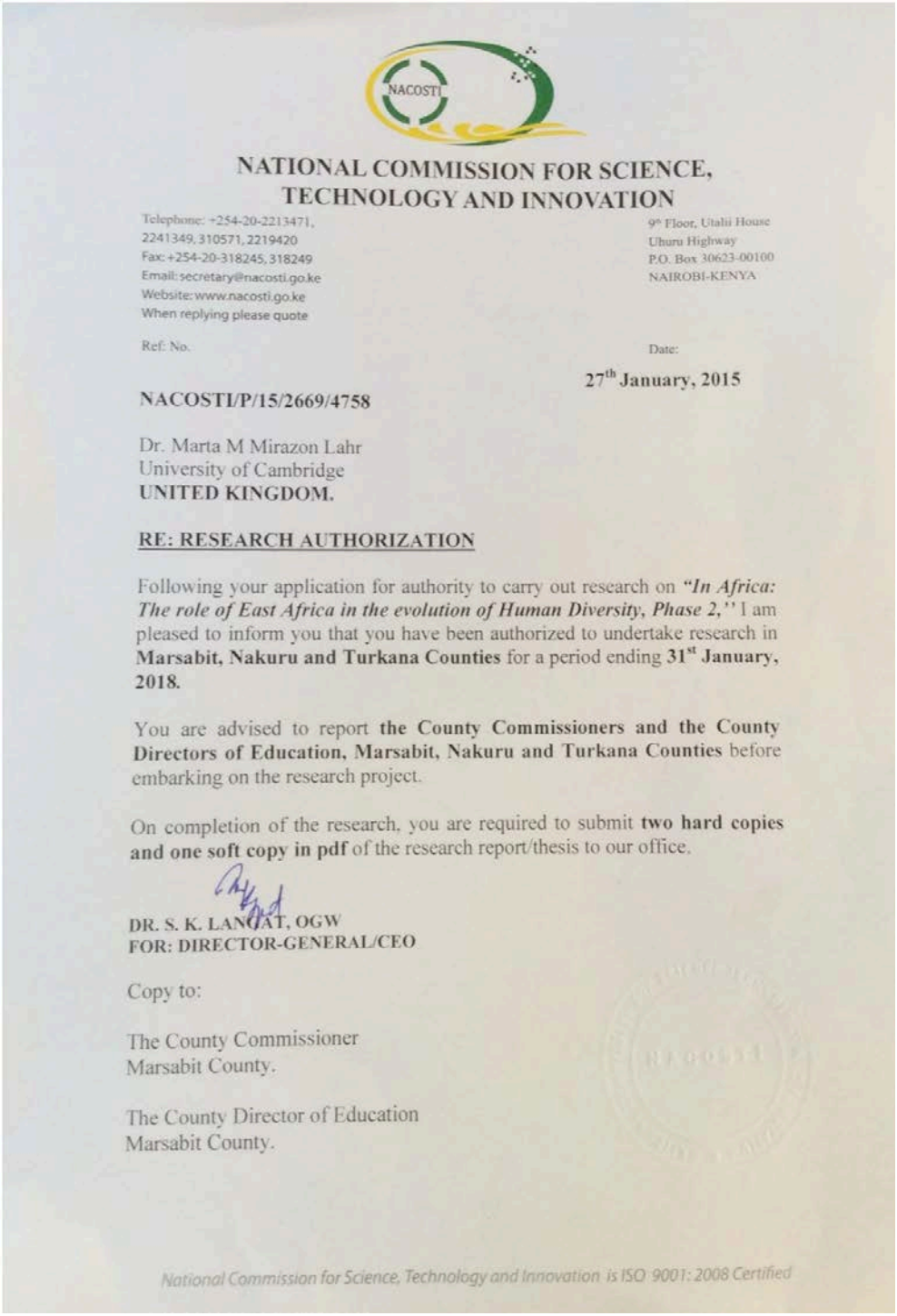
(continues in next page)

Table A.6 Continued

Code	Clan	Age	Sex	Pregnant	Breastfeeding	Diet Survey	Hair	Nail	Breath	Moulds	Comments
BK17	Unknown	30	F	N	N	X	X		X		
BK18	Unknown	19	F	N	Y (12)	X	X		X		
BK19	Unknown	25	F	N	N	X	X		X		
BK20	Unknown	25	F	N	Y (1)	X	X		X		
BK21	Unknown	25	F	N	N	X	X		X		
BK22	Unknown	45	F	N	N	X	X		X		
BK23	Unknown	45	M	NA	NA	X	X		X		
BK24	Unknown	40	M	NA	NA	X	X		X		
BK25	Unknown	20	M	NA	NA	X	X		X		
BK26	Unknown	24	M	NA	NA	X	X		X		
BK27	Unknown	47	F	N	N	X	X		X		
BK28	Unknown	55	M	NA	NA	X	X		X		
BK29	Unknown	57	M	NA	NA	X	X		X		
BK30	Unknown	35	F	N	N	X	X		X		
BK31	Unknown	18	F	N	Y (2)	X	X				
BK32	Unknown	55	M	NA	NA	X	X		X		
BK33	Unknown	35	F	N	Y (4.5)	X	X		X		

Note. M=Male; F=Female; Y=Yes; N=No; NA=Not Applicable; L=Left; R=Right.

A.2 LOGISTICS OF ETHICS AND HEALTH & SAFETY



NACOSTI

NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY AND INNOVATION

Telephone: +254-20-2213471, 2241349, 310571, 2219420
Fax: +254-20-318245, 318249
Email: secretary@nacosti.go.ke
Website: www.nacosti.go.ke
When replying please quote

9th Floor, Utalii House
Uhuru Highway
P.O. Box 30623-00100
NAIROBI-KENYA

Ref: No.

Date:

27th January, 2015

NACOSTI/P/15/2669/4758

Dr. Marta M Mirazon Lahr
University of Cambridge
UNITED KINGDOM.

RE: RESEARCH AUTHORIZATION

Following your application for authority to carry out research on ***"In Africa: The role of East Africa in the evolution of Human Diversity, Phase 2,"*** I am pleased to inform you that you have been authorized to undertake research in **Marsabit, Nakuru and Turkana Counties** for a period ending **31st January, 2018.**

You are advised to report **the County Commissioners and the County Directors of Education, Marsabit, Nakuru and Turkana Counties** before embarking on the research project.

On completion of the research, you are required to submit **two hard copies and one soft copy in pdf** of the research report/thesis to our office.

DR. S. K. LANGAT, OGW
FOR: DIRECTOR-GENERAL/CEO

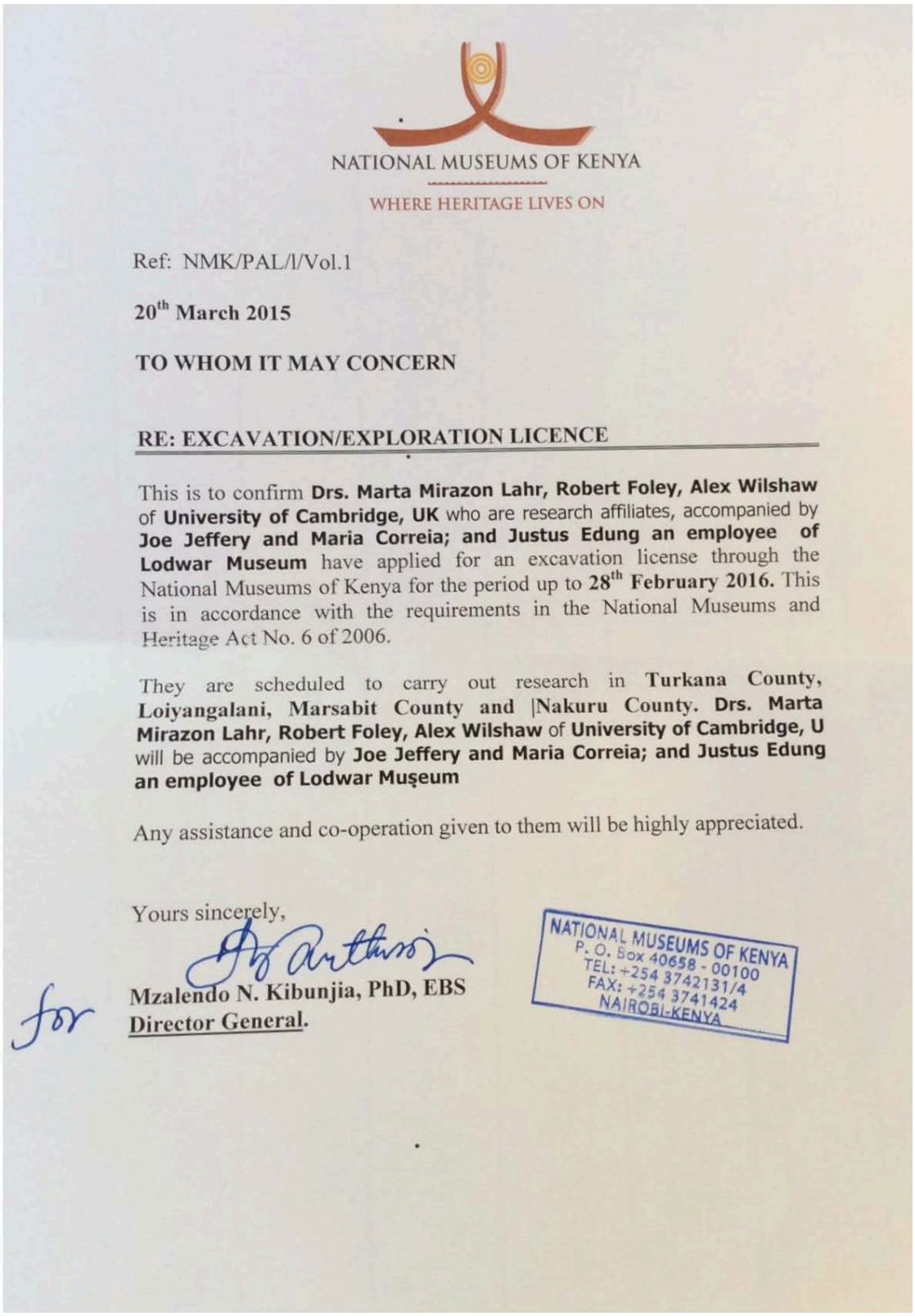
Copy to:

The County Commissioner
Marsabit County.

The County Director of Education
Marsabit County.

National Commission for Science, Technology and Innovation is ISO 9001: 2008 Certified

Figure A.5. Copy of Research Permit 2015-2018 NACOSTI/P?15/2669/4758.



NATIONAL MUSEUMS OF KENYA

WHERE HERITAGE LIVES ON

Ref: NMK/PAL/l/Vol.1

20th March 2015

TO WHOM IT MAY CONCERN

RE: EXCAVATION/EXPLORATION LICENCE

This is to confirm **Drs. Marta Mirazon Lahr, Robert Foley, Alex Wilshaw** of **University of Cambridge, UK** who are research affiliates, accompanied by **Joe Jeffery and Maria Correia; and Justus Edung an employee of Lodwar Museum** have applied for an excavation license through the National Museums of Kenya for the period up to **28th February 2016.** This is in accordance with the requirements in the National Museums and Heritage Act No. 6 of 2006.

They are scheduled to carry out research in **Turkana County, Loiyangalani, Marsabit County and |Nakuru County. Drs. Marta Mirazon Lahr, Robert Foley, Alex Wilshaw** of **University of Cambridge, U** will be accompanied by **Joe Jeffery and Maria Correia; and Justus Edung an employee of Lodwar Mușeum**

Any assistance and co-operation given to them will be highly appreciated.

Yours sincerely,

for **Mzalendo N. Kibunjia, PhD, EBS**
Director General.

NATIONAL MUSEUMS OF KENYA
P. O. Box 40658 - 00100
TEL: +254 3742131/4
FAX: +254 3741424
NAIROBI-KENYA

Figure A.6. Letter of Exploration licence, IN-AFRICA PROJECT, 20th March 2015.

Ethics Statement

The following is an ethics approval statement for Maria Ana Correia's PhD fieldwork project.

Maria Ana Correia's PhD project aims to study diet in Late Quaternary African populations based on a framework built on living Kenyan populations with different subsistence patterns. In this framework, adult individuals will be interviewed (for a brief dietary assessment), their hair will be collected (for isotopic analysis) and dental impressions will be taken (for dental morphology studies).

We have carefully reviewed this PhD project description and concluded that it does not raise any ethical issues and that it complies with the Guiding Principles set on the "University of Cambridge Policy on the Ethics of Research Involving Human Participants and Personal Data".

More specifically and relevant to the project under question:

- any harm to those involved in or affected by research is minimized in all instances and all participants are warned in advance about any potential risks of harm, however slight these might seem;
- free and informed consent will be obtained from all participants in research at an appropriate point in the research process where they will also be informed of the purpose, methods and intended use of the research;
- the PhD student will respect the participant's right to withdraw from research at any time without adverse consequences to the participant;
- any information provided by the participants will be confidential;
- the anonymity of subjects will be respected at all times and documentation protected accordingly;
- the project is approved by the National Museums of Kenya and the Office of the President, and all necessary permits, issued by the Ministry of Science and Technology, are in place.

17. 3. 2015

Prof. Robert Foley
PhD student Supervisor

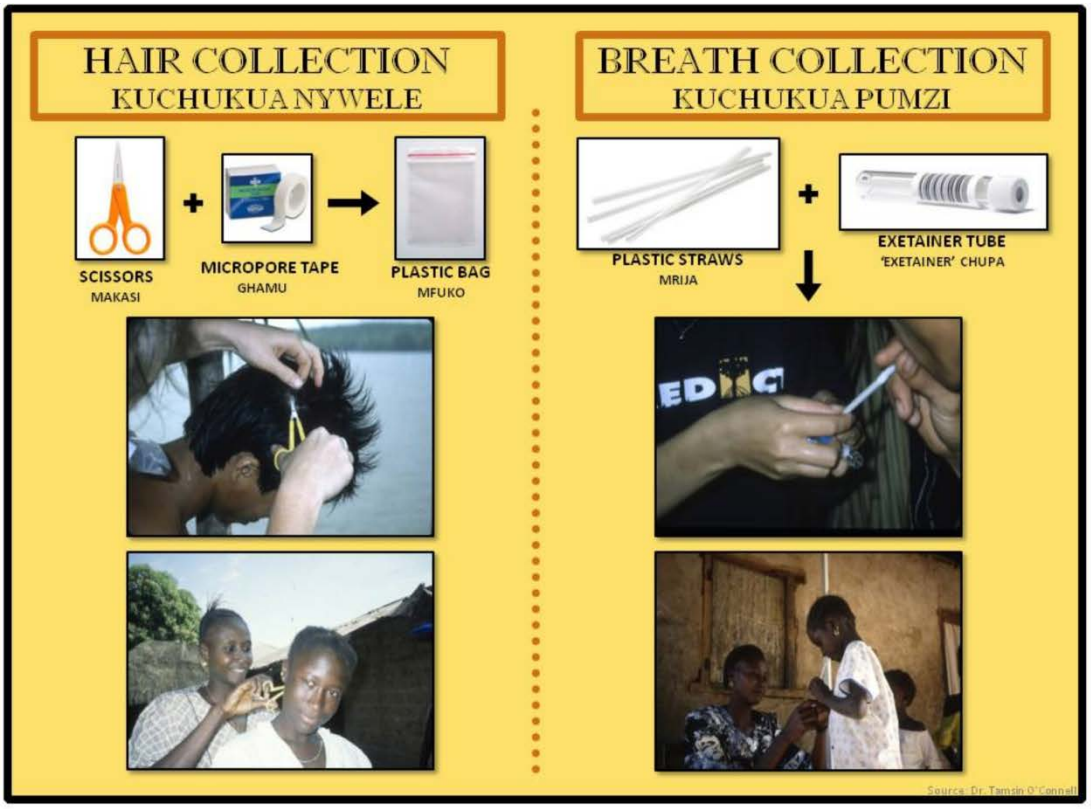
17/03/2015

Dr. Toomas Kivisild
Responsible Officer of the Division of Biological Anthropology

23/03/2015

Prof. Martin Jones
Head of the Department of Archaeology & Anthropology

Figure A.7. Ethics Approval Statement on Fieldwork in Kenya.



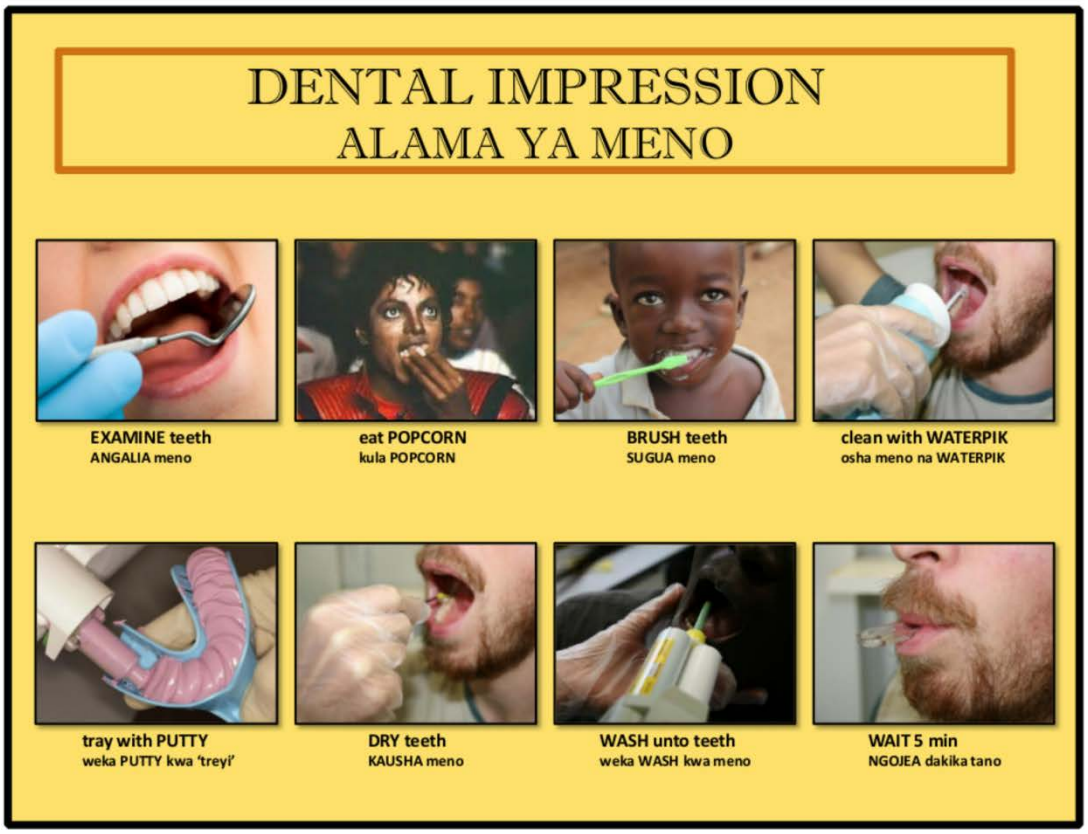
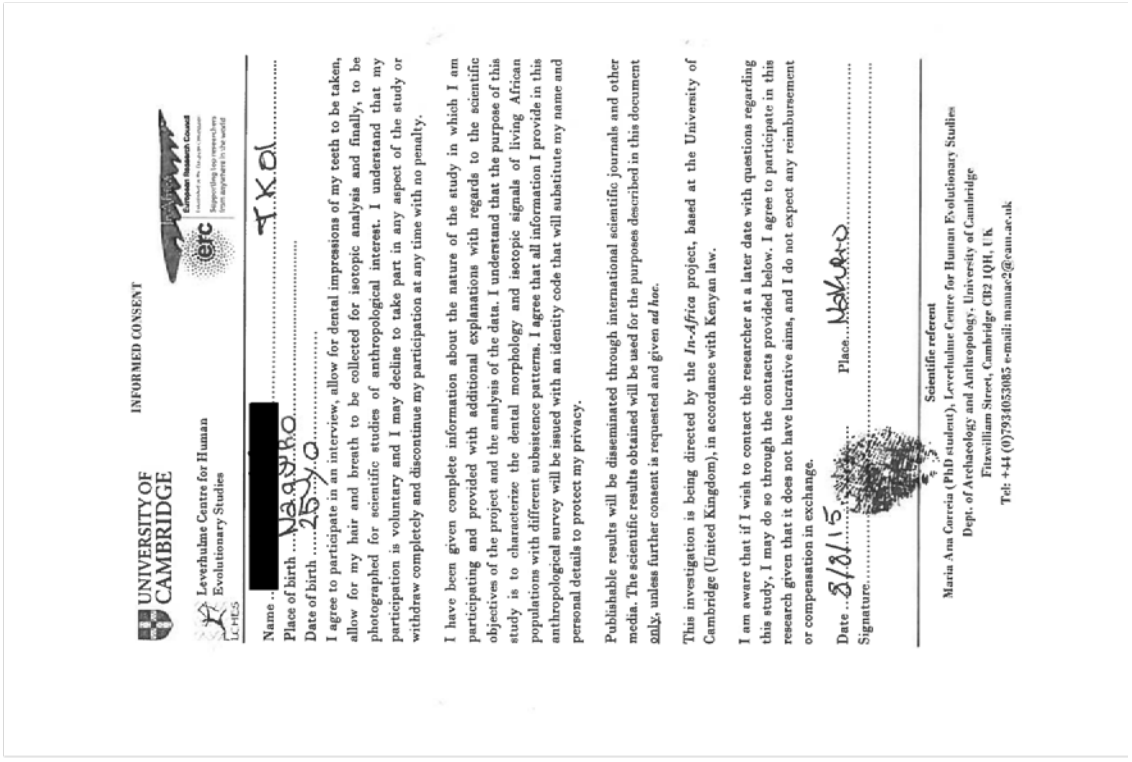


Figure A.8 Posters used to explain hair and breath collection, plus dental impression procedure.



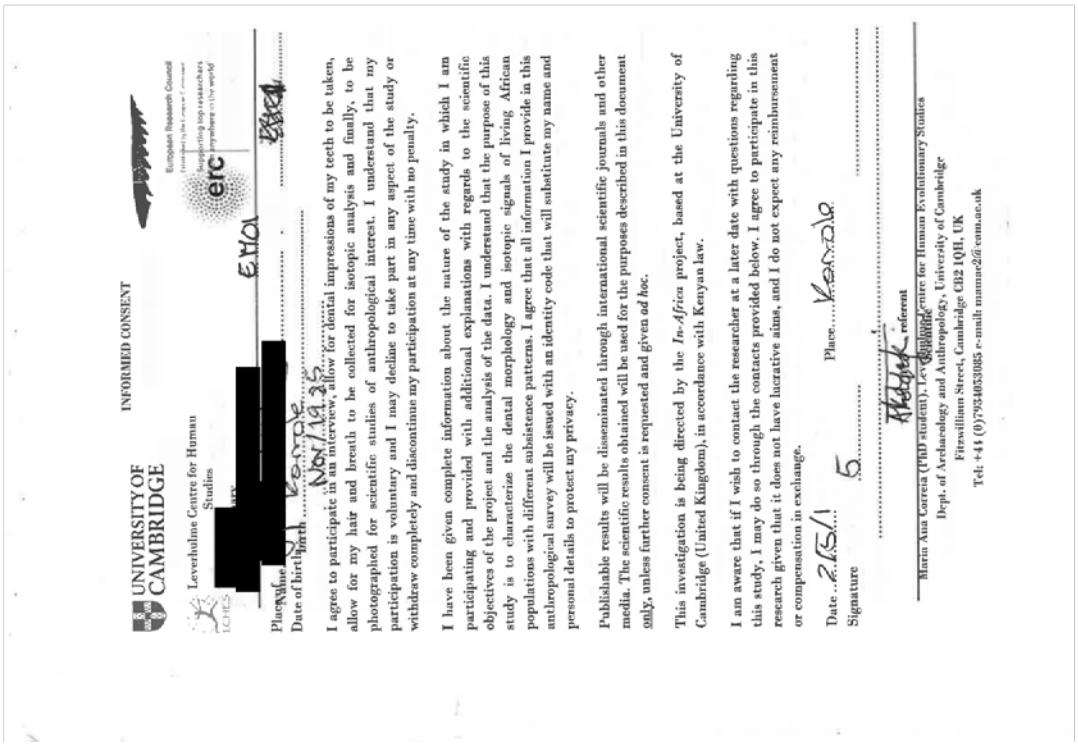


Figure A.9 Examples of signed form consents, with written signature and a thumb print.

Oral Consent Statement

This document confirms that I, Fernando Ramirez Rozzi (undersigned), have received free, individual, informed and oral consent from Baka participants on the study on Diet in Late Quaternary African populations.

This study is part of the PhD project being developed by Maria Ana Correia, at the University of Cambridge, and within the wider project IN-AFRICA, led by Dr. Marta Lahr and Prof. Robert Foley.

Considering participants couldn't read or write and weren't willing to have their fingerprints collected, only oral consent could be and was collected.

During this process, participants agreed to participate in a diet interview and to have their hair and breath collected. They understood their participation was voluntary and that they could decline to take part in any aspect of the study or withdraw completely and discontinue their participation at any time with no penalty.

They were given complete information about the nature of the study and provided with additional explanations with regards to the scientific objectives of the project and the analysis of the data. They were informed that no personal information would be published, and that all data would be in anonymous or aggregated form. They were also informed publishable results are to be disseminated through international scientific journals and other media.

I provide a complete list of the participants, their tribe, age, sex and oral consent confirmation.

Name	Tribe	Age	Sex	Oral Consent
[redacted]	Baka	33	F	Given
[redacted]	Baka	13	F	Given
[redacted]	Baka	24	F	Given
[redacted]	Baka	22	F	Given
[redacted]	Baka	~25	F	Given
[redacted]	Baka	~35	F	Given
[redacted]	Baka	51	F	Given
[redacted]	Baka	~25	M	Given
[redacted]	Baka	60	F	Given
[redacted]	Baka	~30	M	Given
[redacted]	Baka	19	F	Given
[redacted]	Baka	~25	F	Given
[redacted]	Baka	~35	F	Given
[redacted]	Baka	~30	F	Given
[redacted]	Baka	~30	F	Given
[redacted]	Baka	~40	F	Given
[redacted]	Baka	~30	F	Given
[redacted]	Baka	19	F	Given
[redacted]	Baka	~25	F	Given
[redacted]	Baka	25	F	Given
[redacted]	Baka	~25	F	Given
[redacted]	Baka	~45	F	Given
[redacted]	Baka	45	M	Given
[redacted]	Baka	40	M	Given
[redacted]	Baka	20	M	Given
[redacted]	Baka	24	M	Given
[redacted]	Baka	45 - 50	F	Given
[redacted]	Baka	50 - 60	M	Given
[redacted]	Baka	57	M	Given
[redacted]	Baka	30 - 40	F	Given
[redacted]	Baka	8	F	Given
[redacted]	Baka	50 - 60	M	Given
[redacted]	Baka	~30	F	Given

Signature:

Fernando Ramírez Rozzi, CNRS

Date: 21/01/2016

Figure A.10. Dr. Ramírez-Rossi declaration on the informed consent of the Baka.

ETHICS STATEMENT

The following is an ethics approval statement for Fernando Ramirez Rozzi's research fieldworks project.

Fernando Ramirez Rozzi's research project aims to study growth and diet in African populations inhabiting different environments and having different subsistence patterns. For this research, all individuals with the informed consent will be measured, subadult and adult individuals will be interviewed (for familiar relationships and a dietary assessment), their hair will be collected for isotopic analysis and dental impressions will be taken for wear and dental studies.

This project does not raise any ethical issues and it complies with the Guiding Principles set on the 'Policy for individual protection' of the CNRS.

More specifically and relevant to the project under question:

- any arm to those involved in or affected by research is minimized in all instances and all participants are warned in advance about any potential risks of harm, however slight these might seem;
- free and informed consent will be obtained from all participants in research at an appropriate point in the research process where they will also be informed of the purpose, methods and intended use of the research;
- the scientist will respect the participant's right to withdraw from research at any time without adverse consequences to the participant;
- any information provided by the participants will be confidential;
- the anonymity of subjects will be respected at all times and documentation protected accordingly;
- the project is part of the international agreement between the Institut pour la Recherche et le Développement (IRD, France) and the Ministry of Scientific Research and Technology of Cameroon.

.....

Dr. Anne-Marie GUIHARD-COSTA
Head of the Department, UPR2147, CNRS

30/9/2015

Anne-Marie Guihard-Cost
Directeur de l'UPR 2147 du CNRS

Figure A.11. Ethics statement regarding Dr. Ramírez-Rossi work in Cameroon signed by the Head of the CNRS.

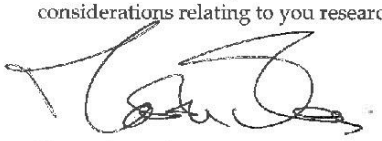
Cambridge, 26th February 2016

Dear Maria,

Thank you for showing me the documentation concerning hair and breath collection amongst the Baka, in Cameroun.

I note that your research involves collaboration with external partics - that seems perfectly reasonable to me. I also note that both you and your collaborator have completed ethics statements within your respective home institutions, and that the collaboration between those institutions has been endorsed by the relevant government bodies.

I therefore conclude you have shown due diligence in the matter, and judge the ethical considerations relating to you research to be in order.



Yours sincerely,

Prof. Martin Jones
Head of Department of Archaeology and Anthropology

Figure A.12. Final ethics statement from the Department of Archaeology and anthropology at Cambridge regarding stable isotope analysis of samples collected in Cameroon.

APPENDIX B

DIET QUESTIONNAIRES DATA ANALYSIS

B.1 DIET QUESTIONNAIRES RESPONSES

All responses to diet questionnaires are in an Excel (.xls) spreadsheet named "APPENDIXB.1_DietQuestionnaires".

B.2 DIET ASSESSMENT STATISTICAL ANALYSES

This section includes the R code that produced all graphical and statistical analysis conducted on diet questionnaires variables. In the cases where the analysis themselves were not fully reported in text (e.g. Shapiro Wilk tests), then the output is included here. Otherwise, only the R code is present and, if applicable, the tables or figures that report the results are mentioned.

The dataset used (APPENDIXB.1_DietQuestionnaires):

```
AllDiet<-read.csv("AllDietData.csv",header=TRUE,sep=",",
            stringsAsFactors = FALSE)

#!making ggplot2 respect the Tribe order on AllDiet
AllDiet$Tribe<-factor(AllDiet$Tribe,levels=unique(AllDiet$Tribe))
AllDiet$Food<-factor(AllDiet$Food,
                levels=c("Animal","Fish",
                      "C3","C4","Other"))
levels(AllDiet$Tribe) <- c("El Molo", "Turkana", "Luhya (Webuye)","Luhya (Port Vict.)","Luo (Po
rt Vict.)")
levels(AllDiet$Food) <- c("Animal","Fish","C3 plant","C4 plant","Other")
```

The functions used:

```
#!setting decimals
fmt_decimals <- function(decimals=0){
   function(x) format(x,nsmall = decimals,scientific = FALSE)
}
#calculate outliers
is_outlier <- function(x) {
  return(x < quantile(x, 0.25,na.rm=TRUE) - 1.5 * IQR(x,na.rm=TRUE) | x > quantile(x, 0.75,na.r
m=TRUE) + 1.5 * IQR(x,na.rm=TRUE))
}
#my  theme
my_theme <- theme(axis.text=element_text(size=11,
                                colour="black"),
      axis.ticks=element_line (size=0.5,colour="black"),
      axis.title=element_text(size=12),
      panel.grid.minor = element_blank(),
      panel.background = element_blank(),
      panel.grid.major = element_blank(),
      panel.border = element_rect(colour = "black",
                           fill=NA, size=0.5))
```

```
#label boxplots with n=
give.n <- function(x){
  return(data.frame(y = median(x)*1.05,
                    label = paste0("n=",length(x))))
  # experiment with the multiplier to find the perfect position
}
```

B.2.1. DIETARY INTAKE

B.2.1.1 Summarising Dietary Intake

First, weight, energy, and macronutrient intake were summarised by calculating classic and robust measures of central tendency and variability: means and standard deviations; and medians and interquartile ranges (IQR) (Table 5.2). These descriptive statistics were calculated by population overall and by gender.

```
#by population
AllDiet2<-AllDiet %>%
  group_by(Row.Labels,Tribe)%>%
  summarise(Wt=sum(Weight,na.rm=T),
            En=sum(Energy,na.rm=T),
            Pr=sum(Protein,na.rm=T),
            Ft=sum(Fat,na.rm=T),
            Cb=sum(Carbohydrate,na.rm=T))
AllDiet3<-AllDiet2%>%
  complete(Tribe)%>%
  group_by(Tribe) %>%
  summarise(meanWeight=round(mean(Wt, na.rm=TRUE),1),
            sdWeight=round(sd(Wt, na.rm=TRUE),1),
            medWeight=round(median(Wt,na.rm=T),1),
            IQRWeight=round(IQR(Wt, na.rm=TRUE),1),
            nWeight=sum(!is.na(Wt)),
            meanEnergy=round(mean(En, na.rm=TRUE),1),
            sdEnergy=round(sd(En, na.rm=TRUE),1),
            medEnergy=round(median(En, na.rm=TRUE),1),
            IQREnergy=round(IQR(En, na.rm=TRUE),1),
            nEnergy=sum(!is.na(En)),
            meanProtein=round(mean(Pr, na.rm=TRUE),1),
            sdProtein=round(sd(Pr,na.rm=TRUE),1),
            medianProtein=round(median(Pr, na.rm=TRUE),1),
            IQRProtein=round(IQR(Pr,na.rm=TRUE),1),
            nProtein=sum(!is.na(Pr)),
            meanFat=round(mean(Ft, na.rm=TRUE),1),
            sdFat=round(sd(Ft, na.rm=TRUE),1),
            medianFat=round(median(Ft, na.rm=TRUE),1),
            IQRFat=round(IQR(Ft, na.rm=TRUE),1),
            nFat=sum(!is.na(Ft)),
            meanCarb=round(mean(Cb, na.rm=TRUE),1),
            sdCarbohydrate=round(sd(Cb,na.rm=TRUE),1),
            medianCarb=round(median(Cb, na.rm=TRUE),1),
```

```
        IQRCarbohydrate=round(IQR(Cb,na.rm=TRUE),1),
        nCarbohydrate=sum(!is.na(Cb))
        )
#by pop+sex
AllDiet4<-AllDiet %>%
 group_by(Row.Labels,Tribe,Sex)%>%
 summarise(Wt=sum(Weight,na.rm=T),
        En=sum(Energy,na.rm=T),
        Pr=sum(Protein,na.rm=T),
        Ft=sum(Fat,na.rm=T),
        Cb=sum(Carbohydrate,na.rm=T))
AllDiet5<-AllDiet4%>%
 complete(Tribe,Sex)%>%
 group_by(Tribe,Sex) %>%
 summarise(meanWeight=round(mean(Wt, na.rm=TRUE),1),
        sdWeight=round(sd(Wt, na.rm=TRUE),1),
        medWeight=round(median(Wt,na.rm=T),1),
        IQRWeight=round(IQR(Wt, na.rm=TRUE),1),
        nWeight=sum(!is.na(Wt)),
        meanEnergy=round(mean(En, na.rm=TRUE),1),
        sdEnergy=round(sd(En, na.rm=TRUE),1),
        medEnergy=round(median(En, na.rm=TRUE),1),
        IQREnergy=round(IQR(En, na.rm=TRUE),1),
        nEnergy=sum(!is.na(En)),
        meanProtein=round(mean(Pr, na.rm=TRUE),1),
        sdProtein=round(sd(Pr,na.rm=TRUE),1),
        medianProtein=round(median(Pr, na.rm=TRUE),1),
        IQRProtein=round(IQR(Pr,na.rm=TRUE),1),
        nProtein=sum(!is.na(Pr)),
        meanFat=round(mean(Ft, na.rm=TRUE),1),
        sdFat=round(sd(Ft, na.rm=TRUE),1),
        medianFat=round(median(Ft, na.rm=TRUE),1),
        IQRFat=round(IQR(Ft, na.rm=TRUE),1),
        nFat=sum(!is.na(Ft)),
        meanCarb=round(mean(Cb, na.rm=TRUE),1),
        sdCarbohydrate=round(sd(Cb,na.rm=TRUE),1),
        medianCarb=round(median(Cb, na.rm=TRUE),1),
        IQRCarbohydrate=round(IQR(Cb,na.rm=TRUE),1),
        nCarbohydrate=sum(!is.na(Cb)))
```

B.2.1.2 Visualising dietary intake

Next, the data was visualised in boxplots, with outliers labelled based on IQR (Figure 5.1).

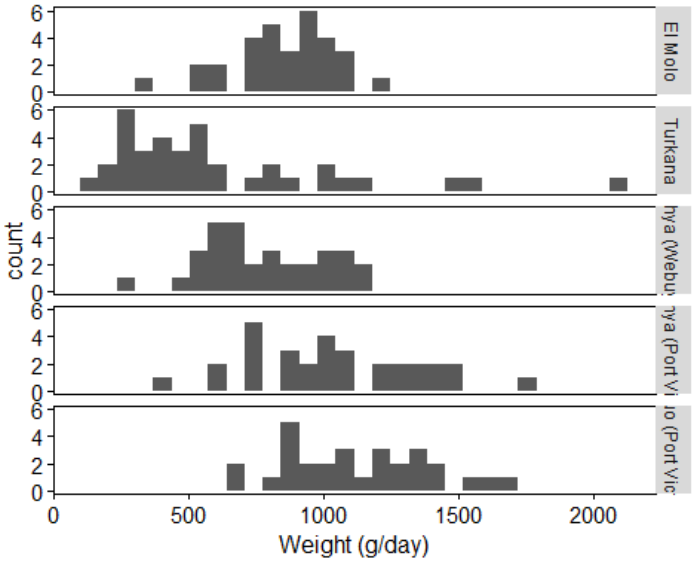
```
#wet weight
AllDiet2%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Wt),
                        as.character(Row.Labels),as.numeric(NA)))) %>%
 ggplot(.,aes(x=Tribe,y=Wt))+
 geom_boxplot()+
```

```
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  stat_summary(fun.data = give.n, geom = "text",size=4)+
  my_theme+
  ylab(" Wet Weight (g/day)")

#wet energy
AllDiet2%>%
  group_by(Tribe)%>%
  mutate(outlier = as.character(ifelse(is_outlier(En),
                                as.character(Row.Labels),as.numeric(NA)))) %>%
  ggplot(.,aes(x=Tribe,y=En))+
  geom_boxplot()+
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  stat_summary(fun.data = give.n, geom = "text",size=4)+
  my_theme+
  ylab(" Wet Energy (kcal/day)")

#protein
AllDiet2%>%
  group_by(Tribe)%>%
  mutate(outlier = as.character(ifelse(is_outlier(Pr),
                                as.character(Row.Labels),as.numeric(NA)))) %>%
  ggplot(.,aes(x=Tribe,y=Pr))+
  geom_boxplot()+
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  stat_summary(fun.data = give.n, geom = "text",size=4)+
  my_theme+
  ylab("Protein (g/day)")

#fat
AllDiet2%>%
  group_by(Tribe)%>%
  mutate(outlier = as.character(ifelse(is_outlier(Ft),
                                as.character(Row.Labels),as.numeric(NA)))) %>%
  ggplot(.,aes(x=Tribe,y=Ft))+
  geom_boxplot()+
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  stat_summary(fun.data = give.n, geom = "text",size=4)+
  my_theme+
  ylab("Fat (g/day)")

#Carbohydrate
AllDiet2%>%
  group_by(Tribe)%>%
  mutate(outlier = as.character(ifelse(is_outlier(Cb),
                                as.character(Row.Labels),as.numeric(NA)))) %>%
  ggplot(.,aes(x=Tribe,y=Cb))+
  geom_boxplot()+
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
```

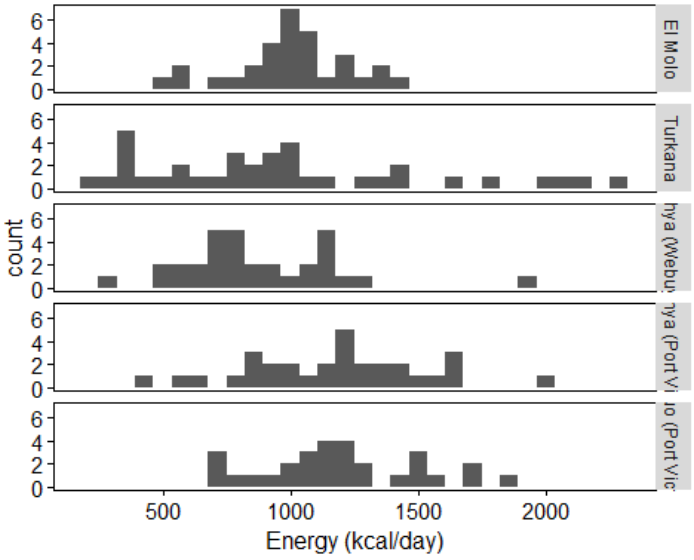
```
  stat_summary(fun.data = give.n, geom = "text",size=4)+
  my_theme+
  ylab("Carbohydrate (g/day)")
```

Histograms were also used in order to visualise their normality.

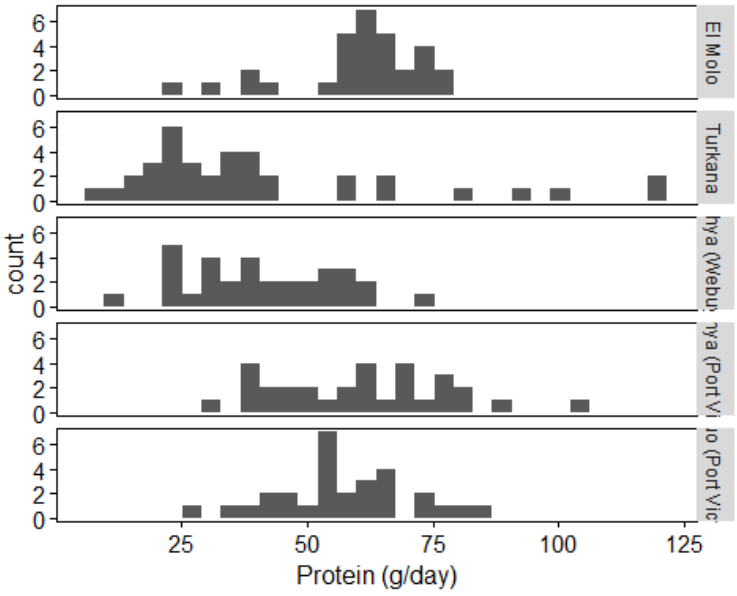
```
AllDiet2%>%
  ggplot(.,aes(x=Wt))+
  geom_histogram()+
  facet_grid(Tribe~.)+
  my_theme+
  xlab("Weight (g/day)")
```



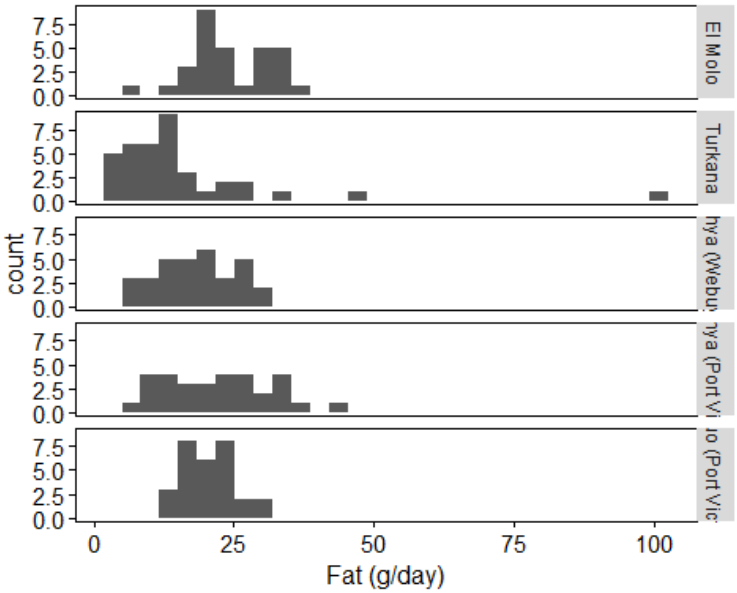

```
AllDiet2%>%
  ggplot(.,aes(x=En))+
  geom_histogram()+
  facet_grid(Tribe~.)+
  my_theme+
  xlab("Energy (kcal/day)")
```



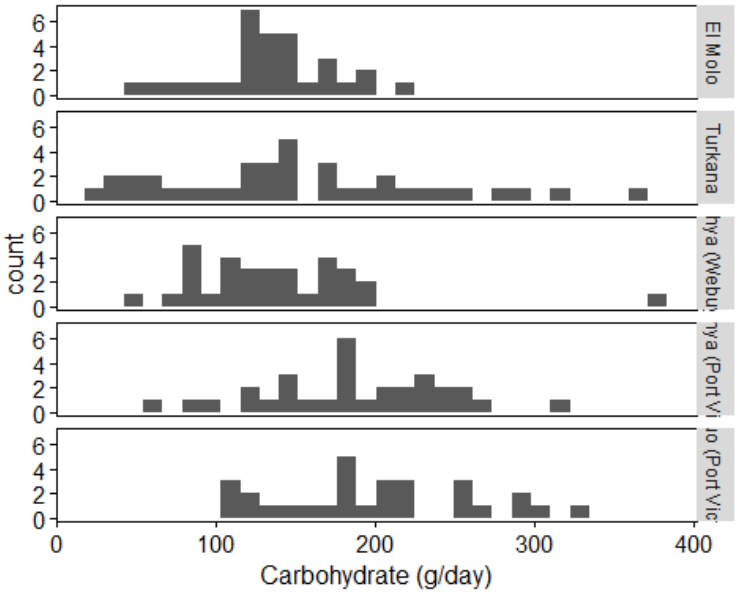
```
AllDiet2%>%
 group_by(Tribe)%>%
 ggplot(.,aes(x=Pr))+
geom_histogram()+
 facet_grid(Tribe~.)+
 my_theme+
 xlab("Protein (g/day)")
```




```
AllDiet2%>%
 group_by(Tribe)%>%
 ggplot(.,aes(x=Ft))+
 geom_histogram()+
 facet_grid(Tribe~.)+
 my_theme+
 xlab("Fat (g/day)")
```

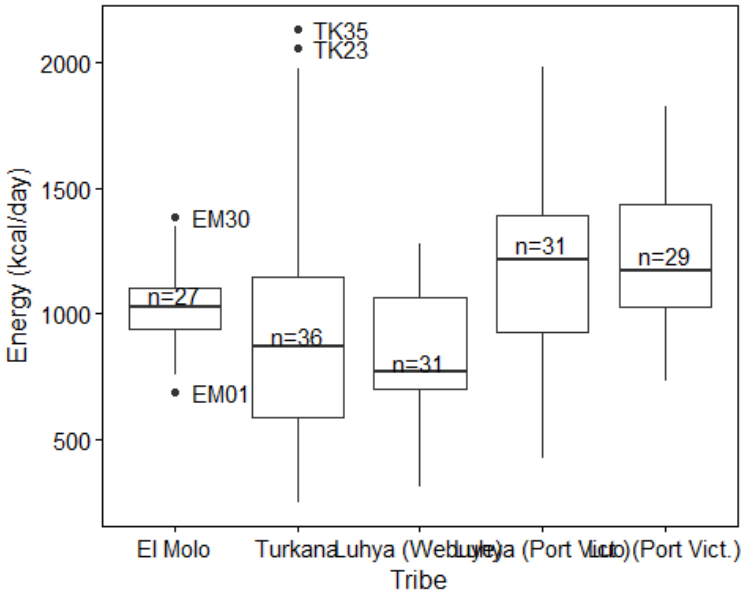


```
AllDiet2%>%
 group_by(Tribe)%>%
 ggplot(.,aes(x=Cb))+
 geom_histogram()+
 facet_grid(Tribe~.)+
 my_theme+
 xlab("Carbohydrate (g/day)")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

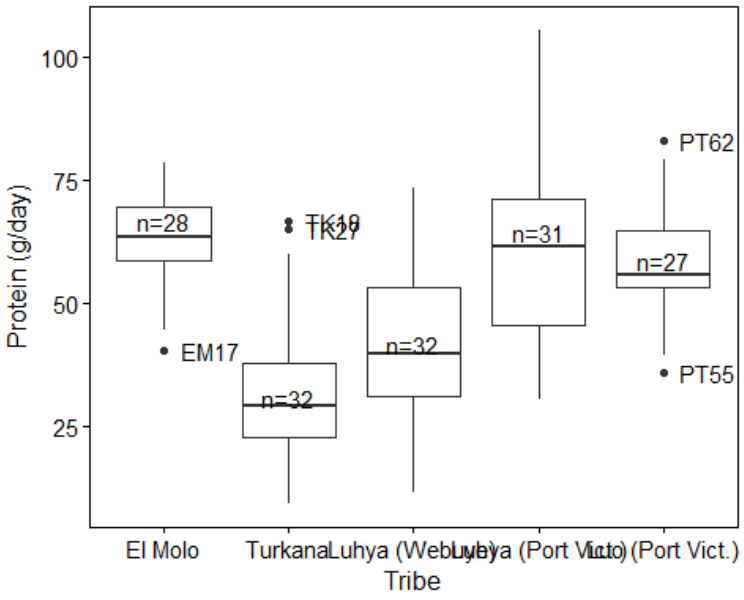


Boxplots were redrawn after filtering outliers identified in Figure 5.1.

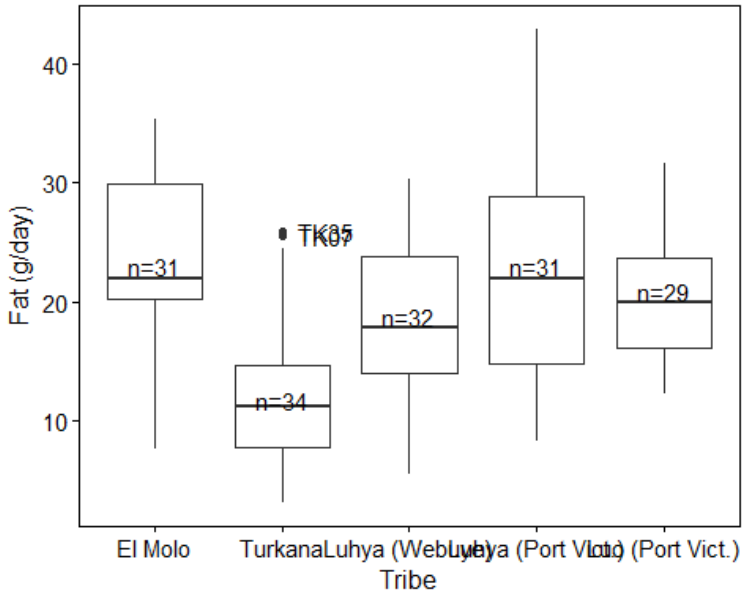
```
Eout<-AllDiet2%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(En),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 filter(is.na(outlier))#data

Eout%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(En),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 ggplot(.,aes(x=Tribe,y=En))+
 geom_boxplot()+
 geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
 stat_summary(fun.data = give.n, geom = "text",size=4)+
 my_theme+
 ylab("Energy (kcal/day)")#graph
```



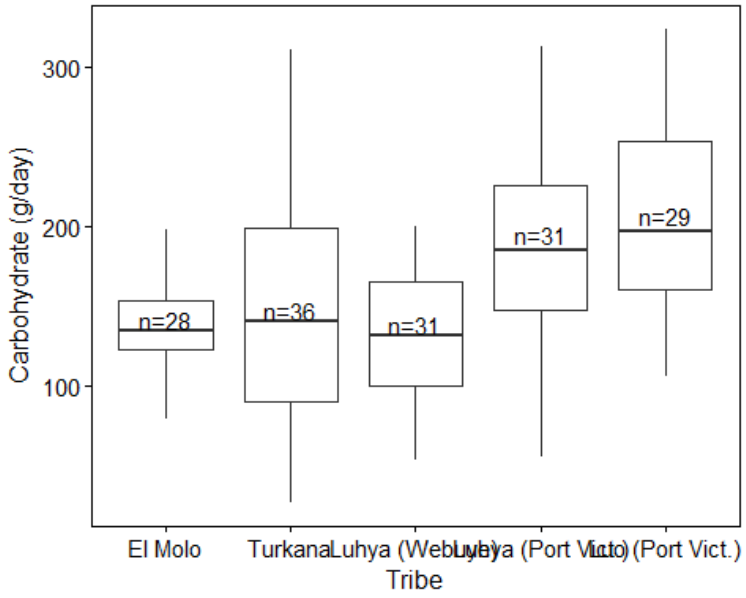

```
Pout<-AllDiet2%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Pr),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 filter(is.na(outlier))#data
Pout%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Pr),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 ggplot(.,aes(x=Tribe,y=Pr))+
 geom_boxplot()+
 geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
 stat_summary(fun.data = give.n, geom = "text",size=4)+
 my_theme+
 ylab("Protein (g/day)")#graph
```



```
Fout<-AllDiet2%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Ft),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 filter(is.na(outlier))#data
Fout%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Ft),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 ggplot(.,aes(x=Tribe,y=Ft))+
 geom_boxplot()+
 geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
 stat_summary(fun.data = give.n, geom = "text",size=4)+
 my_theme+
 ylab("Fat (g/day)")#graph
```




```
Cout<-AllDiet2%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Cb),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 filter(is.na(outlier))

Cout%>%
 group_by(Tribe)%>%
 mutate(outlier = as.character(ifelse(is_outlier(Cb),
                        as.character(Row.Labels),as.numeric(NA))))%>%
 ggplot(.,aes(x=Tribe,y=Cb))+
 geom_boxplot()+
 geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
 stat_summary(fun.data = give.n, geom = "text",size=4)+
 my_theme+
 ylab("Carbohydrate (g/day)")#graph
```



B.2.1.3 Normality

Shapiro Wilk tests were used to check for normality within each variable. However, this test is very sensitive to outliers (Field et al., 2012, p. 182). In addition, also consider that in big samples, it is easy for Shapiro Wilk to get significant results (Field et al., 2012, p.182), whereas in small samples Shapiro Wilk has low power (Field et al., 2012, p. 193) (Table 5.2).

```
sTWt<-AllDiet2 %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Wt)))
pander(sTWt)
```

Tribe	statistic	p.value	method
El Molo	0.9628864	3.471370e-01	Shapiro-Wilk normality test
Turkana	0.8266801	4.701839e-05	Shapiro-Wilk normality test
Luhya (Webuye)	0.9597570	2.704980e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9880175	9.743707e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9769006	7.547558e-01	Shapiro-Wilk normality test

```
sTSwt<-AllDiet4 %>%
  group_by(Tribe,Sex) %>%
  do(tidy(shapiro.test(.$Wt)))
pander(sTSwt)
```

Tribe	Sex	statistic	p.value	method
El Molo	F	0.9118460	0.10748632	Shapiro-Wilk normality test
El Molo	M	0.9501535	0.56309071	Shapiro-Wilk normality test
Turkana	F	0.9223346	0.12499788	Shapiro-Wilk normality test
Turkana	M	0.9102848	0.08703376	Shapiro-Wilk normality test
Luhya (Webuye)	F	0.9582025	0.62931876	Shapiro-Wilk normality test
Luhya (Webuye)	M	0.9544939	0.56403179	Shapiro-Wilk normality test

Luhya (Port Vict.)	F	0.9746412	0.89336404	Shapiro-Wilk normality test
Luhya (Port Vict.)	M	0.9826351	0.98728547	Shapiro-Wilk normality test
Luo (Port Vict.)	F	0.9764916	0.93989345	Shapiro-Wilk normality test
Luo (Port Vict.)	M	0.9554659	0.64827630	Shapiro-Wilk normality test

```
sTEn<-AllDiet2 %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$En)))
pander(sTEn)
```

Tribe	statistic	p.value	method
El Molo	0.9599379	0.29066006	Shapiro-Wilk normality test
Turkana	0.9214703	0.01230944	Shapiro-Wilk normality test
Luhya (Webuye)	0.9100530	0.01127451	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9919273	0.99717686	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9580016	0.29322300	Shapiro-Wilk normality test

```
sTSEn<-AllDiet4 %>%
  group_by(Tribe,Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(sTSEn)
```

Tribe	Sex	statistic	p.value	method
El Molo	F	0.8921261	0.05023484	Shapiro-Wilk normality test
El Molo	M	0.9398698	0.41665849	Shapiro-Wilk normality test
Turkana	F	0.8674748	0.01308403	Shapiro-Wilk normality test
Turkana	M	0.9237448	0.15059643	Shapiro-Wilk normality test
Luhya (Webuye)	F	0.8584072	0.01815792	Shapiro-Wilk normality test
Luhya (Webuye)	M	0.9379490	0.32464665	Shapiro-Wilk normality test
Luhya (Port Vict.)	F	0.9860421	0.99255988	Shapiro-Wilk normality test
Luhya (Port Vict.)	M	0.9806214	0.97824658	Shapiro-Wilk normality test
Luo (Port Vict.)	F	0.9439517	0.43466357	Shapiro-Wilk normality test
Luo (Port Vict.)	M	0.9272574	0.27916519	Shapiro-Wilk normality test

```
sTPr<-AllDiet2 %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Pr)))
pander(sTPr)
```

Tribe	statistic	p.value	method
El Molo	0.8870956	3.485836e-03	Shapiro-Wilk normality test
Turkana	0.8182702	3.104451e-05	Shapiro-Wilk normality test
Luhya (Webuye)	0.9814906	8.415922e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9735942	6.225623e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9812669	8.690353e-01	Shapiro-Wilk normality test

```
sTSPr<-AllDiet4 %>%
  group_by(Tribe, Sex) %>%
  do(tidy(shapiro.test(.$Pr)))
pander(sTSPr)
```

Tribe	Sex	statistic	p.value	method
El Molo	F	0.7672842	0.0007471484	Shapiro-Wilk normality test
El Molo	M	0.9100271	0.1577286053	Shapiro-Wilk normality test
Turkana	F	0.8833332	0.0245510514	Shapiro-Wilk normality test
Turkana	M	0.8588288	0.0117060231	Shapiro-Wilk normality test
Luhya (Webuye)	F	0.9390915	0.3380345348	Shapiro-Wilk normality test
Luhya (Webuye)	M	0.9594136	0.6511583092	Shapiro-Wilk normality test
Luhya (Port Vict.)	F	0.9620585	0.6704422949	Shapiro-Wilk normality test
Luhya (Port Vict.)	M	0.9376215	0.3886967970	Shapiro-Wilk normality test
Luo (Port Vict.)	F	0.9567413	0.6359896771	Shapiro-Wilk normality test
Luo (Port Vict.)	M	0.9609546	0.7387887979	Shapiro-Wilk normality test

```
sTFt<-AllDiet2 %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Ft)))
pander(sTFt)
```

Tribe	statistic	p.value	method
El Molo	0.9391285	7.806435e-02	Shapiro-Wilk normality test
Turkana	0.5878479	5.330173e-09	Shapiro-Wilk normality test
Luhya (Webuye)	0.9684601	4.580189e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9632979	3.557044e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9707362	5.798287e-01	Shapiro-Wilk normality test

```
sTSFt<-AllDiet4 %>%
  group_by(Tribe, Sex) %>%
  do(tidy(shapiro.test(.$Ft)))
pander(sTSFt)
```

Tribe	Sex	statistic	p.value	method
El Molo	F	0.9297501	0.2155580057	Shapiro-Wilk normality test
El Molo	M	0.9064546	0.1399607856	Shapiro-Wilk normality test
Turkana	F	0.8728541	0.0161581875	Shapiro-Wilk normality test
Turkana	M	0.6539024	0.0000243712	Shapiro-Wilk normality test
Luhya (Webuye)	F	0.9736559	0.8937779825	Shapiro-Wilk normality test
Luhya (Webuye)	M	0.9308431	0.2513991379	Shapiro-Wilk normality test
Luhya (Port Vict.)	F	0.8819699	0.0342268785	Shapiro-Wilk normality test
Luhya (Port Vict.)	M	0.9308949	0.3140759836	Shapiro-Wilk normality test
Luo (Port Vict.)	F	0.9318564	0.2908244714	Shapiro-Wilk normality test
Luo (Port Vict.)	M	0.9498380	0.5581921524	Shapiro-Wilk normality test

```
sTCb<-AllDiet2 %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Cb)))
pander(sTCb)
```

Tribe	statistic	p.value	method
El Molo	0.9736210	0.6233621656	Shapiro-Wilk normality test
Turkana	0.9623947	0.2403850614	Shapiro-Wilk normality test
Luhya (Webuye)	0.8222575	0.0001099537	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9897899	0.9887247011	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9625973	0.3802304986	Shapiro-Wilk normality test

```
sTSCb<-AllDiet4 %>%
  group_by(Tribe, Sex) %>%
  do(tidy(shapiro.test(.$Cb)))
pander(sTSCb)
```

Tribe	Sex	statistic	p.value	method
El Molo	F	0.9485309	0.43370402	Shapiro-Wilk normality test
El Molo	M	0.9681095	0.85039262	Shapiro-Wilk normality test
Turkana	F	0.9242376	0.13555068	Shapiro-Wilk normality test
Turkana	M	0.9643527	0.68728342	Shapiro-Wilk normality test
Luhya (Webuye)	F	0.7798617	0.00148952	Shapiro-Wilk normality test
Luhya (Webuye)	M	0.9574604	0.61603836	Shapiro-Wilk normality test
Luhya (Port Vict.)	F	0.9803215	0.95996945	Shapiro-Wilk normality test
Luhya (Port Vict.)	M	0.9732173	0.91647821	Shapiro-Wilk normality test
Luo (Port Vict.)	F	0.9248318	0.22813041	Shapiro-Wilk normality test
Luo (Port Vict.)	M	0.9152022	0.18748264	Shapiro-Wilk normality test

B.2.2 DIETARY ADEQUACY

Macronutrient intake as percentage of total energy were calculated this way (Table 5.3):

```
#Macronutrient:energy ratio by pop
AllDietRat<-AllDiet3%>%
  mutate(PEr= (meanProtein*0.8*4)/meanEnergy,
         FEr= meanFat*9/meanEnergy,
         CEr= meanCarb*4/meanEnergy)%>%
  select(Tribe,PEr,FEr,CEr)
pander(AllDietRat)

#Macronutrient:energy ratio by pop+sex
AllDietRat2<-AllDiet5%>%
  mutate(PEr= (meanProtein*0.8*4)/meanEnergy,
         FEr= meanFat*9/meanEnergy,
         CEr= meanCarb*4/meanEnergy)%>%
```

```
  select(Tribe,PEr,FEr,CEr)
pander (AllDietRat2)
```

B.2.3. DIETARY COMPOSITION

Dietary composition was represented in terms of proportions using stacked charts. Below is the code for all graphs included in this section of Chapter 5, as well as for all tables, and Shapiro-Wilk tests.

B.2.3.1 Contribution of Macronutrients to Population Dietary Intake

First, Fig. 5.2 illustrates the distribution of macronutrients in dry weight, whereas the Fig. 5.3 does the same for dry energy

```
#Dry Weight
AllDietRaw2<-AllDiet2%>%
  gather(Macronutrient,Weight, Pr, Ft, Cb)
AllDietRaw2$Tribe<-factor(AllDietRaw2$Tribe,
                    levels=unique(rev(AllDietRaw2$Tribe)))

ggplot(AllDietRaw2, aes(x = Tribe, y = Weight, fill = Macronutrient)) +
  geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
  geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
  coord_flip()+
  my_theme+
  scale_fill_grey(start = 0.8, end = .05,
                  breaks=c("Pr","Ft","Cb"),
                  labels=c("Protein","Fat","Carbohydrate"))+
  ylab("Dry Weight (g/day)")+
  scale_y_continuous(expand = c(0, 0), labels = scales::percent)

#Dry Energy
AllDietDryWt<-mutate(AllDiet3,
                     PEr= (meanProtein*0.8*4),
                     FEr= meanFat*9,
                     CEr= meanCarb*4)

AllDietDryWt$Tribe<-factor(AllDietDryWt$Tribe, levels=unique(rev(AllDietDryWt$Tribe)))

AllDietDryWt<-AllDietDryWt%>%
  gather(Macronutrient,Energy, PEr, FEr, CEr)

ggplot(AllDietDryWt, aes(x = Tribe, y = Energy,
                         fill = Macronutrient)) +
  geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
  geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
  coord_flip()+
  my_theme+
  scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
  scale_fill_grey(start = 0.8, end = .05,
                  breaks=c("CEr","FEr","PEr"),
```

```
            labels=c("Carbohydrate","Fat","Protein"))+
  guides(fill = guide_legend(reverse = TRUE))+
  ylab("Dry Energy (kcal/day)")
```

The following is the code for Table 5.5, and the Shapiro-Wilk tests for it.

```
#mean, sd, median, and IQR of dry energy by population
AllDietMacDry<-AllDiet %>%
  mutate(Prkc=Protein*4*0.8,
         Ftkc=Fat*9,
         Cbkc=Carbohydrate*4)%>%
  group_by(Row.Labels,Tribe)%>%
  summarise(Pr=sum(Prkc,na.rm=T),
            Ft=sum(Ftkc,na.rm=T),
            Cb=sum(Cbkc,na.rm=T))
AllDietMacDry2<-AllDietMacDry %>%
  complete(Tribe)%>%
  group_by(Tribe) %>%
  summarise(meanProtein=round(mean(Pr, na.rm=TRUE),1),
            sdProtein=round(sd(Pr,na.rm=TRUE),1),
            medProtein=round(median(Pr,na.rm=T),1),
            IQRProtein=round(IQR(Pr, na.rm=TRUE),1),
            nProtein=sum(!is.na(Pr)),
            meanFat=round(mean(Ft, na.rm=TRUE),1),
            sdFat=round(sd(Ft, na.rm=TRUE),1),
            medFat=round(median(Ft,na.rm=T),1),
            IQRFat=round(IQR(Ft, na.rm=TRUE),1),
            nFat=sum(!is.na(Ft)),
            meanCarb=round(mean(Cb, na.rm=TRUE),1),
            sdCarb=round(sd(Cb,na.rm=TRUE),1),
            medCarb=round(median(Cb,na.rm=T),1),
            IQRCarb=round(IQR(Cb, na.rm=TRUE),1),
            nCarb=sum(!is.na(Cb))
            )

#shapiro wilk on Table 5.5 (dry energy/macronutrients)
##protein in kcal
sDPr<-AllDietMacDry %>%
  group_by(Tribe) %>%
  do (tidy(shapiro.test(.$Pr)))
pander(sDPr)
```

Tribe	statistic	p.value	method
El Molo	0.8870956	3.485836e-03	Shapiro-Wilk normality test
Turkana	0.8182702	3.104451e-05	Shapiro-Wilk normality test
Luhya (Webuye)	0.9814906	8.415922e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9735942	6.225623e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9812669	8.690353e-01	Shapiro-Wilk normality test

```
##fat in kcal
sDFt<-AllDietMacDry %>%
  group_by(Tribe) %>%
  do (tidy(shapiro.test(.$Ft)))
pander(sDFt)
```

Tribe	statistic	p.value	method
El Molo	0.9391285	7.806435e-02	Shapiro-Wilk normality test
Turkana	0.5878479	5.330173e-09	Shapiro-Wilk normality test
Luhya (Webuye)	0.9684601	4.580189e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9632979	3.557044e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9707362	5.798287e-01	Shapiro-Wilk normality test

```
##carbs in kcal
sDCb<-AllDietMacDry %>%
  group_by(Tribe) %>%
  do (tidy(shapiro.test(.$Cb)))
pander(sDCb)
```

Tribe	statistic	p.value	method
El Molo	0.9736210	0.6233621656	Shapiro-Wilk normality test
Turkana	0.9623947	0.2403850614	Shapiro-Wilk normality test
Luhya (Webuye)	0.8222575	0.0001099537	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9897899	0.9887247011	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9625973	0.3802304986	Shapiro-Wilk normality test

B.2.3.2 Contribution of Food Categories to Population Dietary Intake

Below is the code for Fig. 5.4 and Table 5.6. Afterwards, one finds the results of the Shapiro-Wilk tests for this table.

```
AllDiet$Tribe<-factor(AllDiet$Tribe,levels=rev(unique(AllDiet$Tribe)))
AllDiet$Food<-factor(AllDiet$Food,levels=rev(c("Animal","Fish","C3 plant","C4 plant","Other"))
)
#rev is from [here](http://stackoverflow.com/questions/33538059/how-to-change-stacking-order-in-stacke
d-bar-chart-in-r)

AllDietOther<-filter(AllDiet,Food!="Other")

ggplot(AllDietOther, aes(x = Tribe, y = Weight, fill = Food)) +
  geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
  geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
  coord_flip()+
  my_theme+
  scale_fill_grey(start = 0.8, end = .05)+
  scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
  guides(fill = guide_legend(reverse = TRUE))+
  ylab("Wet Weight (g/day)")
```

```
#wet weight summary
AllDietCatWt<-AllDiet %>%
 group_by(Row.Labels,Tribe, Food)%>%
 summarise(Wt=sum(Weight,na.rm=T))
AllDietCatWt2<-AllDietCatWt%>%
 group_by(Tribe,Food) %>%
 summarise(meanWeight=round(mean(Wt, na.rm=TRUE),1),
        sdWeight=round(sd(Wt, na.rm=TRUE),1),
        medWeight=round(median(Wt,na.rm=T),1),
        IQRWeight=round(IQR(Wt, na.rm=TRUE),1),
        nWeight=sum(!is.na(Wt)))
#shapiro wilk on Table 5.6 (wetweight/category)
sCWt<-AllDietCatWt%>%
 group_by(Tribe, Food) %>%
 do (tidy(shapiro.test(.$Wt)))
pander(sCWt)
```

Tribe	Food	statistic	p.value	method
El Molo	Animal	0.9251164	3.234086e-02	Shapiro-Wilk normality test
El Molo	Fish	0.5322597	8.347759e-09	Shapiro-Wilk normality test
El Molo	C3 plant	0.9704093	5.304001e-01	Shapiro-Wilk normality test
El Molo	C4 plant	0.9836570	9.042726e-01	Shapiro-Wilk normality test
El Molo	Other	0.2511883	1.818561e-11	Shapiro-Wilk normality test
Turkana	Animal	0.8208331	3.519282e-05	Shapiro-Wilk normality test
Turkana	Fish	0.7688530	3.263917e-06	Shapiro-Wilk normality test
Turkana	C3 plant	0.9509340	1.034285e-01	Shapiro-Wilk normality test
Turkana	C4 plant	0.6003393	7.773679e-09	Shapiro-Wilk normality test
Turkana	Other	0.1553105	2.545362e-13	Shapiro-Wilk normality test
Luhya (Webuye)	Animal	0.8172299	8.740626e-05	Shapiro-Wilk normality test
Luhya (Webuye)	Fish	0.9061813	8.930659e-03	Shapiro-Wilk normality test
Luhya (Webuye)	C3 plant	0.9346053	5.270124e-02	Shapiro-Wilk normality test
Luhya (Webuye)	C4 plant	0.9528753	1.738591e-01	Shapiro-Wilk normality test
Luhya (Webuye)	Other	0.3342032	6.107939e-11	Shapiro-Wilk normality test
Luhya (Port Vict.)	Animal	0.8863959	3.354081e-03	Shapiro-Wilk normality test
Luhya (Port Vict.)	Fish	0.3821409	2.487758e-10	Shapiro-Wilk normality test
Luhya (Port Vict.)	C3 plant	0.9732929	6.135977e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	C4 plant	0.8973464	6.192382e-03	Shapiro-Wilk normality test
Luhya (Port Vict.)	Other	0.1811971	5.112755e-12	Shapiro-Wilk normality test
Luo (Port Vict.)	Animal	0.9285766	5.054130e-02	Shapiro-Wilk normality test
Luo (Port Vict.)	Fish	0.4763222	4.452699e-09	Shapiro-Wilk normality test
Luo (Port Vict.)	C3 plant	0.9793534	8.215168e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	C4 plant	0.9353570	7.581268e-02	Shapiro-Wilk normality test
Luo (Port Vict.)	Other	0.2060851	1.935630e-11	Shapiro-Wilk normality test

Next, the code for Fig. 5.5, Table 5.7, and respective Shapiro Wilks.

```
AllDietCat2<-filter(AllDietCat,Food!="Other")
AllDietCat2$Tribe<-factor(AllDietCat2$Tribe,
                  levels=rev(unique(AllDietCat2$Tribe)))
AllDietCat2$Food<-factor(AllDietCat2$Food,
                levels=rev(unique(AllDietCat2$Food)))
ggplot(AllDietCat2, aes(x = Tribe, y = meanWt, fill = Food)) +
 geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
 geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
 coord_flip()+
 my_theme+
 scale_fill_grey(start = 0.8, end = .05)+
 scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
 guides(fill = guide_legend(reverse = TRUE))+
 ylab("Dry Weight (g/day)")

#adding protein+fat+carb (dry weight)
AllDietCatDry<-AllDiet%>%
 mutate(WtMacro=Protein+Fat+Carbohydrate)%>%
 group_by(Row.Labels,Tribe,Food)%>%
 summarise(Wt=sum(WtMacro,na.rm=T))
AllDietCatDry2<-AllDietCatDry %>%
 group_by(Tribe,Food) %>%
 summarise(meanWt=round(mean(Wt,na.rm=T),1),
        sdWeight=round(sd(Wt, na.rm=TRUE),1),
        medWeight=round(median(Wt,na.rm=T),1),
        IQRWeight=round(IQR(Wt, na.rm=TRUE),1),
        nWeight=sum(!is.na(Wt)))

#shapiro wilk on Table 5.7 (dryweight/category)
sDTFWt<-AllDietCatDry%>%
 group_by(Tribe, Food) %>%
 do (tidy(shapiro.test(.$Wt)))
pander(sDTFWt)
```

Tribe	Food	statistic	p.value	method
El Molo	Animal	0.9634938	3.598455e-01	Shapiro-Wilk normality test
El Molo	Fish	0.5322597	8.347759e-09	Shapiro-Wilk normality test
El Molo	C3 plant	0.9764944	7.099633e-01	Shapiro-Wilk normality test
El Molo	C4 plant	0.9539727	2.007209e-01	Shapiro-Wilk normality test
El Molo	Other	0.2446236	1.608940e-11	Shapiro-Wilk normality test
Turkana	Animal	0.7651010	2.782781e-06	Shapiro-Wilk normality test
Turkana	Fish	0.5995416	7.586852e-09	Shapiro-Wilk normality test
Turkana	C3 plant	0.9505565	1.005710e-01	Shapiro-Wilk normality test
Turkana	C4 plant	0.9098210	5.589641e-03	Shapiro-Wilk normality test
Turkana	Other	0.1553105	2.545362e-13	Shapiro-Wilk normality test
Luhya (Webuye)	Animal	0.9459916	1.108409e-01	Shapiro-Wilk normality test

Luhya (Webuye)	Fish	0.9061813	8.930659e-03	Shapiro-Wilk normality test
Luhya (Webuye)	C3 plant	0.9369687	6.144043e-02	Shapiro-Wilk normality test
Luhya (Webuye)	C4 plant	0.6442412	1.406682e-07	Shapiro-Wilk normality test
Luhya (Webuye)	Other	0.3358246	6.314896e-11	Shapiro-Wilk normality test
Luhya (Port Vict.)	Animal	0.9434914	1.030938e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	Fish	0.3821409	2.487758e-10	Shapiro-Wilk normality test
Luhya (Port Vict.)	C3 plant	0.9674149	4.508772e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	C4 plant	0.8953389	5.524996e-03	Shapiro-Wilk normality test
Luhya (Port Vict.)	Other	0.1805197	5.052270e-12	Shapiro-Wilk normality test
Luo (Port Vict.)	Animal	0.9750186	7.012228e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	Fish	0.4763222	4.452699e-09	Shapiro-Wilk normality test
Luo (Port Vict.)	C3 plant	0.9621785	3.715134e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	C4 plant	0.9451822	1.370565e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	Other	0.3815439	5.520019e-10	Shapiro-Wilk normality test

Then, the code for Fig. 5.6, Table 5.8, and respective Shapiro Wilks.

```
ggplot(AllDietOther, aes(x = Tribe, y = Energy, fill = Food)) +
 geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
 geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
 coord_flip()+
 my_theme+
 scale_fill_grey(start = 0.8, end = .05)+
 scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
 guides(fill = guide_legend(reverse = TRUE))+
 ylab("Wet Energy (kcal/day)")

#wet energy
AllDietCatEn<-AllDiet %>%
 group_by(Row.Labels,Tribe, Food)%>%
 summarise(En=sum(Energy,na.rm=T))
AllDietCatEn2<-AllDietCatEn%>%
 group_by(Tribe,Food) %>%
 summarise(meanEN=round(mean(En, na.rm=TRUE),1),
        sdEn=round(sd(En, na.rm=TRUE),1),
        medEn=round(median(En,na.rm=T),1),
        IQREn=round(IQR(En, na.rm=TRUE),1),
        nEn=sum(!is.na(En)))

#shapiro wilk on Table 5.8 (wetenergy/category)
sCTFEn<-AllDietCatEn%>%
 group_by(Tribe, Food) %>%
 do (tidy(shapiro.test(.$En)))
pander(sCTFEn)
```

Tribe	Food	statistic	p.value	method
El Molo	Animal	0.9660865	4.183032e-01	Shapiro-Wilk normality test

El Molo	Fish	0.5322597	8.347759e-09	Shapiro-Wilk normality test
El Molo	C3 plant	0.9710039	5.470963e-01	Shapiro-Wilk normality test
El Molo	C4 plant	0.9485657	1.424947e-01	Shapiro-Wilk normality test
El Molo	Other	0.2290203	1.206168e-11	Shapiro-Wilk normality test
Turkana	Animal	0.7636147	2.613472e-06	Shapiro-Wilk normality test
Turkana	Fish	0.4926563	3.751322e-10	Shapiro-Wilk normality test
Turkana	C3 plant	0.8265884	4.680341e-05	Shapiro-Wilk normality test
Turkana	C4 plant	0.9086366	5.168108e-03	Shapiro-Wilk normality test
Turkana	Other	0.1553105	2.545362e-13	Shapiro-Wilk normality test
Luhya (Webuye)	Animal	0.9485064	1.307014e-01	Shapiro-Wilk normality test
Luhya (Webuye)	Fish	0.9061813	8.930659e-03	Shapiro-Wilk normality test
Luhya (Webuye)	C3 plant	0.9338230	5.009893e-02	Shapiro-Wilk normality test
Luhya (Webuye)	C4 plant	0.6674052	2.930122e-07	Shapiro-Wilk normality test
Luhya (Webuye)	Other	0.3408280	7.001248e-11	Shapiro-Wilk normality test
Luhya (Port Vict.)	Animal	0.9423418	9.580238e-02	Shapiro-Wilk normality test
Luhya (Port Vict.)	Fish	0.3821409	2.487758e-10	Shapiro-Wilk normality test
Luhya (Port Vict.)	C3 plant	0.9677442	4.592184e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	C4 plant	0.9000452	7.226828e-03	Shapiro-Wilk normality test
Luhya (Port Vict.)	Other	0.1791623	4.933319e-12	Shapiro-Wilk normality test
Luo (Port Vict.)	Animal	0.9713412	5.966036e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	Fish	0.4763222	4.452699e-09	Shapiro-Wilk normality test
Luo (Port Vict.)	C3 plant	0.9645106	4.220968e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	C4 plant	0.9409161	1.059652e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	Other	0.3815439	5.520019e-10	Shapiro-Wilk normality test

B.2.3.3 Contribution of Food Categories to Population Protein Intake

Following, the code for Fig. 5.7, Table 5.9, and respective Shapiro Wilks.

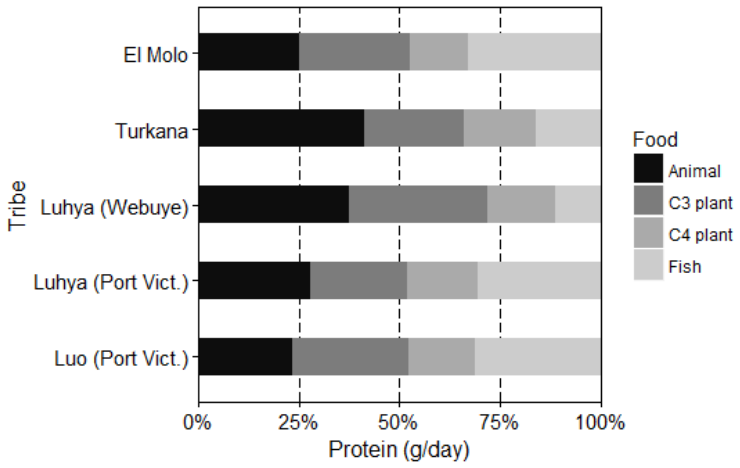
```
ggplot(AllDietOther, aes(x = Tribe, y = Protein, fill = Food)) +
 geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
 geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
 coord_flip()+
 my_theme+
 scale_fill_grey(start = 0.8, end = .05)+
 scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
 guides(fill = guide_legend(reverse = TRUE))+
 ylab("Protein (g/day)")

#protein
AllDietCatPr<-AllDiet %>%
 group_by(Row.Labels,Tribe, Food)%>%
 summarise(Pr=sum(Protein,na.rm=T))
AllDietCatPr2<-AllDietCatPr%>%
 group_by(Tribe,Food) %>%
```

```
 summarise(meanPr=round(mean(Pr, na.rm=TRUE),1),
        sdPr=round(sd(Pr, na.rm=TRUE),1),
        medPr=round(median(Pr,na.rm=T),1),
        IQRPr=round(IQR(Pr, na.rm=TRUE),1),
        nPr=sum(!is.na(Pr)))

#shapiro wilk on Table 5.9 (protein/category)
sCPr<-AllDietCatPr%>%
 group_by(Tribe, Food) %>%
 do (tidy(shapiro.test(.$Pr)))
pander(sCPr)
```

Tribe	Food	statistic	p.value	method
El Molo	Animal	0.9757567	6.876664e-01	Shapiro-Wilk normality test
El Molo	Fish	0.5322597	8.347759e-09	Shapiro-Wilk normality test
El Molo	C3 plant	0.9284639	3.982703e-02	Shapiro-Wilk normality test
El Molo	C4 plant	0.8591672	7.993189e-04	Shapiro-Wilk normality test
El Molo	Other	0.1810828	5.102491e-12	Shapiro-Wilk normality test
Turkana	Animal	0.7495997	1.461866e-06	Shapiro-Wilk normality test
Turkana	Fish	0.7651278	2.785934e-06	Shapiro-Wilk normality test
Turkana	C3 plant	0.8537176	1.924299e-04	Shapiro-Wilk normality test
Turkana	C4 plant	0.9141512	7.467027e-03	Shapiro-Wilk normality test
Turkana	Other	0.1553105	2.545362e-13	Shapiro-Wilk normality test
Luhya (Webuye)	Animal	0.9415858	8.304649e-02	Shapiro-Wilk normality test
Luhya (Webuye)	Fish	0.9061813	8.930659e-03	Shapiro-Wilk normality test
Luhya (Webuye)	C3 plant	0.9063249	9.007745e-03	Shapiro-Wilk normality test
Luhya (Webuye)	C4 plant	0.8186565	9.325698e-05	Shapiro-Wilk normality test
Luhya (Webuye)	Other	0.2762208	1.921935e-11	Shapiro-Wilk normality test
Luhya (Port Vict.)	Animal	0.9405749	8.559558e-02	Shapiro-Wilk normality test
Luhya (Port Vict.)	Fish	0.3821409	2.487758e-10	Shapiro-Wilk normality test
Luhya (Port Vict.)	C3 plant	0.9645966	3.838742e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	C4 plant	0.9478668	1.362903e-01	Shapiro-Wilk normality test
Luhya (Port Vict.)	Other	0.1759404	4.662604e-12	Shapiro-Wilk normality test
Luo (Port Vict.)	Animal	0.9768039	7.520382e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	Fish	0.4763222	4.452699e-09	Shapiro-Wilk normality test
Luo (Port Vict.)	C3 plant	0.9364212	8.082217e-02	Shapiro-Wilk normality test
Luo (Port Vict.)	C4 plant	0.9093714	1.654442e-02	Shapiro-Wilk normality test
Luo (Port Vict.)	Other	0.3815439	5.520019e-10	Shapiro-Wilk normality test

Considering the number of outliers identified in boxplots for protein intake, the stacked chart for protein intake was repeated without the outliers (EM20, TK03, EM06, EM21, TK22, TK23, TK35, TK37, PT18, PT25), with minimal changes.

```
AllDietOther<-filter(AllDiet,Food!="Other")
AllDietPout<-filter(AllDietOther,
                    Row.Labels!="EM20"&
                      Row.Labels!="TK03"&
                      Row.Labels!="EM06"&
                      Row.Labels!="EM21"&
                      Row.Labels!="TK22"&
                      Row.Labels!="TK23"&
                      Row.Labels!="TK35"&
                      Row.Labels!="TK37"&
                      Row.Labels!="PT18"&
                      Row.Labels!="PT25")
AllDietPout$Tribe<-factor(AllDietPout$Tribe,
                          levels=rev(unique(AllDietPout$Tribe)))
ggplot(AllDietPout, aes(x = Tribe, y = Protein, fill = Food)) +
  geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
  geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
  coord_flip()+
  my_theme+
  scale_fill_grey(start = 0.8, end = .05)+
  ylab("Protein (g/day)")+
  scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
  guides(fill = guide_legend(reverse = TRUE))
```



B.2.3.4 Comparison of Food Categories across Gender

Finally the, the code for Fig. 5.8, Table 5.10, and respective Shapiro Wilks.

```
ggplot(AllDietOther, aes(x = Sex, y = Energy, fill = Food)) +
  geom_hline(yintercept = c(0.25,0.50,0.75), linetype = "longdash")+
  geom_bar(stat = 'identity', position = 'fill', width = 0.5) +
  facet_grid(Tribe~.)+
  coord_flip()+
  ylab("Wet Energy (kcal/day)")+
  my_theme+
  scale_fill_grey(start = 0.8, end = .05)+
```

```
  scale_y_continuous(expand = c(0, 0), labels = scales::percent)+
  guides(fill = guide_legend(reverse = TRUE))

#wet energy by sex
AllDietCatEnS<-AllDiet %>%
  group_by(Row.Labels,Tribe, Food,Sex)%>%
  summarise(En=sum(Energy,na.rm=T))
AllDietCatEnS2<-AllDietCatEnS%>%
  group_by(Tribe,Food,Sex) %>%
  summarise(meanEN=round(mean(En, na.rm=TRUE),1),
            sdEn=round(sd(En, na.rm=TRUE),1),
            medEn=round(median(En,na.rm=T),1),
            IQREn=round(IQR(En, na.rm=TRUE),1),
            nEn=sum(!is.na(En)))

#shapiro wilk on Table 5.10 (wet energy/category+sex)
#I had to filter by Tribe first, otherwise returns "Error in shapiro.test(.$En) : all 'x' values are identical"
#tried this https://stackoverflow.com/questions/43748663/shapiro-test-plyr-all-x-values-are-identical, but didn't work
AllDietCatEl<-AllDietCatEnS%>%
  filter(Tribe=="El Molo")%>%
  group_by(Food,Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatEl)
```

Food	Sex	statistic	p.value	method
Animal	F	0.9616481	6.626407e-01	Shapiro-Wilk normality test
Animal	M	0.9598988	7.214277e-01	Shapiro-Wilk normality test
Fish	F	0.4949163	1.173073e-06	Shapiro-Wilk normality test
Fish	M	0.4212805	1.519034e-06	Shapiro-Wilk normality test
C3 plant	F	0.9374874	2.895893e-01	Shapiro-Wilk normality test
C3 plant	M	0.9129993	1.741990e-01	Shapiro-Wilk normality test
C4 plant	F	0.9061476	8.613947e-02	Shapiro-Wilk normality test
C4 plant	M	0.9557012	6.521381e-01	Shapiro-Wilk normality test
Other	F	0.2622232	2.166910e-08	Shapiro-Wilk normality test
Other	M	0.2968407	2.189847e-07	Shapiro-Wilk normality test

```
#Turkana returned same problem, so split it by food
AllDietCatTkA<-AllDietCatEnS%>%
  filter(Food=="Animal",Tribe=="Turkana")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatTkA)
```

Sex	statistic	p.value	method
F	0.8409431	0.004795083	Shapiro-Wilk normality test
M	0.8023093	0.001637654	Shapiro-Wilk normality test

```
AllDietCatTkF<-AllDietCatEnS%>%
  filter(Food=="Fish",Tribe=="Turkana")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatTkF)
```

Sex	statistic	p.value	method
F	0.7984808	1.086318e-03	Shapiro-Wilk normality test
M	0.5365558	1.692365e-06	Shapiro-Wilk normality test

```
AllDietCatTkC3<-AllDietCatEnS%>%
  filter(Food=="C3 plant",Tribe=="Turkana")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatTkC3)
```

Sex	statistic	p.value	method
F	0.7514776	0.0002445385	Shapiro-Wilk normality test
M	0.8997737	0.0569382227	Shapiro-Wilk normality test

```
AllDietCatTkC4<-AllDietCatEnS%>%
  filter(Food=="C4 plant",Tribe=="Turkana")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatTkC4)
```

Sex	statistic	p.value	method
F	0.8944767	0.03869325	Shapiro-Wilk normality test
M	0.9094066	0.08398647	Shapiro-Wilk normality test

```
#Luhya returned same problem, so split it by food
AllDietCatLhA<-AllDietCatEnS%>%
  filter(Food=="Animal",Tribe=="Luhya (Port Vict.)")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatLhA)
```

Sex	statistic	p.value	method
F	0.8903982	0.04703888	Shapiro-Wilk normality test
M	0.9503823	0.56665701	Shapiro-Wilk normality test

```
#fish returned same problem, so split it by sex
#that returned same problem bc everyone claimed to eat the same amount of fish, and shapiro.test cannot compute with that; but the data set is normal

AllDietCatLhC3<-AllDietCatEnS%>%
  filter(Food=="C3 plant",Tribe=="Luhya (Port Vict.)")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatLhC3)
```

Sex	statistic	p.value	method
F	0.9475363	0.4187398	Shapiro-Wilk normality test
M	0.9687827	0.8599562	Shapiro-Wilk normality test

```
AllDietCatLhC4<-AllDietCatEnS%>%
  filter(Food=="C4 plant",Tribe=="Luhya (Port Vict.)")%>%
  group_by(Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatLhC4)
```

Sex	statistic	p.value	method
F	0.8943290	0.05464017	Shapiro-Wilk normality test
M	0.8550211	0.02600374	Shapiro-Wilk normality test

```
AllDietCatLo<-AllDietCatEnS%>%
  filter(Tribe=="Luo (Port Vict.)")%>%
  group_by(Food,Sex) %>%
  do(tidy(shapiro.test(.$En)))
pander(AllDietCatLo)
```

Food	Sex	statistic	p.value	method
Animal	F	0.9368626	3.445542e-01	Shapiro-Wilk normality test
Animal	M	0.9432477	4.615117e-01	Shapiro-Wilk normality test
Fish	F	0.5190536	5.031645e-06	Shapiro-Wilk normality test
Fish	M	0.4419608	2.147306e-06	Shapiro-Wilk normality test
C3 plant	F	0.9145860	1.592023e-01	Shapiro-Wilk normality test
C3 plant	M	0.9598477	7.205862e-01	Shapiro-Wilk normality test
C4 plant	F	0.9275289	2.505535e-01	Shapiro-Wilk normality test
C4 plant	M	0.9335300	3.417139e-01	Shapiro-Wilk normality test
Other	F	0.5247644	5.610872e-06	Shapiro-Wilk normality test
Other	M	0.6388460	9.202771e-05	Shapiro-Wilk normality test

```
## R version 3.3.1 (2016-06-21)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 14393)
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
```

```
## other attached packages:
## [1] gmodels_2.16.2 broom_0.4.1    pander_0.6.0   tidyr_0.6.0
## [5] dplyr_0.5.0    ggplot2_2.2.1  MASS_7.3-45
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.7      knitr_1.16       magrittr_1.5     mnormt_1.5-5
##  [5] munsell_0.4.3    lattice_0.20-33  colorspace_1.2-7 R6_2.2.0
##  [9] stringr_1.2.0    plyr_1.8.4       tools_3.3.1      parallel_3.3.1
## [13] grid_3.3.1       nlme_3.1-128     gtable_0.2.0     psych_1.6.9
## [17] DBI_0.5-1        gtools_3.5.0     htmltools_0.3.6  yaml_2.1.14
## [21] lazyeval_0.2.0   rprojroot_1.2    digest_0.6.10    assertthat_0.1
## [25] tibble_1.2       reshape2_1.4.2   evaluate_0.10    rmarkdown_1.6
## [29] labeling_0.3     gdata_2.17.0     stringi_1.1.2    scales_0.4.1
## [33] backports_1.0.5  foreign_0.8-66
```

APPENDIX C
DMTA DATA ANALYSIS

C.1 DMTA SCANS FROM LIVING SUBJECTS

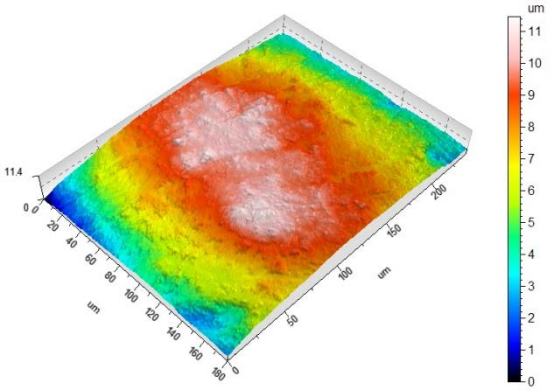


Figure C.1 Three-dimensional representation of microwear surface (EM01, M_1, f9).

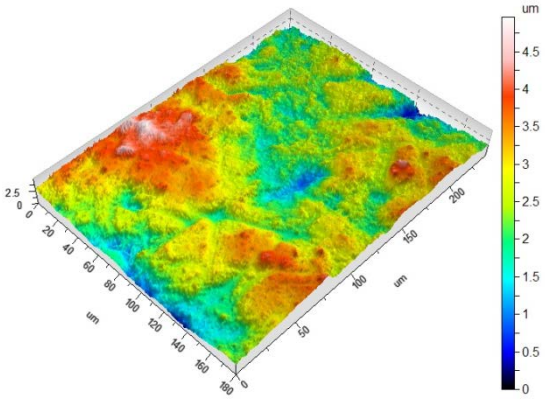


Figure C.2 Three-dimensional representation of microwear surface (EM05, M_1, f9).

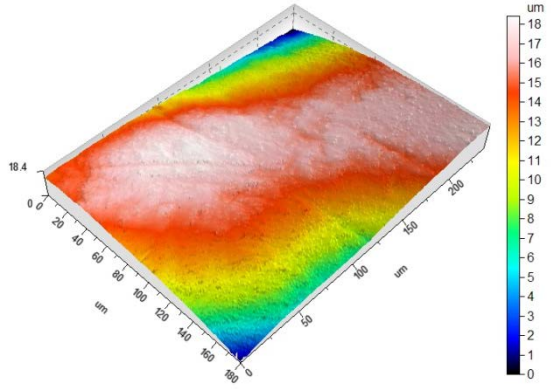


Figure C.3 Three-dimensional representation of microwear surface (EM08, M_1, f9).

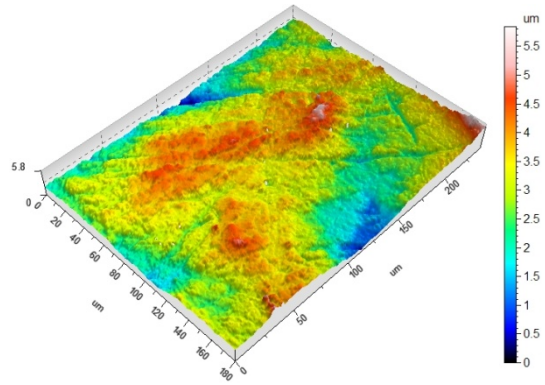


Figure C.4 Three-dimensional representation of microwear surface (EM10, M_1, fx).

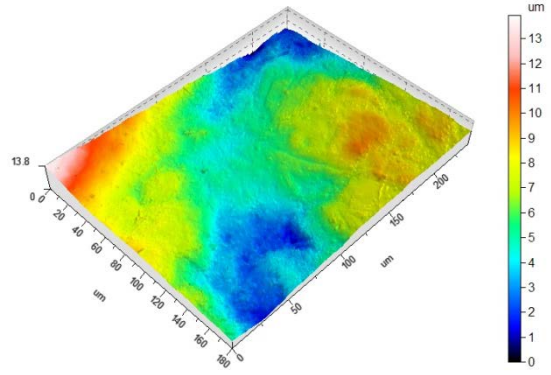


Figure C.5 Three-dimensional representation of microwear surface (EM14, M_3, fx).

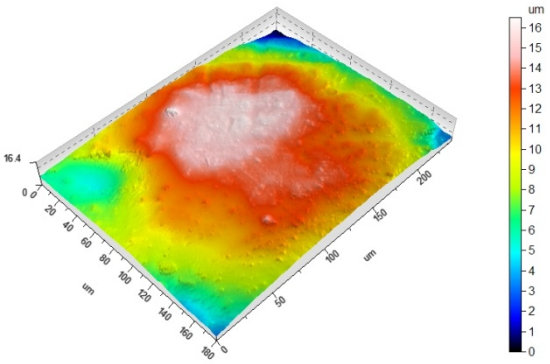


Figure C.6 Three-dimensional representation of microwear surface (EM18, M_1, f10n).

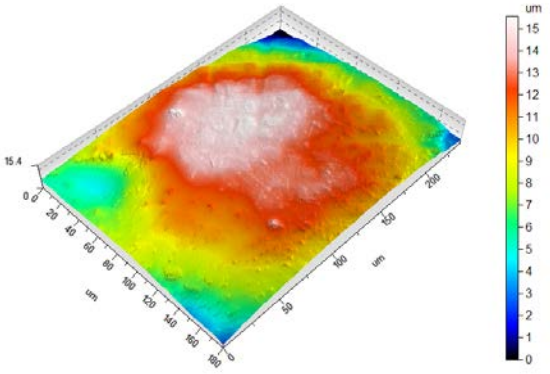


Figure C.7 Three-dimensional representation of microwear surface (EM19, M_3, f10n).

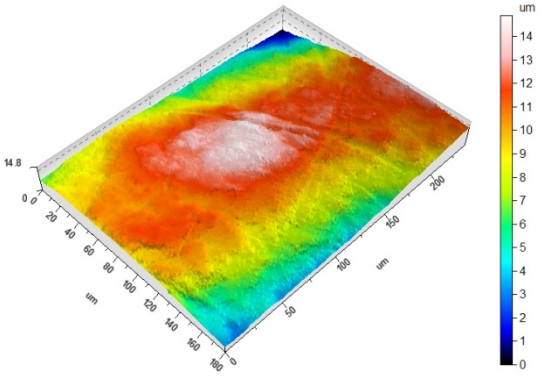


Figure C.8 Three-dimensional representation of microwear surface (EM24, M_1, f9).

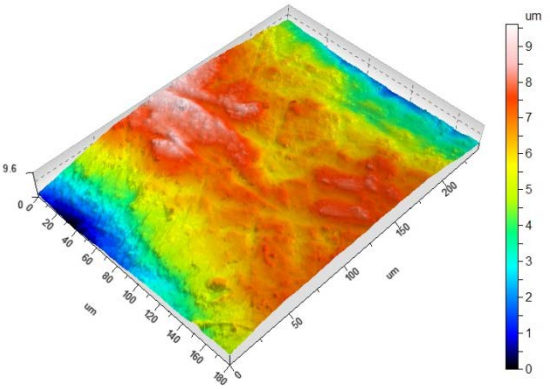


Figure C.9 Three-dimensional representation of microwear surface (EM25, M_1, fx).

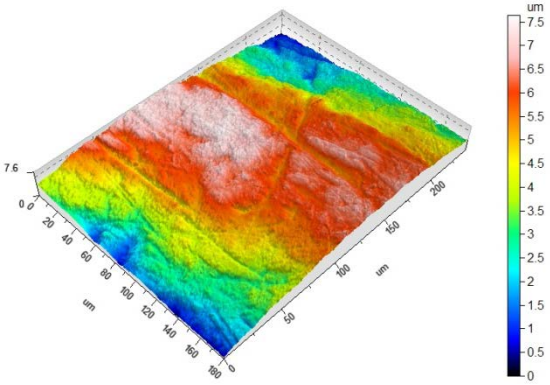


Figure C.10 Three-dimensional representation of microwear surface (EM31, M_1, fx).

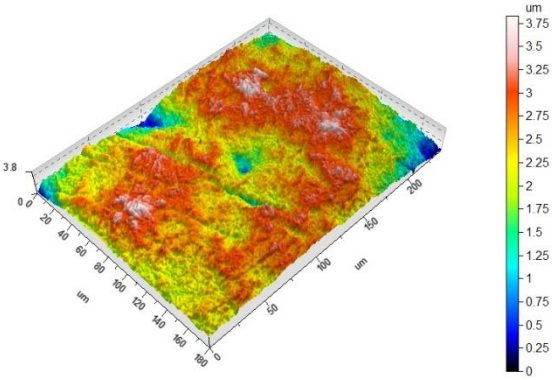


Figure C.11 Three-dimensional representation of microwear surface (TK02, M_1, f10n).

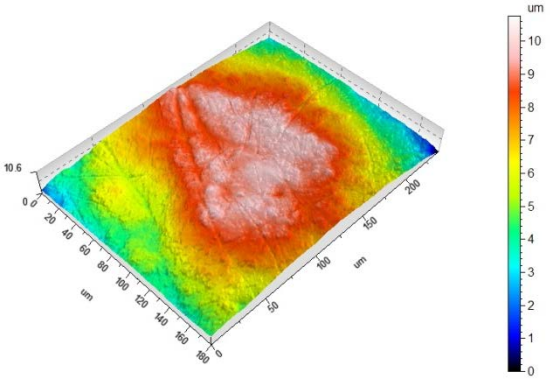


Figure C.12 Three-dimensional representation of microwear surface (TK03, M_1, f9).

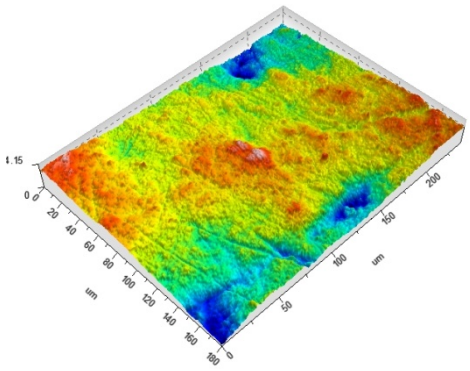


Figure C.13 Three-dimensional representation of microwear surface (TK09, M_1, f9).

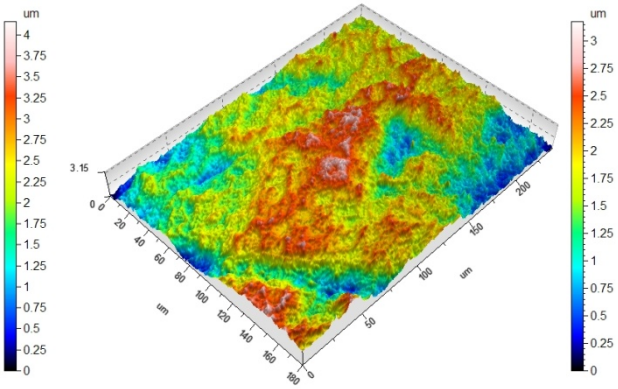


Figure C.14 Three-dimensional representation of microwear surface (TK16, M_1, f10n).

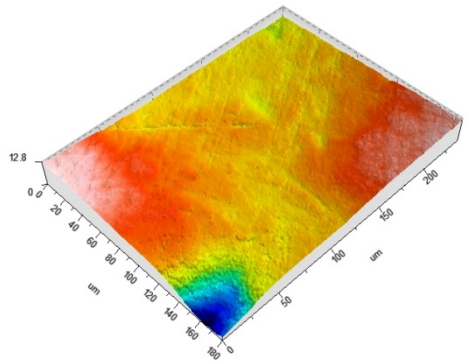


Figure C.15 Three-dimensional representation of microwear surface (TK19, M_1, f9).

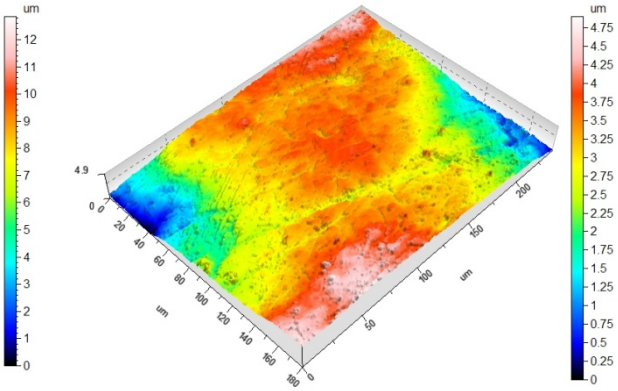


Figure C.16 Three-dimensional representation of microwear surface (TK29, M_1, f9).

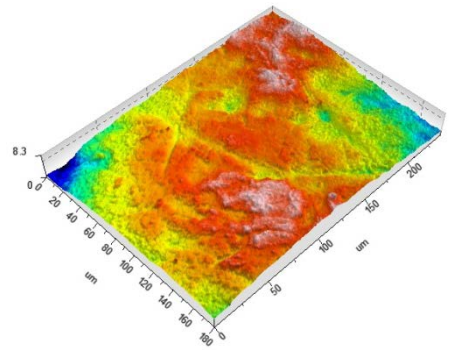


Figure C.17 Three-dimensional representation of microwear surface (TK35, M_1, f10n).

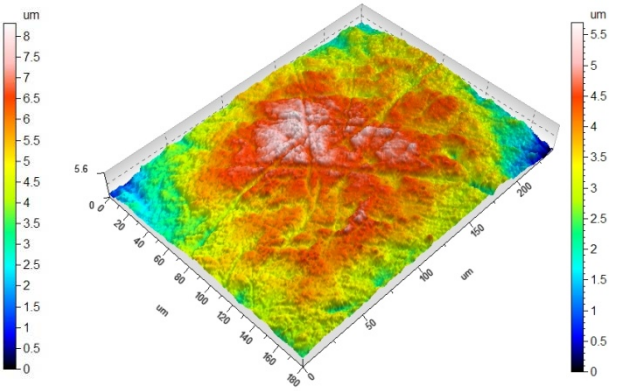


Figure C.18 Three-dimensional representation of microwear surface (WB01, M_1, f9).

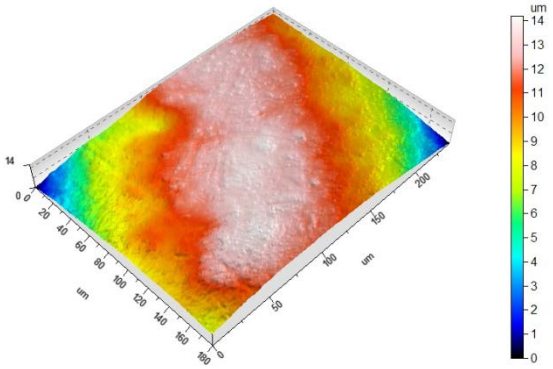


Figure C.19 Three-dimensional representation of microwear surface (WB02, M_1, f9).

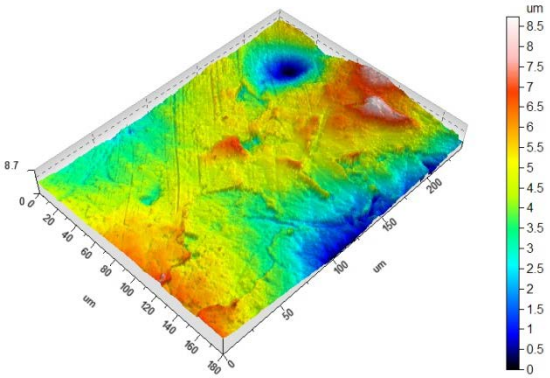


Figure C.20 Three-dimensional representation of microwear surface (WB05, M_1, f10n).

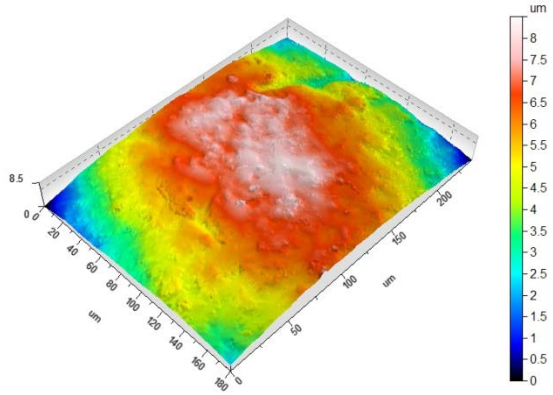


Figure C.21 Three-dimensional representation of microwear surface (WB23, M_1, f9).

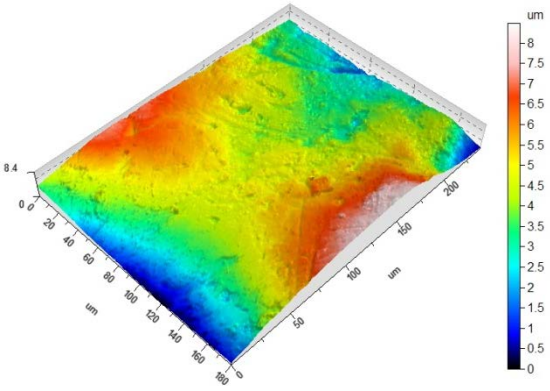


Figure C.22 Three-dimensional representation of microwear surface (WB28, M_1, f10n).

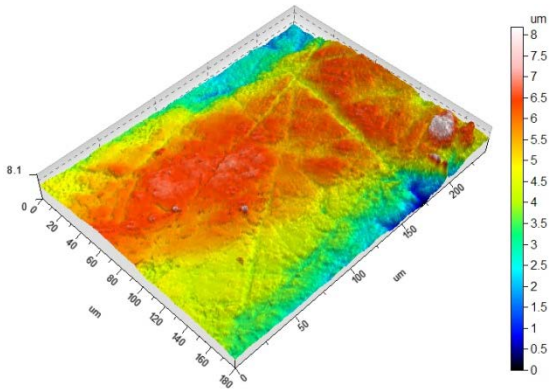


Figure C.23 Three-dimensional representation of microwear surface (WB29, M_1, f10n).

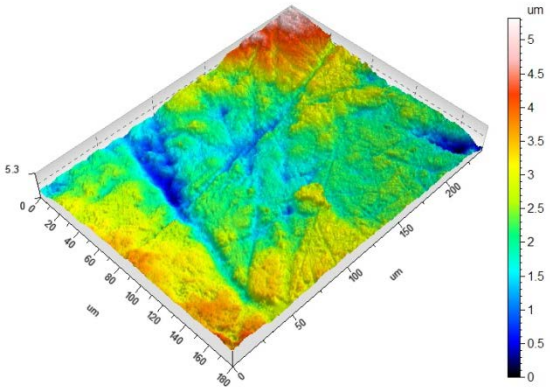


Figure C.24 Three-dimensional representation of microwear surface (WB30, M_1, f9).

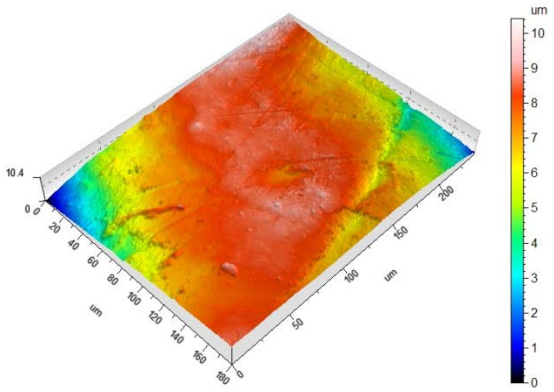


Figure C.25 Three-dimensional representation of microwear surface (PT02, M_1, fx).

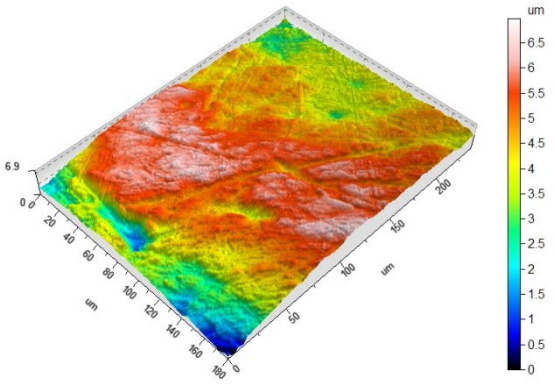


Figure C.26 Three-dimensional representation of microwear surface (PT03, M_2, f9).

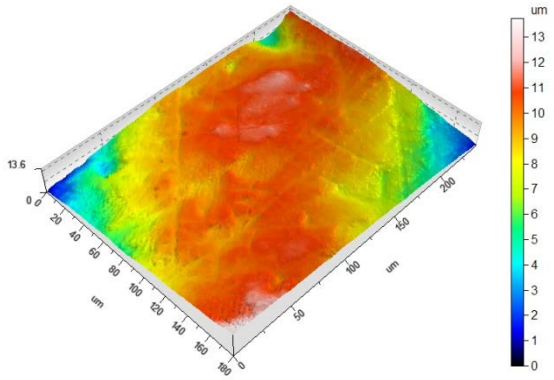


Figure C.27 Three-dimensional representation of microwear surface (PT26, L M_1 9).

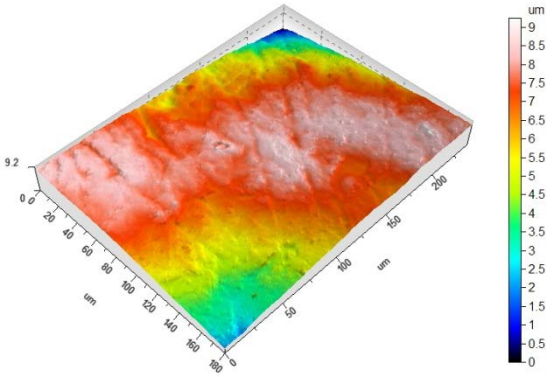


Figure C.28 Three-dimensional representation of microwear surface (PT29, M_1, fx).

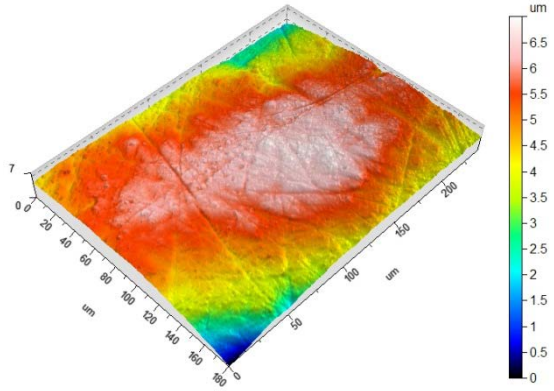


Figure C.29 Three-dimensional representation of microwear surface (PT45, M_2, f9).

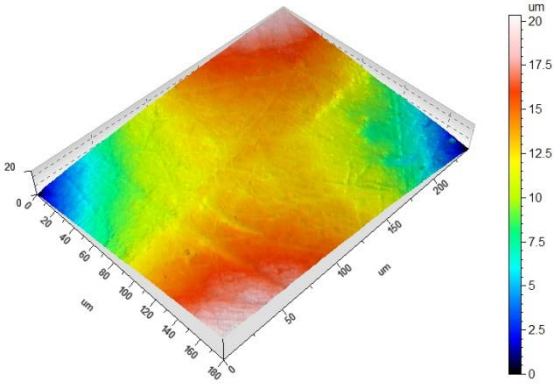


Figure C.30 Three-dimensional representation of microwear surface (PT48, M_1, f9).

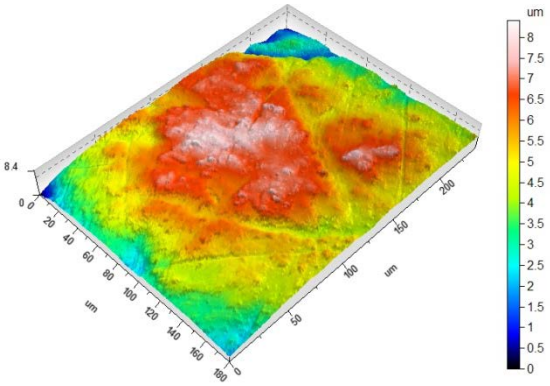


Figure C.31 Three-dimensional representation of microwear surface (PT10, M_1, fx).

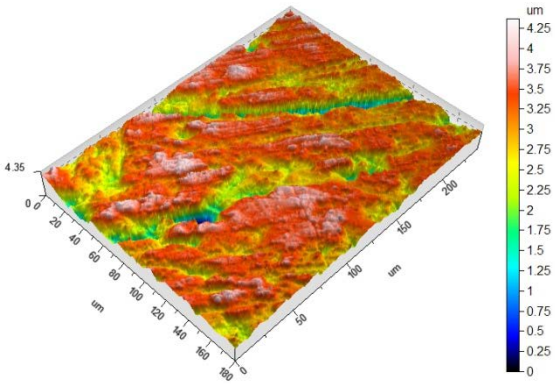


Figure C.32 Three-dimensional representation of microwear surface (PT25, M_1, f10n).

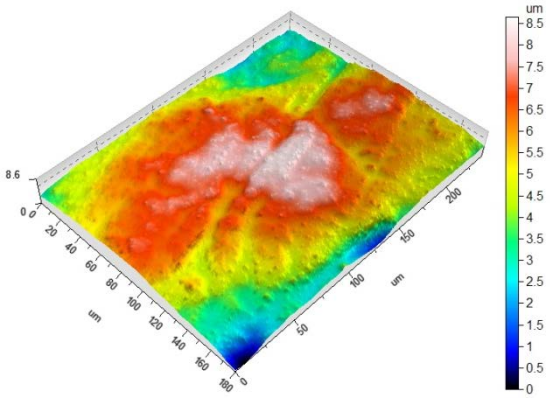


Figure C.33 Three-dimensional representation of microwear surface (PT27, M_1, f10n).

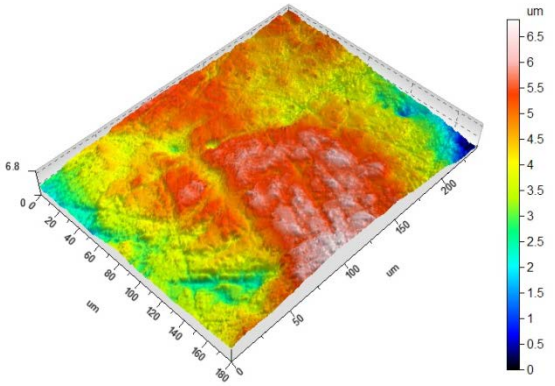


Figure C.34 Three-dimensional representation of microwear surface (PT32, M_1, f10n).

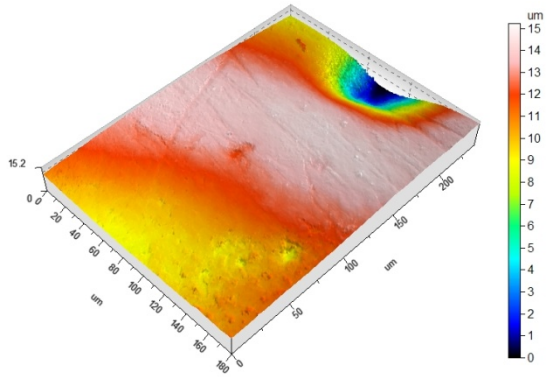


Figure C.35 Three-dimensional representation of microwear surface (PT50, M_2, f9).

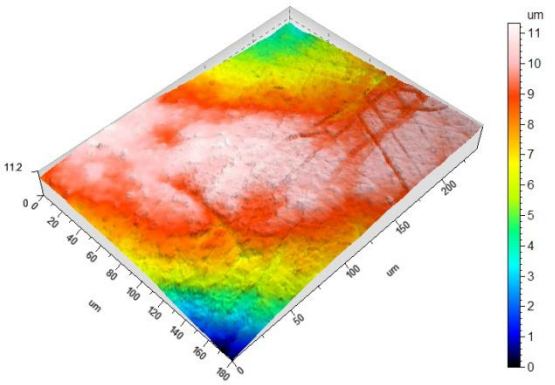


Figure C.36 Three-dimensional representation of microwear surface (PT55, M_1, f9).

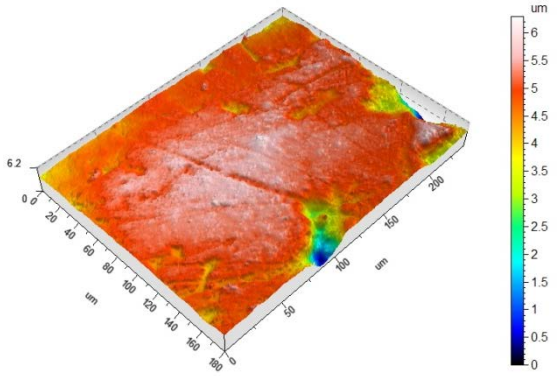


Figure C.37 Three-dimensional representation of microwear surface (PT61, M_1, fx).

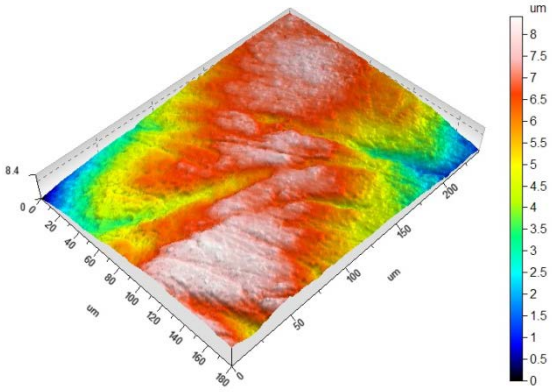


Figure C.38 Three-dimensional representation of microwear surface (PT62, M_1, fx).

C.2 DMTA STATISTICAL ANALYSES

This section includes the R code that produced graphical and statistical analysis on DMTA parameters. In the cases were the analysis themselves were not fully reported in text (e.g. Shapiro Wilk tests) the output is included here. Otherwise, only the R code is present and, if applicable, the tables or figures that report the results are mentioned.

The dataset used is in Table 6.2, whereas the functions used were:

```
#setting decimals
fmt_decimals <- function(decimals=0){
   function(x) format(x,nsmall = decimals,scientific = FALSE)
}
#graphical settings for ggplot
my_theme <- theme(axis.text=element_text(size=11,
                                colour="black"),
      axis.ticks=element_line (size=0.5,colour="black"),
      axis.title=element_text(size=12),
      panel.grid.minor = element_blank(),
      panel.background = element_blank(),
      panel.grid.major = element_blank(),
      panel.border = element_rect(colour = "black",
                           fill=NA, size=0.5))
#calculate outliers
is_outlier <- function(x) {
  return(x < quantile(x, 0.25,na.rm=TRUE) - 1.5 * IQR(x,na.rm=TRUE) | x > quantile(x, 0.75,na.r
m=TRUE) + 1.5 * IQR(x,na.rm=TRUE))
}
#label boxplots with `n=`
give.n <- function(x){
  return(data.frame(y = median(x)*1.05,
                label = paste0("n=",length(x))))
  #change multiplier to find the perfect position
}
#function to calculate omega squared
omega_sq <- function(lm){
```

```
  sum_stats <- summary.aov(lm)[[1]]# made few changes to take "lm" objects
  SSm <- sum_stats[["Sum Sq"]][1]
  SSr <- sum_stats[["Sum Sq"]][2]
  DFm <- sum_stats[["Df"]][1]
  MSr <- sum_stats[["Mean Sq"]][2]
  W2 <- (SSm-DFm*MSr)/(SSm+SSr+MSr)
  return(W2)
}
```

C.2.1 EXPLORING THE DATA AND CHOOSING A TEST

First, the data (i.e. each DMTA parameter) was summarised by calculating classic and robust measures of central tendency and variability: means and standard deviations; and medians and interquartile ranges (Table 6.3).

```
#mean and sd by population
MW2<-MW %>%
 group_by(Tribe)%>%
 summarise(meanAsfc=round(mean(Asfc, na.rm=TRUE),2),
        sdAsfc=round(sd(Asfc, na.rm=TRUE),2),
        medAsfc=round(median(Asfc, na.rm=TRUE),2),
        IQRAsfc=round(IQR(Asfc, na.rm=TRUE),2),
        nAsfc=sum(!is.na(Asfc)),
        meanepLsar=round(mean(epLsar, na.rm=TRUE),4),
        sdepLsar=round(sd(epLsar, na.rm=TRUE),4),
        medepLsar=round(median(epLsar, na.rm=TRUE),4),
        IQRepLsar=round(IQR(epLsar, na.rm=TRUE),4),
        nepLsar=sum(!is.na(epLsar)),
        meanHAsfc=round(mean(HAsfc, na.rm=TRUE),3),
        sdHAsfc=round(sd(HAsfc,na.rm=TRUE),3),
        medHAsfc=round(median(HAsfc, na.rm=TRUE),3),
        IQRHAsfc=round(IQR(HAsfc, na.rm=TRUE),3),
        nHAsfc=sum(!is.na(HAsfc)),
        meanTfv=round(mean(Tfv, na.rm=TRUE),1),
        sdTfv=round(sd(Tfv, na.rm=TRUE),1),
        medTfv=round(median(Tfv, na.rm=TRUE),1),
        IQRTfv=round(IQR(Tfv, na.rm=TRUE),1),
        nTfv=sum(!is.na(Tfv))
        )
pander(MW2)
```

Next, the data was visualised in dotplots, with outliers labelled based on IQR (Figure 6.3). Boxplots were not used due to the low *n*'s.

```
MW%>%
 group_by(Tribe) %>%
 mutate(outlier = as.character(ifelse(is_outlier(Asfc),
as.character(Code),as.numeric(NA)))) %>%
 ggplot(., aes(x=factor(Tribe,levels=unique(MW$Tribe)),
           y = Asfc)) +
  geom_dotplot(binaxis='y', stackdir='center', dotsize=0.7,fill="white")+
 stat_summary(fun.data = give.n, geom = "text",size=4)+
```

```
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  my_theme+
  scale_y_continuous(name=expression(Asfc),
                     labels=fmt_decimals(1))+
 scale_x_discrete(name="Tribe",
                  labels=c("El Molo","Turkana","Luhya (Webuye)",
                           "Luhya (Port Victoria)", "Luo (Port Victoria)"))

MW%>%
 group_by(Tribe) %>%
 mutate(outlier = as.character(ifelse(is_outlier(Smc),                    as.character(Code),as.
numeric(NA)))) %>%
 ggplot(., aes(x=factor(Tribe,levels=unique(MW$Tribe)),
               y = Smc)) +
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
 geom_dotplot(binaxis='y', stackdir='center', dotsize=0.7,fill="white")+
  my_theme+
  scale_y_continuous(name=expression(Smc),
                     labels=fmt_decimals(1))+
 scale_x_discrete(name="Tribe",
                  labels=c("El Molo","Turkana","Luhya (Webuye)",
                           "Luhya (Port Victoria)", "Luo (Port Victoria)"))

MW%>%
 group_by(Tribe) %>%
 mutate(outlier = as.character(ifelse(is_outlier(epLsar),                    as.character(Code),
as.numeric(NA)))) %>%
 ggplot(., aes(x=factor(Tribe,levels=unique(MW$Tribe)),
               y = epLsar)) +
  geom_dotplot(binaxis='y', stackdir='center', dotsize=0.7,fill="white")+
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  my_theme+
  scale_y_continuous(name=expression(epLsar),
                     labels=fmt_decimals(1))+
 scale_x_discrete(name="Tribe",
                  labels=c("El Molo","Turkana","Luhya (Webuye)",
                           "Luhya (Port Victoria)", "Luo (Port Victoria)"))

MW%>%
 group_by(Tribe) %>%
 mutate(outlier = as.character(ifelse(is_outlier(Tfv),                    as.character(Code),as.
numeric(NA)))) %>%
 ggplot(., aes(x=factor(Tribe,levels=unique(MW$Tribe)),
               y = Tfv)) +
  geom_dotplot(binaxis='y', stackdir='center', dotsize=0.7,fill="white")+
  geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
  my_theme+
  scale_y_continuous(name=expression(Tfv),
                     labels=fmt_decimals(1))+
```

```
  scale_x_discrete(name="Tribe",
                   labels=c("El Molo","Turkana","Luhya (Webuye)",
                            "Luhya (Port Victoria)", "Luo (Port Victoria)"))

MW%>%
  group_by(Tribe) %>%
  mutate(outlier = as.character(ifelse(is_outlier(HAsfc),                    as.character(Code),
as.numeric(NA)))) %>%
  ggplot(., aes(x=factor(Tribe,levels=unique(MW$Tribe)),
                y = HAsfc)) +
    geom_dotplot(binaxis='y', stackdir='center', dotsize=0.7,fill="white")+
    geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=4)+
    my_theme+
    scale_y_continuous(name=expression(HAsfc),
                       labels=fmt_decimals(1))+
  scale_x_discrete(name="Tribe",
                   labels=c("El Molo","Turkana","Luhya (Webuye)",
                            "Luhya (Port Victoria)", "Luo (Port Victoria)"))
```

Now, considering that each microwear parameter describes different attributes of the wear surface, then they may be treated as independent variables (Scott et al., 2006, p. 345).
Thus, since the aim is to compare the different groups in relation to each microwear parameter, the adequate test is ANOVA. However, prior to conducting this statistical test, one must check if the data meets ANOVA assumptions.

C.2.2. ANOVA ASSUMPTIONS

In *italics* the ones that may cause problems (according to Field et al., 2012, p.412)
a. Independence
b. Homogeneity of variances
c. *Normality* of the residuals (consider this; Field et al. (2012) actually addresses normality within groups; here both are included)

ADDITIONAL COMMENTS on ASSUMPTIONS
a. *Unequal sample sizes*: when group sizes are equal and bigger than 10, the F-statistic is robust to violations of normality; and a sample sizes are unequal if the largest sample is more than 50% bigger than the smallest (1).
b. *Outliers*: partially addressed in the normality assumption

C.2.2.1 Normality

Shapiro Wilk tests were used to check for normality within each parameter. However, this test is very sensitive to outliers (Field et al., 2012, p. 182). In addition, also consider that in big samples, it is easy for Shapiro Wilk to get significant results (Field et al., 2012, p.182), whereas in small samples Shapiro Wilk has low power (Field et al., 2012, p. 193) (Table 6.4).

```
sAsfc<-MW %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Asfc)))
pander(sAsfc)
```

Tribe	statistic	p.value	method
El Molo	0.9337369	0.48563701	Shapiro-Wilk normality test
Turkana	0.7695046	0.02019359	Shapiro-Wilk normality test

Luhya (Webuye)	0.9015221	0.34029803	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9442405	0.69351718	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9101976	0.35545702	Shapiro-Wilk normality test

```
sepLsar<-MW %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$epLsar)))
pander(sepLsar)
```

Tribe	statistic	p.value	method
El Molo	0.9383603	0.53491708	Shapiro-Wilk normality test
Turkana	0.8393100	0.09789661	Shapiro-Wilk normality test
Luhya (Webuye)	0.9500320	0.72999059	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.8585468	0.18423617	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9153743	0.39346490	Shapiro-Wilk normality test

```
sHAsfc<-MW %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$HAsfc)))
pander(sHAsfc)
```

Tribe	statistic	p.value	method
El Molo	0.9139502	0.309241821	Shapiro-Wilk normality test
Turkana	0.7387435	0.009663937	Shapiro-Wilk normality test
Luhya (Webuye)	0.9868524	0.985896994	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9323291	0.598205128	Shapiro-Wilk normality test
Luo (Port Vict.)	0.7144480	0.003277943	Shapiro-Wilk normality test

```
sTfv<-MW %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Tfv)))
pander(sTfv)
```

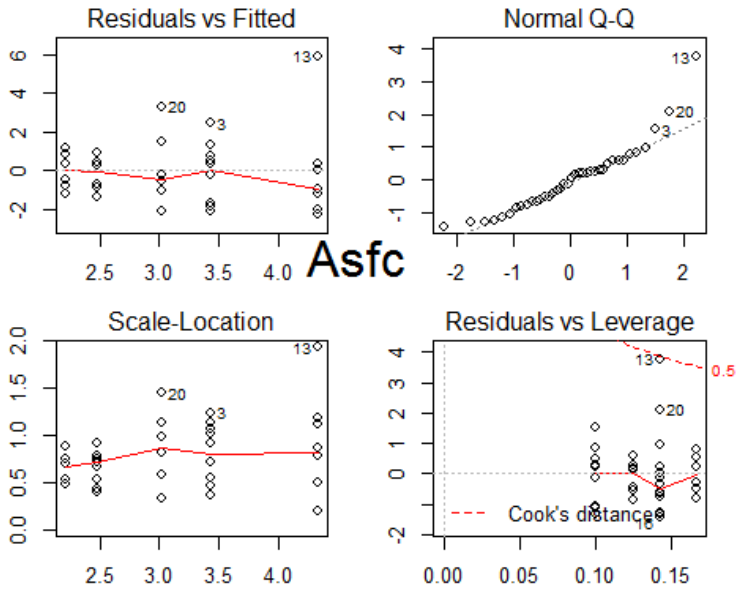
Tribe	statistic	p.value	method
El Molo	0.9145362	0.31362534	Shapiro-Wilk normality test
Turkana	0.9312840	0.56187242	Shapiro-Wilk normality test
Luhya (Webuye)	0.9117179	0.40789746	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.8513254	0.16133193	Shapiro-Wilk normality test
Luo (Port Vict.)	0.8139403	0.04027997	Shapiro-Wilk normality test

As expected, Shapiro Wilk tests indicates that all microwear parameters break the normality assumption, except for *epLsar*. These tests are important because standard deviations do not accurately report the variability of the sample if this sample is non normal. However, in the case of ANOVA tests, the normality of residuals seems to be more important.

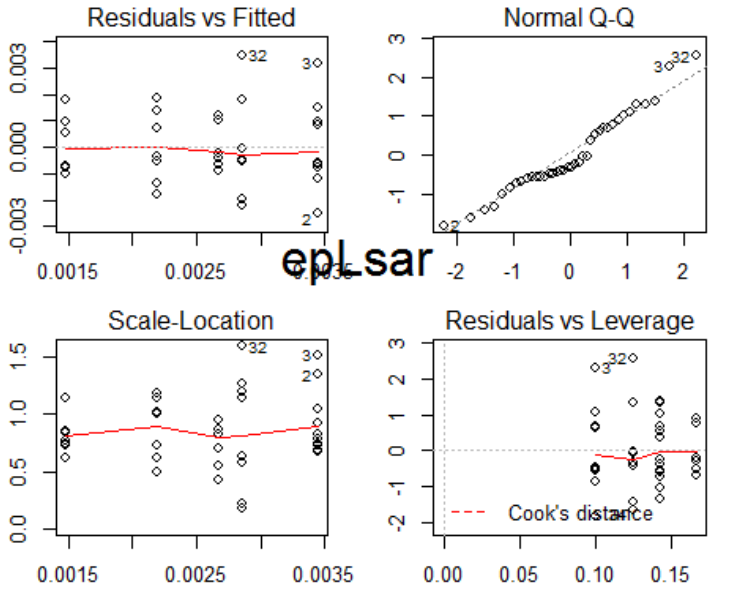
To check residuals assumptions, one may use the plots below, in which:
The *Residuals vs. Fitted* plot shows if residuals have non-linear patterns (and hence if they violate the assumption of linearity). If you find equally spread residuals around a horizontal line without distinct patterns, that is a good indication you do not have non-linear relationships.
The *Normal Q-Q* plot shows if residuals are normally distributed. Do residuals follow a straight line well or do they deviate severely?
The *Scale-Location* shows if residuals are spread equally along the ranges of predictors. This is how you can check the assumption of equal variance (homoscedasticity). It's good if you see a horizontal line with equally (randomly) spread points.
The *Residuals vs. Leverage* plot helps us to find influential cases (i.e., subjects) if any. This time patterns are not relevant. We watch out for outlying values at the upper right corner or at the

lower right corner, outside of a dashed line, Cook's distance. When cases are outside (meaning they have high Cook's distance scores), the cases are influential to the test results.

```
aAsfc<-aov(MW$Asfc~MW$Tribe,data=MW)
par(mfrow=c(2,2), mai=c(0.5,0.5,0.2,0.2)) #c(bottom, left, top, right) in inches
plot(aAsfc)
mtext("Asfc", side=3, outer=T, line=-10.9, cex=2)
```

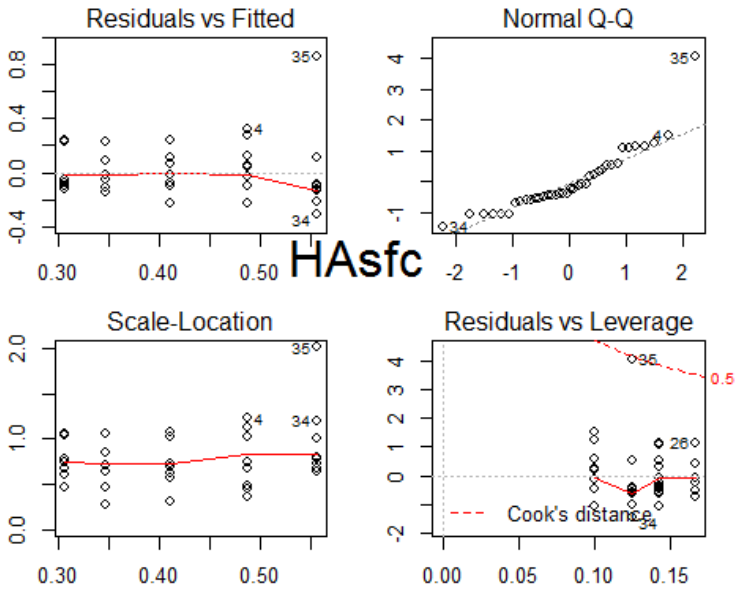



```
aepLsar<-aov(MW$epLsar~MW$Tribe,data=MW)
par(mfrow=c(2,2), mai=c(0.5,0.5,0.2,0.2))
plot(aepLsar)
mtext("epLsar", side=3, outer=T, line=-10.9, cex=2)
```

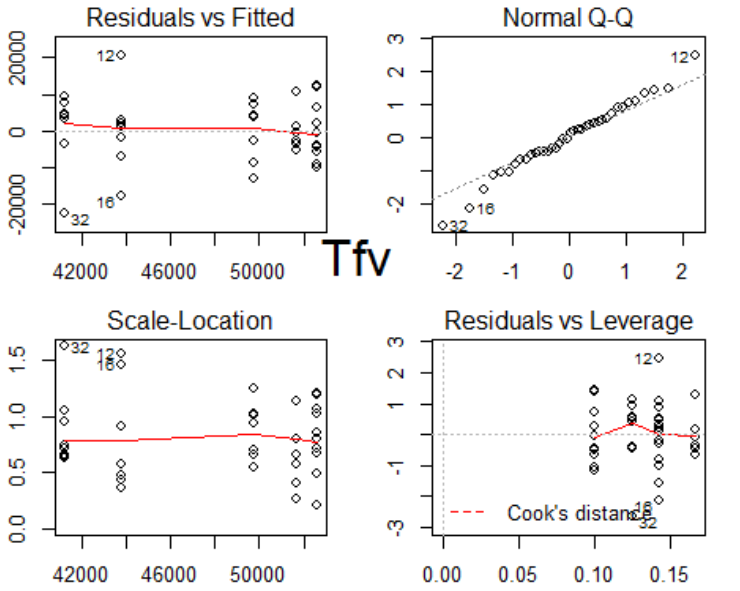



```
aHAsfc<-aov(MW$HAsfc~MW$Tribe,data=MW)
par(mfrow=c(2,2), mai=c(0.5,0.5,0.2,0.2))
```

```
plot(aHAsfc)
mtext("HAsfc", side=3, outer=T, line=-10.9, cex=2)
```




```
aTfv<-aov(MW$Tfv~MW$Tribe,data=MW)
par(mfrow=c(2,2), mai=c(0.5,0.5,0.2,0.2))
plot(aTfv)
mtext("Tfv", side=3, outer=T, line=-10.9, cex=2)
```



From the above graphs, it becomes clear that:
For *Asfc*, EM08 (3), TK09 (13), and WB05 (20) are outliers. Interestingly, EM08 is not identified as an outlier using IQR (see plots above), while the other two are. In addition, the QQ plot indicates that the residuals are non-normal, while the scale-location looks homoscedastic.
For *epLsar*, EM05 (2), and PT25 (32) are outliers (but only PT25 identified as outlier using IQR), but none reach the 0.5 Cook's distance, so none is particularly influential. The QQ and scale-location plot indicate that the residuals of this parameter is normal and homoscedastic.

For *HAsfc*, PT50 (35), seems to be the most influential case, with a Cook's distance of 0.5. Here, the residuals seem non-normal, but homoscedastic, as for the first parameter.
For *Tfv*, TK03 (12), TK 29(16), PT25 (32) are outliers (besides these IQR also identifies PT45 as outlier). Here, once again, the residuals seem non-normal, but homoscedastic.

In conclusion, the analysis of residuals confirms the results of the Shapiro-Wilk tests conducted within the samples.

C.2.2.2 Homoscedasticity

Despite checking for the homoscedasticity of residuals above, one may also check for the equality of variances within each parameter, using Levene tests.

```
lAsfc<-MW %>%
  do(tidy(leveneTest(.$Asfc~.$Tribe)))
pander(lAsfc)
```

term	df	statistic	p.value
group	4	0.6656	0.6204
	33	NA	NA

```
lepLsar<-MW %>%
  do(tidy(leveneTest(.$epLsar~.$Tribe)))
pander(lepLsar)
```

term	df	statistic	p.value
group	4	0.4197	0.7932
	33	NA	NA

```
lHAsfc<-MW %>%
  do(tidy(leveneTest(.$HAsfc~.$Tribe)))
pander(lHAsfc)
```

term	df	statistic	p.value
group	4	0.4093	0.8006
	33	NA	NA

```
lTfv<-MW %>%
  do(tidy(leveneTest(.$Tfv~.$Tribe)))
pander(lTfv)
```

term	df	statistic	p.value
group	4	0.2473	0.9092
	33	NA	NA

Levene tests confirm that all parameters are homoscedastic, as suggested by the analysis of residuals. Considering that several of the group sample sizes were inferior to 10, ROBUST ANOVAs were used for *Asfc*, *HAsfc*, and *Tfv*, due to lack normality and to deal with the presence of outliers, while normal ANOVA was used for *epLsar*.

Now, common ways to rectify problems with assumptions is to transform all of the data and reanalyse transformed values. This, however, changes the hypothesis being test. Instead, with distributional problems, one may use robust methods, which use bootstrapping [where inference about a population from sample data (sample → population) can be modelled by resampling the sample data and performing inference on (resample → sample)] and trimmed means or M-estimators (Field et al., 2012, p. 414).

C.2.3. CONDUCTING ANOVA

In this case, robust ANOVA's were conducted using a 5% trimmed mean and 999 bootstrap samples. From Field et al. (2012, p. 414): "A trimmed mean is simply a mean based on the distribution of scores after some percentage [in the case 5%] of scores has been removed from each extreme of the distribution.". In other words, 10% of the results were excluded from the analysis, which further decreased the sample size. On the other hand, "bootstrapping estimates the properties of the sampling distribution from the sample data. In effect, the sample data are treated as a population, from which smaller samples are taken (with reposition). Then, the statistic of interest [the mean, in this case] is calculated in each sample, and by taking many samples [999 here], the sampling distribution can be estimated. The standard error of the statistic is estimated from the standard deviation of this sampling distribution created from the bootstrap samples. From this standard error, confidence intervals, and significance tests can be computed."

Note that if you use 1000 or 2000, usually one bootstrap estimate fails (returns, for instance, "effective number of bootstrap samples was 1999", and warning "Some bootstrap estimates of the test statistic could not be computed"). For that reason, I used 999 bootstrap samples, although the results are similar, independently of the number of bootstrap samples.

```
raAsfc<-t1waybt(MW$Asfc~MW$Tribe,tr=0.05,nboot=999)
raAsfc

## Call:
## t1waybt(formula = MW$Asfc ~ MW$Tribe, tr = 0.05, nboot = 999)
## Effective number of bootstrap samples was 999.
## Test statistic: 1.5872
## p-value: 0.36837
## Variance explained 0.345
## Effect size 0.587

aepLsar<-lm(MW$epLsar~MW$Tribe,data=MW)
summary.aov(aepLsar)

##             Df    Sum Sq   Mean Sq F value Pr(>F)
## MW$Tribe     4 1.771e-05 4.427e-06   2.079  0.106
## Residuals   33 7.026e-05 2.129e-06

raHAsfc<-t1waybt(MW$HAsfc~MW$Tribe,tr=0.05,nboot=999)
raHAsfc

## Call:
## t1waybt(formula = MW$HAsfc ~ MW$Tribe, tr = 0.05, nboot = 999)
## Effective number of bootstrap samples was 999.
## Test statistic: 1.4889
## p-value: 0.32332
## Variance explained 0.352
## Effect size 0.593

raTfv<-t1waybt(MW$Tfv~MW$Tribe,tr=0.05,nboot=999)
raTfv
```

```
## Call:
## t1waybt(formula = MW$Tfv ~ MW$Tribe, tr = 0.05, nboot = 999)
## Effective number of bootstrap samples was 999.
## Test statistic: 2.0594
## p-value: 0.21522
## Variance explained 0.294
## Effect size 0.543
```

In addition to reporting ANOVA results, one must include effect sizes.
The best measure of effect sizes for standard ANOVA's is the omega squared, which is an unbiased version of the eta squared (this is particularly important for small sample sizes). Usually speaking, an omega squared of 0.01, 0.06 and 0.14 represent small, medium, and large effect sizes, respectively. The omega squared is essentially an unbiased estimate of Pearson's correlation coefficient *r*, although it usually leads to a slightly lower estimate. Below is the code for both omega squared and omega for *epLsar*.

```
omega_sq(aepLsar)
sqrt(omega_sq(aepLsar))
```

Field et al. (2012, p. 57) claims that the Pearson's correlation coefficient (and similar) are preferable to Cohen's effect sizes because they vary between 0 and 1. However, there are situations in which Cohen's *d*, another very common effect size statistic, may be favoured (e.g. when sample sizes are very discrepant).
In addition, the Robust ANOVA's provide a robust explanatory measure of effect size (ξ), which is based on Cohen's *d* (see Wilcox, 2012, p. 166, 295 - note that t1waybt$effectsize uses t1wayv2). More exactly, this effect size takes into account uses the trimming and bootstrapping used. It is noted that under normality and homoscedasticity, Cohen's *d* 0.2, 0.5, and 0.8 roughly correspond to ξ 0.15, 0.35, and 0.50, respectively. Furthermore, note that, just as Cohen's *d*, the estimate of ξ can exceed 1 when there are more than two groups and the amount of trimming is greater than 0 (Wilcox, 2012, p. 166, 294).

In this case, all ANOVA's (robust or otherwise) are non significant, but effect sizes are high (all reported in Table 6.4). As a side note, the robust or non robust equivalent for each parameter gives the same result. Thus, it seems that the microwear parameters do not differ significantly across groups. The interpretation of this result is in Chapter 6.

Now, this result might be biased by the low sample size. And therefore, an interesting idea is to conduct power analysis to calculate more appropriate sample sizes to the analysis conducted here.

C.2.4. CALCULATING THE POWER OF THE TEST AND A BETTER SAMPLE SIZE

Despite the above considerations on effect sizes, the package pwr, used for power analysis with ANOVA, uses Cohen's *f* as a measure of effect size for ANOVA. The same is true for the program G*Power. This effect size is a generalization of Cohen's *d* to several samples, and can be broadly defined as "the standard deviation of standardized means" (Cohen, 1988, p. 276) (The defining formula is found in Cohen (1988, p.275)). Just like Cohen's *d*, it can take values between 0 and an indefinitely large number. By convention, 0.10, 0.25, and 0.40 correspond to small, medium, and large sample sizes, respectively (Cohen, 1988, 355). In addition, Cohen's *f* can easily be calculated from eta squared (which is the name given to Pearson r^2 in ANOVA or omega-squared (Cohen, 1988, p. 284).

With unequal sample sizes, if the samples whose means are extreme also have large *n*, then *f* will be larger than with equal sample sizes; conversely, if extreme populations have low *n*'s, *f* will be smaller. The formula for unequal sample sizes is in Cohen (1988, p. 360) and here. Furthermore, the convention in regard to small, medium, and large *f* values continue to be applicable, as does the relationship between eta-squared and *f* (Cohen, 1988, p. 361).

However, there is no inbuilt function in R to calculate Cohen's *f* directly (not from other measurements) when the group sizes are not equal. Particularly, because it is hard to calculate the numerator of the equation. For this reason, the program G*Power was used to calculate this effect size.

```
#calculate power of Asfc analysis
##f calculated through G*Power
pwr.anova.test(k=5,n=7.6, f=0.408,sig.level = 0.05)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 7.6
##              f = 0.408
##      sig.level = 0.05
##          power = 0.433264
##
## NOTE: n is number in each group

#calculate power of epLsar analysis
##f calculated through G*Power
pwr.anova.test(k=5,n=7.6, f=0.448,sig.level = 0.05)

## ##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 7.6
##              f = 0.448
##      sig.level = 0.05
##          power = 0.5142631
##
## NOTE: n is number in each group

#calculate power of HAsfc analysis
##f calculated through G*Power
pwr.anova.test(k=5,n=7.6, f=0.390,sig.level = 0.05)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 7.6
##              f = 0.39
##      sig.level = 0.05
##          power = 0.3979187
```

```
##
## NOTE: n is number in each group

#calculate power of Tfo analysis
##f calculated through G*Power
pwr.anova.test(k=5,n=7.6, f=0.472,sig.level = 0.05)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 7.6
##              f = 0.472
##      sig.level = 0.05
##          power = 0.5633237
##
## NOTE: n is number in each group
```

All power calculations match G*Power calculations (Table 6.5). After finding that the statistical power was too low to confidently detect the effects present in the population, it is also interesting to calculate a more adequate sample size.

```
#calculate sample size for Asfc analysis
#f calculated through G*Power
pwr.anova.test(k=5,f=0.408,sig.level = 0.05, power=0.8)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 15.31896
##              f = 0.408
##      sig.level = 0.05
##          power = 0.8
##
## NOTE: n is number in each group

#calculate sample size for epLsar analysis
#f calculated through G*Power
pwr.anova.test(k=5,f=0.448,sig.level = 0.05, power=0.8)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 12.8785
##              f = 0.448
##      sig.level = 0.05
##          power = 0.8
##
## NOTE: n is number in each group
```

```
#calculate sample size for HAsfc analysis
#f calculated through G*Power
pwr.anova.test(k=5,f=0.390,sig.level = 0.05, power=0.8)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 16.67064
##              f = 0.39
##      sig.level = 0.05
##          power = 0.8
##
## NOTE: n is number in each group

#calculate sample size for Tfo analysis
#f calculated through G*Power
pwr.anova.test(k=5,f=0.472,sig.level = 0.05, power=0.8)

##     Balanced one-way analysis of variance power calculation
##
##              k = 5
##              n = 11.70352
##              f = 0.472
##      sig.level = 0.05
##          power = 0.8
##
## NOTE: n is number in each group
```

Once again, all sample calculations match G*Power (Table 6.6).

```
## R version 3.3.1 (2016-06-21)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 14393)
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] pwr_1.2-1       WRS2_0.9-1      pander_0.6.0    pgirmess_1.6.5
##  [5] broom_0.4.1     gridExtra_2.2.1 car_2.1-3       tidyr_0.6.0
##  [9] dplyr_0.5.0     ggplot2_2.2.1
##
## loaded via a namespace (and not attached):
##  [1] gtools_3.5.0      reshape2_1.4.2    splines_3.3.1
```

```
##  [4] lattice_0.20-33    mc2d_0.1-18        expm_0.999-1
##  [7] colorspace_1.2-7   htmltools_0.3.6    yaml_2.1.14
## [10] mgcv_1.8-12        nloptr_1.0.4       foreign_0.8-66
## [13] DBI_0.5-1          sp_1.2-4           plyr_1.8.4
## [16] stringr_1.2.0      MatrixModels_0.4-1 rgeos_0.3-22
## [19] munsell_0.4.3      gtable_0.2.0       mvtnorm_1.0-5
## [22] coda_0.19-1        psych_1.6.9        evaluate_0.10
## [25] knitr_1.16         SparseM_1.72       maptools_0.9-1
## [28] quantreg_5.29      pbkrtest_0.4-6     parallel_3.3.1
## [31] spdep_0.6-11       Rcpp_0.12.7        scales_0.4.1
## [34] backports_1.0.5    gdata_2.17.0       deldir_0.1-12
## [37] lme4_1.1-12        mnormt_1.5-5       digest_0.6.10
## [40] gmodels_2.16.2     stringi_1.1.2      splancs_2.01-39
## [43] grid_3.3.1         rprojroot_1.2      rgdal_1.2-5
## [46] tools_3.3.1        LearnBayes_2.15    magrittr_1.5
## [49] lazyeval_0.2.0     tibble_1.2         MASS_7.3-45
## [52] Matrix_1.2-6       reshape_0.8.6      assertthat_0.1
## [55] minqa_1.2.4        rmarkdown_1.6      R6_2.2.0
## [58] boot_1.3-18        nnet_7.3-12        nlme_3.1-128
```

C.3 SHORT REPORT ON DMTA IN THE IN-AFRICA COLLECTION

Table C.1
List of IN AFRICA samples on which DMTA was attempted.

Accession Number [a]	Location/Site	Arch	Side	Comments
WT71849a	Natome	Upper	Right	Scan failed
WT71991	Locher Akwan 1	Lower	Left	Scan failed
WT72063	Locher Akwan 2	Lower	Right	Scan failed
WT71871	Natome	Upper	Left	Scan failed
WT71251l	Nataruk	Lower	Left	Scan failed
WT71251r	Nataruk	Lower	Right	Scan failed
WT71888l	Natome	Lower	Left	Scan failed
WT71888r	Natome	Lower	Right	Scan taken of tooth M_2, facet 9 (Figure D.40)
WT70078	Locher Akwan 1	Lower	Left	Scan failed
WT72062	Locher Akwan 2	Lower	Right	Scan taken of tooth M_3, facet 10n (Figure D.41)
WT70078	Locher Akwan 1	Lower	Left	Scan taken of tooth M_1, facet 9 (Figure D.39)
WT71890l	Natome	Lower	Left	Scan failed
WT71890r	Natome	Lower	Right	Scan failed
WT70076	Locher Akwan 1	Lower	Left	Scan failed
WT71889	Natome	Lower	Right	Scan failed

[a] Accession number given to collections of the National Museums of Kenya. 'WT' stands for West Turkana, followed by an sequential number; the 'l' or 'r' were added in order to distinguish between the left or right side of the same individual.

Besides the DMTA conducted on the ethnic groups studied in this work, a pilot study was also conducted on a few of the archaeological human teeth found in the context of the IN-AFRICA Project. More exactly, these remains belonged to late Pleistocene/early Holocene sites (Lahr &

Foley, 2014). In addition, dental moulding and scanning followed standard procedures for DMTA (Scott et al., 2006)

Unfortunately, this pilot study failed and it was impossible to obtain any DMTA results from this collection (Table C.1). Most of the time the microscope simply crashed during scanning, and in the three instances in which an image was produced, it was clear that the teeth were too altered and that any microwear features had been erased (Figure C.39, Figure C.40, Figure C.41). Considering that this sample consisted of surface finds, the alterations to the teeth are probably the result of post mortem wear. This works highlights a common issue when conducting DMTA in archaeological samples, which most studies fail to report on.

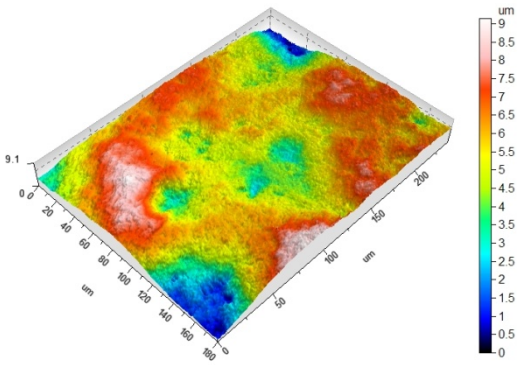


Figure C.39 Three-dimensional representation of microwear surface (WT70078, M_1, f9).

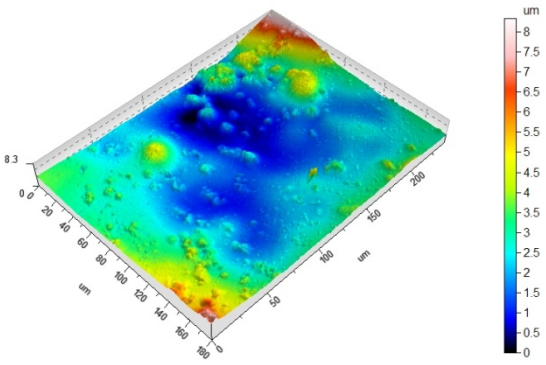


Figure C.40 Three-dimensional representation of microwear surface (WT71888l, M_2, f9).

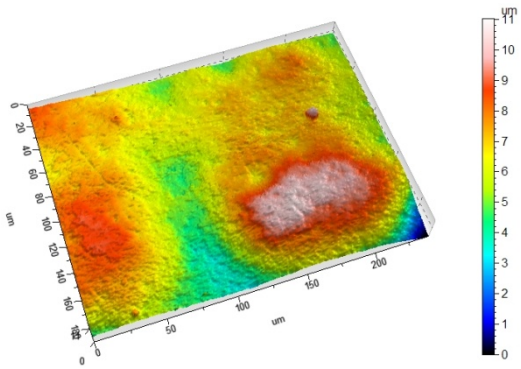


Figure C.41 Three-dimensional representation of microwear surface (WT72062, M_3, f10n).

APPENDIX D
SIRA DATA ANALYSIS

D.1 SIRA RESULTS

The full sample details and isotope results are in an Excel (.xls) spreadsheet named "APPENDIXD.1_StableIsotopeAnalysis".

D.2 SIRA STATISTICAL ANALYSES

This section includes the R code that produced all graphical and statistical analysis conducted on diet questionnaires variables. In the cases where the analysis themselves were not fully reported in text (e.g. Shapiro Wilk tests), then the output is included here. Otherwise, only the R code is present and, if applicable, the tables or figures that report the results are mentioned. In addition, for the sake of brevity, once the code for a type of test (e.g MANOVA) has been reported once, it is not repeated.

The datasets used are in APPENDIXD.1_StableIsotopeAnalysis whereas the functions used were:

```
#!setting decimals
fmt_decimals <- function(decimals=0){
   function(x) format(x,nsmall = decimals,scientific = FALSE)
}
#graphical settings for ggplot
my_theme    <- theme(axis.text=element_text(size=9,colour="black"),
     #makes numbers smaller and black (consider final display)
       axis.ticks=element_line (size=0.5,colour="black"),#same for ticks
       axis.title=element_text(size=11),#and for axis titles
       panel.grid.minor = element_blank(),
       panel.background = element_blank(),
       panel.grid.major = element_blank(),
       panel.border = element_rect(colour = "black", fill=NA, size=0.5),
       legend.key = element_blank())
#calculate outliers
is_outlier <- function(x) {
  return(x < quantile(x, 0.25,na.rm=TRUE) - 1.5 * IQR(x,na.rm=TRUE) | x > quantile(x, 0.75,na.r
m=TRUE) + 1.5 * IQR(x,na.rm=TRUE))
}
#label boxplots with `n=`
give.n <- function(x){
  return(data.frame(y = median(x)*1.05,
                    label = paste0("n=",length(x))))
  #change multiplier to find the perfect position
}
#function to calculate omega squared (effect size)(https://stats.stackexchange.com/questions/2962/omega-
squared-for-measure-of-effect-in-r)
omega_sq <- function(lm){
   sum_stats <- summary.aov(lm)[[1]]# made few changes to take "lm" objects
   SSm <- sum_stats[["Sum Sq"]][1]
   SSr <- sum_stats[["Sum Sq"]][2]
```

```
  DFm <- sum_stats[["Df"]][1]
  MSr <- sum_stats[["Mean Sq"]][2]
  W2 <- (SSm-DFm*MSr)/(SSm+SSr+MSr)
  return(W2)
}
#partial omega squared
partialOmegas <- function(mod){
 aovMod <- mod
 if(!any(class(aovMod) %in% 'aov')) aovMod <- aov(mod)
 sumAov    <- summary(aovMod)[[1]]
 residRow  <- nrow(sumAov)
 dfError   <- sumAov[residRow,1]
 msError   <- sumAov[residRow,3]
 nTotal    <- nrow(model.frame(aovMod))
 dfEffects <- sumAov[1:{residRow-1},1]
 ssEffects <- sumAov[1:{residRow-1},2]
 msEffects <- sumAov[1:{residRow-1},3]
 partOmegas <- abs((dfEffects*(msEffects-msError)) /
                (ssEffects + (nTotal -dfEffects)*msError))
 names(partOmegas) <- rownames(sumAov)[1:{residRow-1}]
 partOmegas
}
#calculate effect size standard factorial ANOVA (Field et al., p. 542)
omega_factorial<-function(n,a,b,SSa,SSb,SSab,SSr){
 MSa<-SSa/(a-1)
 MSb<-SSb/(b-1)
 MSab<-SSab/((a-1)*(b-1))
 MSr<-SSr/(a*b*(n-1))
 varA<-((a-1)*(MSa-MSr))/(n*a*b)
 varB<-((b-1)*(MSb-MSr))/(n*a*b)
 varAB<-((a-1)*(b-1)*(MSab-MSr))/(n*a*b)
 varTotal<-varA+varB+varAB+MSr
 print(paste("Omega-Squared A: ",varA/varTotal))
 print(paste("Omega-Squared B: ",varB/varTotal))
 print(paste("Omega-Squared AB: ",varAB/varTotal))
}
##The following code automates producing an equation
lm_eqn = function(m) {
 l <- list(a = format(coef(m)[1], digits = 2),
    b = format(abs(coef(m)[2]), digits = 2),
    r2 = format(summary(m)$adj.r.squared, digits = 3));
 if (coef(m)[2] >= 0)  {
  eq <- substitute(italic(y) == a + b %.% italic(x)*","~~italic(r)^2~"="~r2,l)
 } else {
  eq <- substitute(italic(y) == a - b %.% italic(x)*","~~italic(r)^2~"="~r2,l)
 }
 as.character(as.expression(eq));
}
```

```
#Analysing PCA residuals (Field et al., p. 787)
residual.stats<-function(matrix){
 residuals<-as.matrix(matrix[upper.tri(matrix)])
 large.resid<-abs(residuals)>0.05
 numberLargeResid<-sum(large.resid)
 propLargeResid<-numberLargeResid/nrow(residuals)
 rmsr<-sqrt(mean(residuals^2))

 cat("Roots means squared residual = ", rmsr,"\n")
 cat("Number of absolute residuals > 0.05 = ", numberLargeResid,"\n")
 cat("Proportion of absolute residuals > 0.05 = ",propLargeResid,"\n")
 hist(residuals)
}
```

D.2.1 EXPLORING THE DATA

In the thesis, section 7.1 summarises and graphs all stable isotope ratio analysis, whose code is included here. Furthermore, this part encompasses all tests of normality conducted on the data.

D.2.1.1 Food Results

The thesis includes all food results, without summary. Thus, this section shows only the code for Figure 7.1.

```
ggplot(Food, aes(x=mean.d13C, y=mean.d15N)) +
 geom_point(aes(colour=Category, shape=Location),size=2.5) +
 coord_fixed()+
 scale_shape_manual(values=c(21,24,25,22))+
 scale_color_manual(values=c("#56b4e9","#009e73","#e69f00"))+
 scale_x_continuous(name=expression(delta^13*C["Food"]*" "("\u2030")),
                    limits=c(-31,-10),
                    breaks=seq(-30,-10,5),
                    labels=fmt_decimals(1))+
 scale_y_continuous(name=expression(delta^15*N["Food"]*" "("\u2030")),
                    limits=c(0,20),
                    breaks=seq(0,20,5),
                    labels=fmt_decimals(1))+
 my_theme
```

D.2.1.2 Hair Results

This code produces Figure 7.2 and Table 7.2.

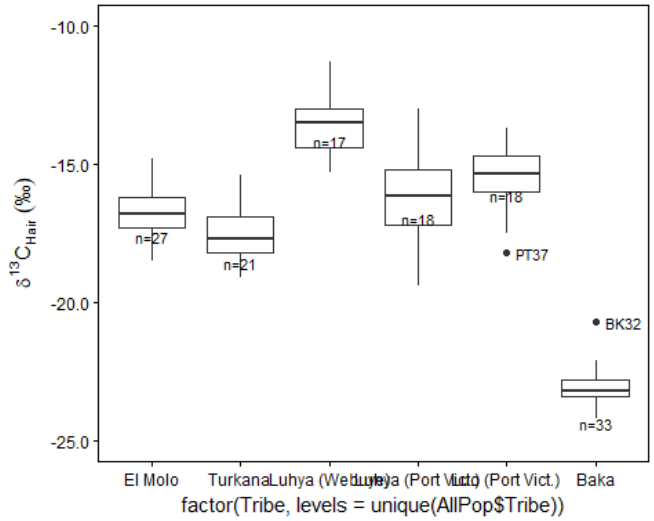
```
# Subsetting AllPop acc to tribe to produce  bagplot
EM <- filter(AllPop,Tribe=="El Molo")
TK <- filter(AllPop,Tribe=="Turkana")
WB <- filter(AllPop,Tribe=="Luhya (Webuye)")
PT_Luhya <- filter(AllPop,Tribe=="Luhya (Port Vict.)")
PT_Luo <- filter(AllPop,Tribe=="Luo (Port Vict.)")
BK <- filter(AllPop,Tribe=="Baka")

#Compute the bagplots
EMHair<-compute.bagplot(EM$Hair.d13C,EM$Hair.d15N,na.rm=TRUE)
```

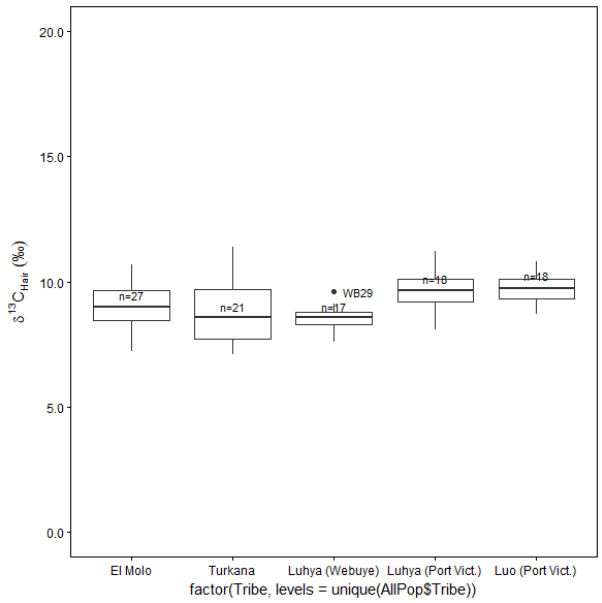
```
TKHair<-compute.bagplot(TK$Hair.d13C,TK$Hair.d15N,na.rm=TRUE)
WBHair<-compute.bagplot(WB$Hair.d13C,WB$Hair.d15N,na.rm=TRUE)
PTLhHair<-compute.bagplot(PT_Luhya$Hair.d13C,PT_Luhya$Hair.d15N,na.rm=TRUE)
PTLoHair<-compute.bagplot(PT_Luo$Hair.d13C,PT_Luo$Hair.d15N,na.rm=TRUE)
BKHair<-compute.bagplot(BK$Hair.d13C,BK$Hair.d15N,na.rm=TRUE)

#Now extracting bag hull data (..$hull.bag) and putting it into a dataframe
EMbag <- data.frame(matrix(EMHair$hull.bag, ncol = 2))
colnames(EMbag) <-c("C", "N")
TKbag <- data.frame(matrix(TKHair$hull.bag, ncol = 2))
colnames(TKbag) <-c("C", "N")
WBbag <- data.frame(matrix(WBHair$hull.bag, ncol = 2))
colnames(WBbag) <-c("C", "N")
PTLhbag <- data.frame(matrix(PTLhHair$hull.bag, ncol = 2))
colnames(PTLhbag) <-c("C", "N")
PTLobag <- data.frame(matrix(PTLoHair$hull.bag, ncol = 2))
colnames(PTLobag) <-c("C", "N")
BKbag <- data.frame(matrix(BKHair$hull.bag, ncol = 2))
colnames(BKbag) <-c("C", "N")

#colour acc to Tribe; shape acc to Diet
ggplot(AllPop, aes(x=Hair.d13C, y=Hair.d15N)) +
 geom_point(aes(colour=Tribe, shape=Livelihood,
           fill=Tribe), size=2) +
 scale_shape_manual(values=c(21,24,25,22))+
 scale_fill_manual(values=c("#f0e442","#cc79a7","#009e73",
                    "#56b4e9","#0072b2","#e69f00"))+
 scale_color_manual(values=c("#f0e442","#cc79a7","#009e73",
                    "#56b4e9","#0072b2","#e69f00"))+
 scale_x_continuous(name=expression(delta^13*C["Hair"]*" "("\u2030")),
               limits=c(-25,-10),
               breaks=round(seq(-25, -10,by=5),digits=3),
               labels=fmt_decimals(1))+
 scale_y_continuous(name=expression(delta^15*N["Hair"]*" "("\u2030")),
               limits=c(5,15),
               breaks=round(seq(5,15,5),2),
               labels=fmt_decimals(1))+
 coord_fixed()+
 geom_polygon(aes(C, N), fill = "#cc79a7",
           alpha=0.5, data = TKbag) +
 geom_polygon(aes(C, N), fill = "#f0e442",
           alpha=0.5, data = EMbag) +
 geom_polygon(aes(C, N), fill = "#009e73",
           alpha=0.5, data = WBbag)+
 geom_polygon(aes(C, N), fill = "#56b4e9",
           alpha=0.5, data = PTLhbag)+
 geom_polygon(aes(C, N), fill = "#0072b2",
           alpha=0.5, data = PTLobag)+
 geom_polygon(aes(C, N), fill = "#e69f00",
```

```
          alpha=0.5, data = BKbag)+
  my_theme
#Table 7.2
tab<-AllPop %>%
  group_by(Tribe) %>%
  summarise(meanHairC=round(mean(Hair.d13C, na.rm=TRUE),1),
          sdHairC=round(sd(Hair.d13C, na.rm=TRUE),1),
          medHairC=round(median(Hair.d13C, na.rm=TRUE),1),
          IQRHairC=round(IQR(Hair.d13C, na.rm=TRUE),1),
          nHC=sum(!is.na(Hair.d13C)),
          meanHairN=round(mean(Hair.d15N, na.rm=TRUE),1),
          sdHairN=round(sd(Hair.d15N,na.rm=TRUE),1),
          medHairN=round(median(Hair.d15N, na.rm=TRUE),1),
          IQRHairN=round(IQR(Hair.d15N, na.rm=TRUE),1),
          nHN=sum(!is.na(Hair.d15N)))
pander(tab)
```

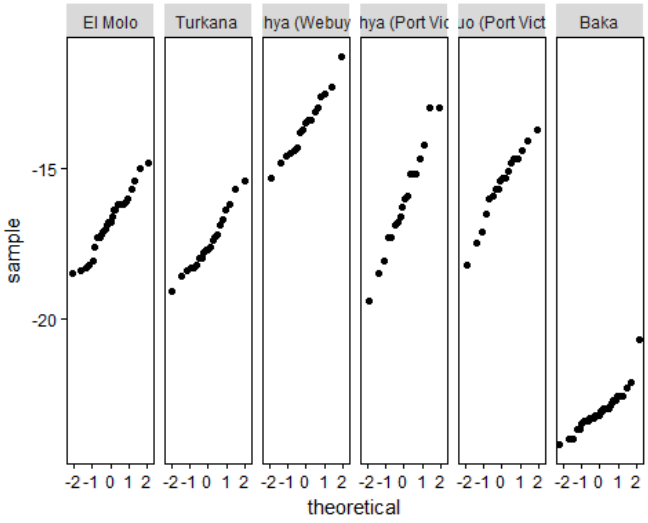
The data was visualised in boxplots, with outliers labelled based on IQR, and in qqplots, to better visualise normality. In addition, normality was formally tested using Shapiro Wilk tests.

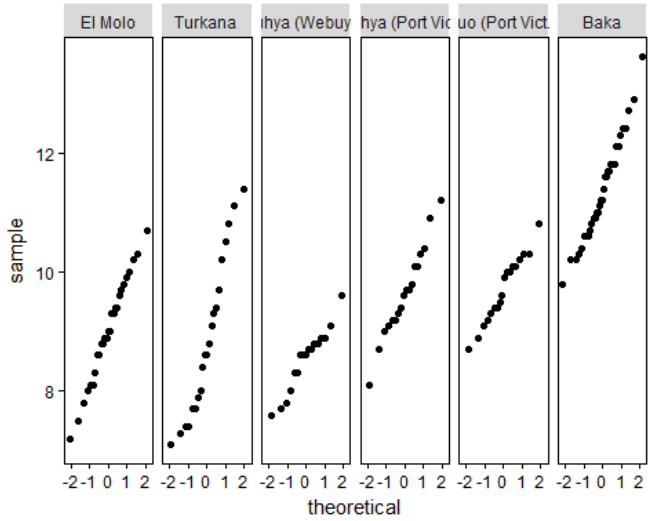
```
AllPop%>%
  group_by(Tribe) %>%
  mutate(outlier = as.character(ifelse(is_outlier(Hair.d13C),
as.character(Code),as.numeric(NA)))) %>%
  ggplot(., aes(x=factor(Tribe,levels=unique(AllPop$Tribe)),
             y = Hair.d13C)) +
    geom_boxplot() +
    geom_text(aes(label = outlier), na.rm = TRUE, hjust = -0.3,size=3)+
    my_theme+
    scale_y_continuous(name=expression(delta^13*C["Hair"]*" "("\u2030")),
                       limits=c(-25,-10),
                       breaks=seq(-25, -10,5),
                       labels=fmt_decimals(1))+
   stat_summary(fun.data = give.n, geom = "text",size=3)
```






```
ggplot(AllPop, aes(sample=Hair.d13C))+
 stat_qq()+
 facet_grid(.~Tribe)+
 my_theme
```






```
sC_tribe<-AllPop %>%
  group_by(Tribe) %>%
  do(tidy(shapiro.test(.$Hair.d13C)))
pander(sC_tribe)
```

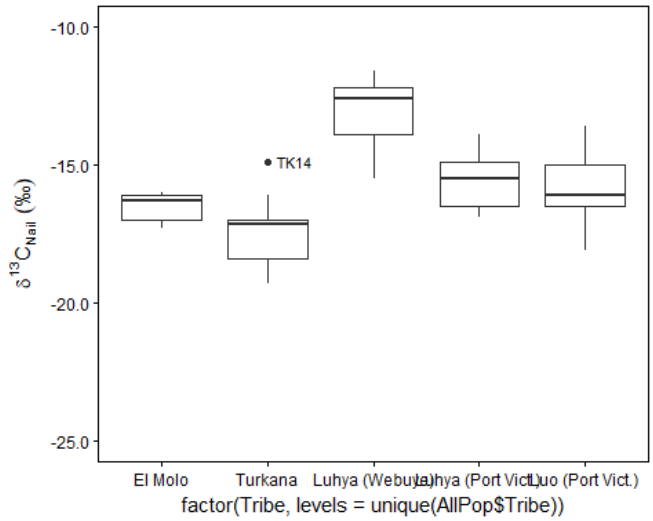
Tribe	statistic	p.value	method
El Molo	0.9679083	0.547596599	Shapiro-Wilk normality test
Turkana	0.9602508	0.521191059	Shapiro-Wilk normality test
Luhya (Webuye)	0.9795998	0.953375777	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9787816	0.936571308	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9591805	0.585886341	Shapiro-Wilk normality test
Baka	0.8853646	0.002270343	Shapiro-Wilk normality test

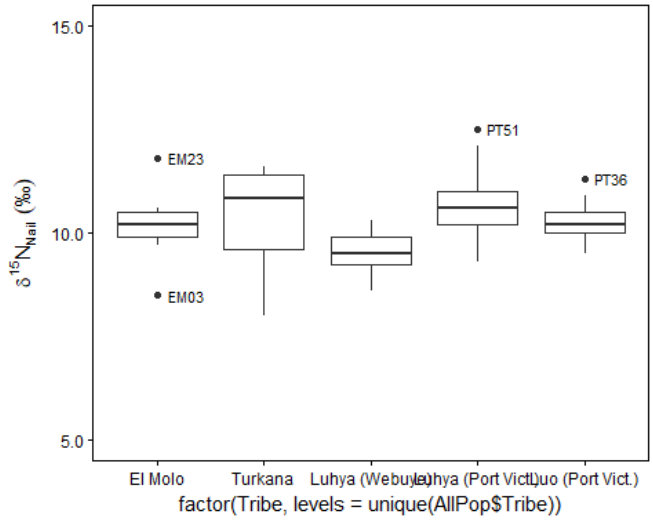
Shapiro Wilk results for for nail $\delta^{15}N$ by tribe

Tribe	statistic	p.value	method
El Molo	0.9877534	0.9820685	Shapiro-Wilk normality test
Turkana	0.9355085	0.1774237	Shapiro-Wilk normality test
Luhya (Webuye)	0.9497999	0.4533631	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9855067	0.9889772	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9711970	0.8198769	Shapiro-Wilk normality test
Baka	0.9772207	0.6995212	Shapiro-Wilk normality test

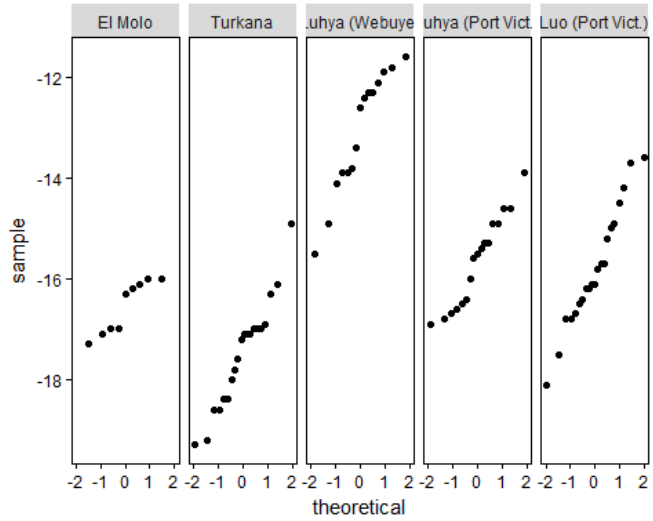
D.2.1.3 Nail Results

Similar code produces Figure 7.3 and Table 7.3, as well as the boxplots, qqplots, and Shapiro-Wilk tests. Therefore, the graphs and test results are included, but not the respective code.

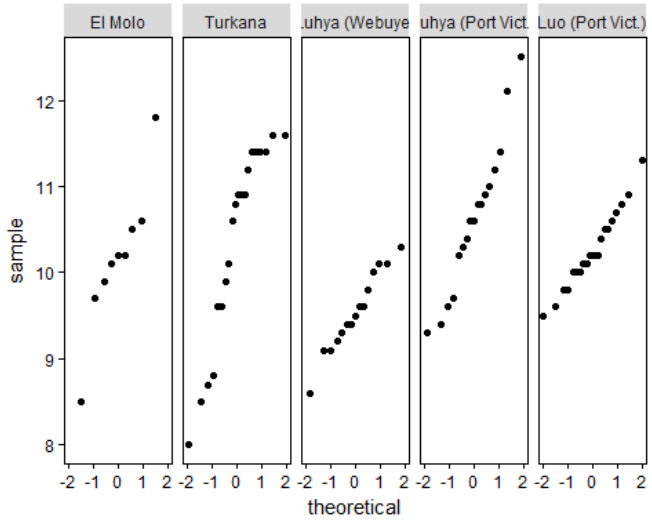




Qqplot for nail $\delta^{13}C$:



Qqplot for nail $\delta^{15}N$:



Shapiro Wilk results for nail $\delta^{13}C$ by tribe

Tribe	statistic	p.value	method
El Molo	0.8374537	0.05410387	Shapiro-Wilk normality test
Turkana	0.9546790	0.44371436	Shapiro-Wilk normality test
Luhya (Webuye)	0.9180049	0.17958518	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9436094	0.36352337	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9717115	0.77058697	Shapiro-Wilk normality test

Shapiro Wilk results for for nail $\delta^{15}N$ by tribe

Tribe	statistic	p.value	method
El Molo	0.9243514	0.42947851	Shapiro-Wilk normality test
Turkana	0.8802948	0.01790981	Shapiro-Wilk normality test
Luhya (Webuye)	0.9707565	0.86909721	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9618785	0.66701894	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9736890	0.81253781	Shapiro-Wilk normality test

D.2.1.4 Breath Results

The same reasoning was applied to breath results, where boxplots, qq plots, and the results of Shapiro Wilk tests, but not the code itself. In contrast, the code to conduct Kendall and Pearson correlation tests is shown, since this is the first time this test appears in the Appendix.

Section 7.1.4 reported the correlation between breath and time of day.

```
c<-tidy(cor.test(as.numeric(BT3$Chms),
                 BT3$Breath.d13C,
                 method="kendall",
                 use="pairwise.complete.obs"))
pander(c)
```

estimate	statistic	p.value	method	alternative
-0.136	-2.708	0.006764	Kendall's rank correlation tau	two.sided

```
#by tribe
p_tribe2<-BT3 %>%
  group_by(Tribe) %>%
  do(tidy(cor.test(.$Breath.d13C, as.numeric(.$Chms),method="pearson",use="pairwise.complet
```

```
e.obs")))
pander(p_tribe2)
```

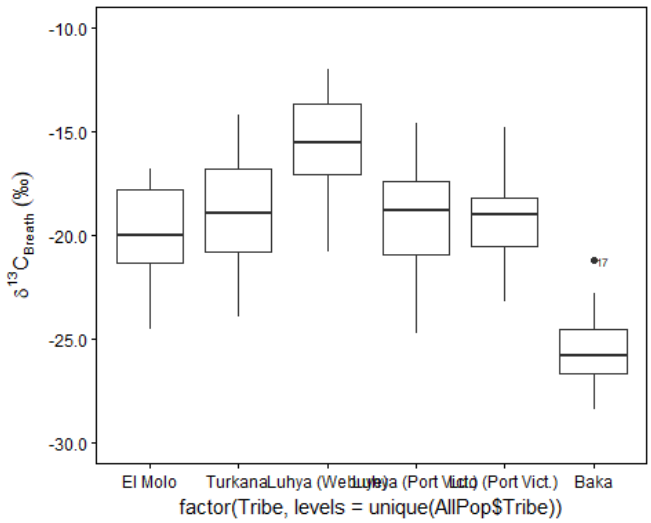
Table continues below

Tribe	estimate	statistic	p.value	parameter	conf.low
El Molo	0.35349615	1.9268916	0.06498862	26	-0.02255542
Turkana	0.02390328	0.1309611	0.89668038	30	-0.32752055
Luhya (Webuye)	-0.03540792	-0.1940589	0.84743786	30	-0.37941741
Luhya (Port Vict.)	-0.36498809	-2.0744470	0.04734571	28	-0.64097320
Luo (Port Vict.)	0.11089833	0.5798211	0.56684266	27	-0.26643663
Baka	-0.20304100	-1.1357589	0.26504978	30	-0.51525532

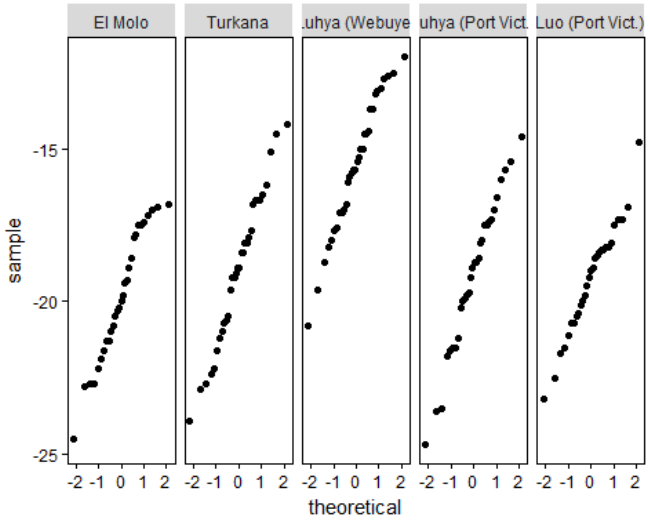
conf.high	method	alternative
0.641916343	Pearson's product-moment correlation	two.sided
0.369517349	Pearson's product-moment correlation	two.sided
0.317202395	Pearson's product-moment correlation	two.sided
-0.005433321	Pearson's product-moment correlation	two.sided
0.458757859	Pearson's product-moment correlation	two.sided
0.156750801	Pearson's product-moment correlation	two.sided

This line of code applies an FDR (false discovery rate) correction to p-values.

```
p.adjust(p_tribe2$p.value, method="fdr")#p.value correction
## [1] 0.1949659 0.8966804 0.8966804 0.1949659 0.8502640 0.5300996
```



Qqplot for breath δ15N:



Shapiro Wilk results for breath δ13C by tribe

Tribe	statistic	p.value	method
El Molo	0.9519752	0.2059198	Shapiro-Wilk normality test
Turkana	0.9791253	0.7737835	Shapiro-Wilk normality test
Luhya (Webuye)	0.9729835	0.5854875	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9788702	0.7947801	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9826401	0.8995382	Shapiro-Wilk normality test
Baka	0.9628880	0.3288246	Shapiro-Wilk normality test

D.2.2 COMPARISON ACROSS POPULATIONS

In this case:

Dependent variables: $delta^{15}N$ and $delta^{13}C$

Independent variables: Tribe

Since there are two dependent variables, one should conduct MANOVA because it takes into account the relationship between the two dependent variables. However, there might not be a need to use MANOVA if the dependent variables are not correlated (Field et al., 2012, p. 698).

Shapiro Wilk results for hair δ13C overall

statistic	p.value	method
0.9046	9.616e-08	Shapiro-Wilk normality test

Shapiro Wilk results for hair δ15N overall

statistic	p.value	method
0.9846	0.1369	Shapiro-Wilk normality test

Kendalls'tau for hair δ13C and δ15N:

estimate	statistic	p.value	method	alternative
-0.2719	-4.593	4.377e-06	Kendall's rank correlation tau	two.sided

Shapiro Wilk results for nail δ13C overall

statistic	p.value	method
0.9646	0.02332	Shapiro-Wilk normality test

APPENDIX D

Shapiro Wilk results for nail $\delta^{15}N$ overall

statistic	p.value	method
0.9938	0.9652	Shapiro-Wilk normality test

Kendalls'tau for nail $\delta^{13}C$ and $\delta^{15}N$:

estimate	statistic	p.value	method	alternative
-0.06704	-0.8708	0.3839	Kendall's rank correlation tau	two.sided

A low correlation was found between carbon and nitrogen in hair isotopic results, but not in nail isotopic results. Nevertheless, a MANOVA was still used, to account to any possible interaction between the two variables. On the other an ANCOVA was used to account for the influence of the time of collection on breath isotopic results.

The assumptions of the MANOVA are (in *italics* the ones that may cause problems, acc. To Field et al., 2012, p. 717)

a. Independence
b. Random Sampling
c. Multivariate normality (within groups)
d. Homogeneity of variances
e. *Linearity* (linear relationships among all pairs of dependent variables, all pairs of covariates, and all dependent variable-covariate pairs in each cell)

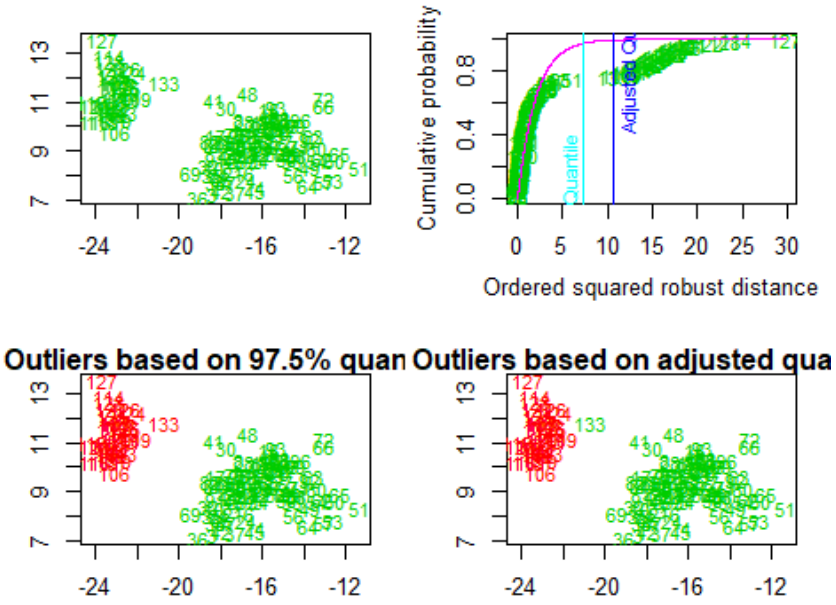
ADDITIONAL LIMITATIONS

a. *Unequal sample sizes* (when cells in a factorial MANOVA have different sample sizes, the sum of squares for effect plus error does not equal the total sum of squares. This causes tests of main effects and interactions to be correlated)
b. *Outliers* (may produce either a Type I or Type II error and give no indication as to which type of error is occurring in the analysis; check for multivariate outliers using Mahalanobis distance)

D.2.2.1 Hair Data

Univariate outliers are identified in the box plots above, whereas the plots below identify multivariate outliers in red. The aq.plot() identfies multivariate outliers by plotting the ordered squared robust Mahalanobis distances of the observations against the empirical distribution function of the MD^2_i. Input consists of a matrix or data frame. The function produces 4 graphs and returns a boolean vector identifying the outliers.

```
# Detect multivariate outliers
AllPopH<-filter(AllPop,!is.na(Hair.d13C),!is.na(Hair.d15N))
outliers <- aq.plot(AllPopH[,c("Hair.d13C","Hair.d15N")])
```



Univariate normality is addressed in the previous section. Thus, one needs to test for *multivariate normality* per group. To achieve this, one may use the mshapiro test. However, if the samples are sufficiently large (say at least 20 elements for each dependent × independent variable combination), then the Multivariate Central Limit Theorem holds, which is almost the case here. In addittion, Mahalanobis distance is used below to visualise multivariate normality.

```
#mshapiro needs data in rows rather than columns, so use t()
EM<-t(dplyr::select(EM, Hair.d13C, Hair.d15N))
TK<-t(dplyr::select(TK,Hair.d13C,Hair.d15N))
WB<-t(dplyr::select(WB,Hair.d13C,Hair.d15N))
LhP<-t(dplyr::select(LhP,Hair.d13C,Hair.d15N))
LoP<-t(dplyr::select(LoP,Hair.d13C,Hair.d15N))
BK<-t(dplyr::select(BK,Hair.d13C,Hair.d15N))

#now apply the test
mEM<-mshapiro.test(EM)
pander(mEM)
```

Shapiro-Wilk normality test: Z

Test statistic	P value
0.9867	0.9729

```
mTK<-mshapiro.test(TK)
pander(mTK)
```

Shapiro-Wilk normality test: Z

Test statistic	P value
0.9816	0.9465

```
mWB<-mshapiro.test(WB)
pander(mWB)
```

Shapiro-Wilk normality test: Z

Test statistic	P value
0.9786	0.943

```
mLhP<-mshapiro.test(LhP)
pander(mLhP)
```

Shapiro-Wilk normality test: Z

Test statistic	P value
0.9609	0.6199

```
mLoP<-mshapiro.test(LoP)
pander(mLoP)
```

Shapiro-Wilk normality test: Z

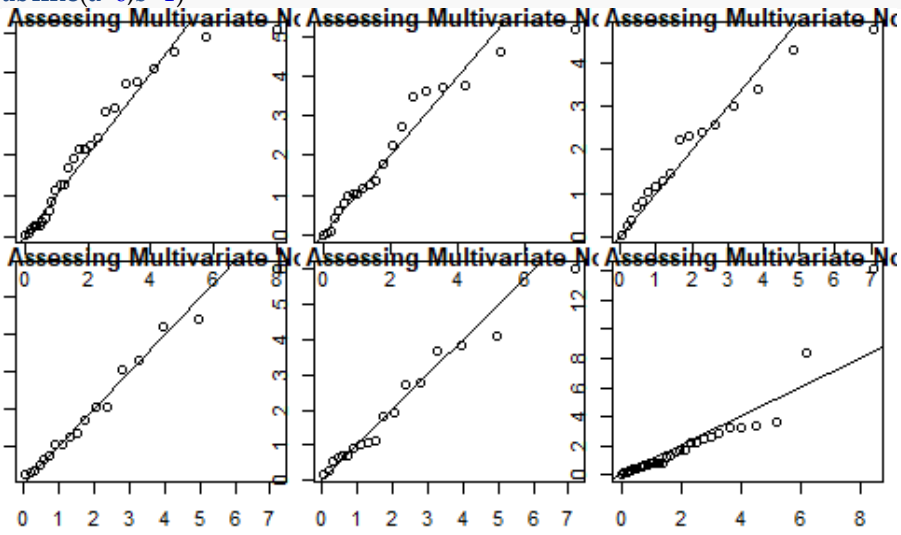
Test statistic	P value
0.9451	0.3534

```
mBK<-mshapiro.test(BK)
pander(mBK)
```

Shapiro-Wilk normality test: Z

Test statistic	P value
0.8884	0.002704 * *

```
# Graphical Assessment of Multivariate Normality
par(mfrow=c(3,3),pin=c(1,1), mar=c(0.1,0.5,0.5,0.5)+0.1)
x <- as.matrix(t(EM)) # n x p numeric matrix
center <- colMeans(x) # centroid
n <- nrow(x); p <- ncol(x); cov <- cov(x);
d <- mahalanobis(x,center,cov) # distances
qqplot(qchisq(ppoints(n),df=p),d,
 main="QQ Plot Assessing Multivariate Normality",
 ylab="Mahalanobis D2")
abline(a=0,b=1)
```



The only group that deviates significantly from multivariate normality is the Baka; and removing BK26 does not improve results substantially.

Following this, one may test for uni and mult homoscedasticity (2012, p. 717).Field et al. (2012, p. 725) does not recommend a formal test for covariance matrices but simply compare values within matrices. Alternatively, one may use Box's test to test for multivariate homoscedasticity.

```
lHC<-AllPop %>%
 do(tidy(leveneTest(.$Hair.d13C~.$Tribe)))
pander(lHC)
```

term	df	statistic	p.value
group	5	4.714	0.0005508
	128	NA	NA

Levene test for nail $\delta^{15}N$

term	df	statistic	p.value
group	5	4.269	0.001267
	128	NA	NA

```
# the matrixes are very different!
c<-by(AllPopH[,c(17,20)], AllPopH$Tribe,cov)
c
## AllPopH$Tribe: El Molo
##           Hair.d13C Hair.d15N
## Hair.d13C 1.0038462 0.3328205
## Hair.d15N 0.3328205 0.7799430
## ------------------------------------------------------------
## AllPopH$Tribe: Turkana
##           Hair.d13C Hair.d15N
## Hair.d13C 0.9351429 0.4497143
## Hair.d15N 0.4497143 1.7729048
## ------------------------------------------------------------
## AllPopH$Tribe: Luhya (Webuye)
##            Hair.d13C   Hair.d15N
## Hair.d13C  1.07382353 -0.09378676
## Hair.d15N -0.09378676  0.27720588
## ------------------------------------------------------------
## AllPopH$Tribe: Luhya (Port Vict.)
##           Hair.d13C Hair.d15N
## Hair.d13C  3.071634 1.0140523
## Hair.d15N  1.014052 0.5884967
## ------------------------------------------------------------
## AllPopH$Tribe: Luo (Port Vict.)
##           Hair.d13C Hair.d15N
## Hair.d13C  1.394281 0.2607190
## Hair.d15N  0.260719 0.3151634
## ------------------------------------------------------------
## AllPopH$Tribe: Baka
##           Hair.d13C  Hair.d15N
## Hair.d13C 0.40001894 0.08685606
## Hair.d15N 0.08685606 0.77189394
#formal test
boxM(AllPopH[, c(17,20)], AllPopH[, 3])
##  Box's M-test for Homogeneity of Covariance Matrices
## data:  AllPopH[, c(17, 20)]
## Chi-Sq (approx.) = 43.812, df = 9, p-value = 1.529e-06
```

In the end, a ROBUST MANOVA was used due to the lack of uni and multivariate homocedasticity. This test was conducted using the Minimum Covariance Determinant Estimator, implemented in R using the Wilks.test() function (Todorov and Filzmoser, 2009). There was a significant effect of group on $\delta^{13}C$ and $\delta^{15}N$. Note that low Wilks' lamda are associated with statistical significance and not the inverse (Field et al., p. 716).

```
Wilks.test(Tribe~cbind(Hair.d13C,Hair.d15N),
        data=AllPopH,method="mcd")
##  Robust One-way MANOVA (Bartlett Chi2)
## data:  x
## Wilks' Lambda = 0.042549, Chi2-Value = 285.1000, DF = 9.4432,
## p-value < 2.2e-16
## sample estimates:
##                 cbind(Hair.d13C, Hair.d15N)Hair.d13C
## Baka                                       -23.05625
## El Molo                                     -16.76667
## Luhya (Port Vict.)                          -16.12143
## Luhya (Webuye)                               -13.70000
## Luo (Port Vict.)                           -15.56111
## Turkana                                     -17.43000
##                 cbind(Hair.d13C, Hair.d15N)Hair.d15N
## Baka                                       11.306250
## El Molo                                     9.007407
## Luhya (Port Vict.)                          9.585714
## Luhya (Webuye)                               8.543750
## Luo (Port Vict.)                           9.711111
## Turkana                                     8.765000
```

Field et al. (2012, p. 719,738) recommend following up MANOVA with both univariate tests and discriminant analysis.
A significant MANOVA reflects a significant difference for one but not all the dependent variables (and hence only that dependet variable is "protected" by the MANOVA). Also, by following up MANOVA with ANOVA's it is assumed that the significant MANOVA is not due to the dependent variables representing a set of underlying dimensions that differentiate the groups. Thus, it's advisable to also conduct discriminat function analysis, which reduces and explains the dependent variables in terms of a set of underlying dimensions thought to reflect substantive theoretical dimensions.
Now, considering that only the Baka for Hair Carbon is non-normal but that there's no univariate heterocedasticity for either $\delta^{13}C$ and $\delta^{15}N$ for Hair, I will use robust ANOVA for both $\delta^{13}C$ and $\delta^{15}N$.

```
rANOVAC<-t1waybt(AllPop$Hair.d13C~AllPop$Tribe,tr=.05,nboot=999)
rANOVAC
## Call:
## t1waybt(formula = AllPop$Hair.d13C ~ AllPop$Tribe, tr = 0.05,
##     nboot = 999)
## Effective number of bootstrap samples was 999.
## Test statistic: 481.53
## p-value: 0
## Variance explained 0.911
## Effect size 0.954

## Call:
## t1waybt(formula = AllPop$Hair.d15N ~ AllPop$Tribe, tr = 0.05,
##     nboot = 999)
## Effective number of bootstrap samples was 999.
```

```
## Test statistic: 42.8521
## p-value: 0
## Variance explained 0.63
## Effect size 0.793
```

One follows robust ANOVA's using `t1waybt` with `mcppb20` post hoc tests from the same package (Table 7.5).

```
phC<-mcppb20(rANOVAC,tr= .2, nboot=5000)
phC
## Call:
## mcppb20(formula = rANOVAC, tr = 0.2, nboot = 5000)
##                                  psihat ci.lower ci.upper p-value
## El Molo vs. Turkana                0.87647 -0.12036  1.59638  0.0160
## El Molo vs. Luhya (Webuye)        -3.11444 -4.12727 -2.23957  0.0000
## El Molo vs. Luhya (Port Vict.)    -0.60686 -2.05686  0.65441  0.1848
## El Molo vs. Luo (Port Vict.)      -1.29853 -2.21275 -0.26029  0.0004
## El Molo vs. Baka                   6.39552  5.70924  6.96415  0.0000
## Turkana vs. Luhya (Webuye)        -3.99091 -4.94615 -3.06084  0.0000
## Turkana vs. Luhya (Port Vict.)    -1.48333 -2.85256 -0.07244  0.0028
## Turkana vs. Luo (Port Vict.)      -2.17500 -3.00192 -0.99359  0.0000
## Turkana vs. Baka                   5.51905  4.95458  6.32674  0.0000
## Luhya (Webuye) vs. Luhya (Port Vict.)  2.50758  1.05758  3.92348  0.0000
## Luhya (Webuye) vs. Luo (Port Vict.)    1.81591  0.87424  2.96818  0.0000
## Luhya (Webuye) vs. Baka                9.50996  8.74719 10.30130  0.0000
## Luhya (Port Vict.) vs. Luo (Port Vict.) -0.69167 -2.11667  0.90000  0.1964
## Luhya (Port Vict.) vs. Baka            7.00238  5.76786  8.27619  0.0000
## Luo (Port Vict.) vs. Baka              7.69405  6.71429  8.39762  0.0000

## Call:
## mcppb20(formula = rANOVAN, tr = 0.2, nboot = 5000)
##                                  psihat ci.lower ci.upper p-value
## El Molo vs. Turkana                0.30633 -0.88235  1.29864  0.4740
## El Molo vs. Luhya (Webuye)         0.45668 -0.17487  1.11230  0.0464
## El Molo vs. Luhya (Port Vict.)    -0.59559 -1.34314  0.10833  0.0196
## El Molo vs. Luo (Port Vict.)      -0.69559 -1.34608 -0.03235  0.0024
## El Molo vs. Baka                  -2.28487 -3.03613 -1.58487  0.0000
## Turkana vs. Luhya (Webuye)         0.15035 -0.75734  1.33357  0.6790
## Turkana vs. Luhya (Port Vict.)    -0.90192 -1.93205  0.25321  0.0350
## Turkana vs. Luo (Port Vict.)      -1.00192 -1.93782  0.16282  0.0152
## Turkana vs. Baka                  -2.59121 -3.60476 -1.41905  0.0000
## Luhya (Webuye) vs. Luhya (Port Vict.)  -1.05227 -1.73864 -0.46818  0.0000
## Luhya (Webuye) vs. Luo (Port Vict.)    -1.15227 -1.74697 -0.62348  0.0000
## Luhya (Webuye) vs. Baka                -2.74156 -3.42078 -2.23420  0.0000
## Luhya (Port Vict.) vs. Luo (Port Vict.) -0.10000 -0.71667  0.62500  0.6546
## Luhya (Port Vict.) vs. Baka            -1.68929 -2.37500 -0.99762  0.0000
## Luo (Port Vict.) vs. Baka              -1.58929 -2.22976 -1.00357  0.0000
```

After this, one must calculate effect sizes for robust post hocs. Here, only one calculation is shown, since the results are all reported in text.

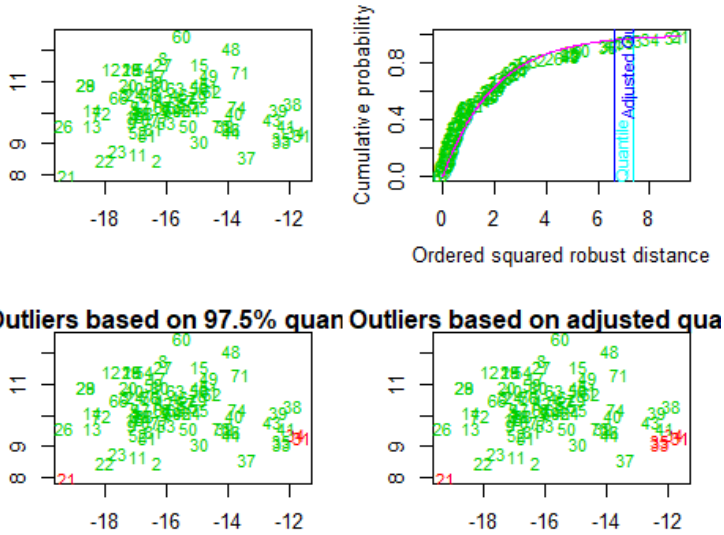
```
#use yuen.effect.ci but have to slice data first by each two comparisons
AllPopET<-AllPop%>%
  filter(Tribe == "El Molo"|Tribe=="Turkana")
yuen.effect.ci(Hair.d13C~Tribe,data=AllPopET)
## $effsize
## [1] 0.5020516
## $CI
## [1] 0.1636573 0.7295578
```

Last, the MANOVA was followed by discriminant analysis using MVE and MCD (Figure 7.5).

```
levels(AllPop3$Tribe) <- c('EM', 'TK','WB','Lh','Lo','BK')
rdfaH2<-lda(Tribe~Hair.d13C+Hair.d15N,data=AllPop3,
        prior=c(27,21,17,18,17,33)/133,
        method="mve")
rdfaH2.values<-predict(rdfaH2)
tab<-table(AllPop3$Tribe,rdfaH2.values$class)
error<-1-sum(diag(tab)/sum(tab))
error
plot(rdfaH2)
par(mar=c(0,0,0,0),pin=c(4,0.5))
ldahist(data=rdfaH2.values$x[,1],g=rdfaH2.values$class)
ldahist(data=rdfaH2.values$x[,2],g=rdfaH2.values$class)
rlda<-Linda(Tribe~Hair.d13C+Hair.d15N, data=AllPop3, method = "mcdC")
predict(rlda)
## Apparent error rate 0.3507
## Classification table
##       Predicted
## Actual EM TK WB Lh Lo BK
##     EM 14  7  0  3  3  0
##     TK  5 13  0  2  1  0
##     WB  0  0 16  0  1  0
##     Lh  6  4  3  1  4  0
##     Lo  2  2  1  3 10  0
##     BK  0  0  0  0  0 33
## Confusion matrix
##       Predicted
## Actual    EM    TK    WB    Lh    Lo BK
##     EM 0.519 0.259 0.000 0.111 0.111  0
##     TK 0.238 0.619 0.000 0.095 0.048  0
##     WB 0.000 0.000 0.941 0.000 0.059  0
##     Lh 0.333 0.222 0.167 0.056 0.222  0
##     Lo 0.111 0.111 0.056 0.167 0.556  0
##     BK 0.000 0.000 0.000 0.000 0.000  1
```

D.2.2.2 Nail Data

In the same way of hair, univariate outliers are identified in the box plots above, whereas the plots below identify multivariate outliers in red.



There's a few multivariate outliers: TK29 (21), WB01 (31), WB07(33), WB09(35).
On multivariate normality:

Shapiro-Wilk normality test: ElMolo

Test statistic	P value
0.9358	0.5383

Shapiro-Wilk normality test: Turkana

Test statistic	P value
0.9635	0.6158

Shapiro-Wilk normality test: Luhya (Webuye)

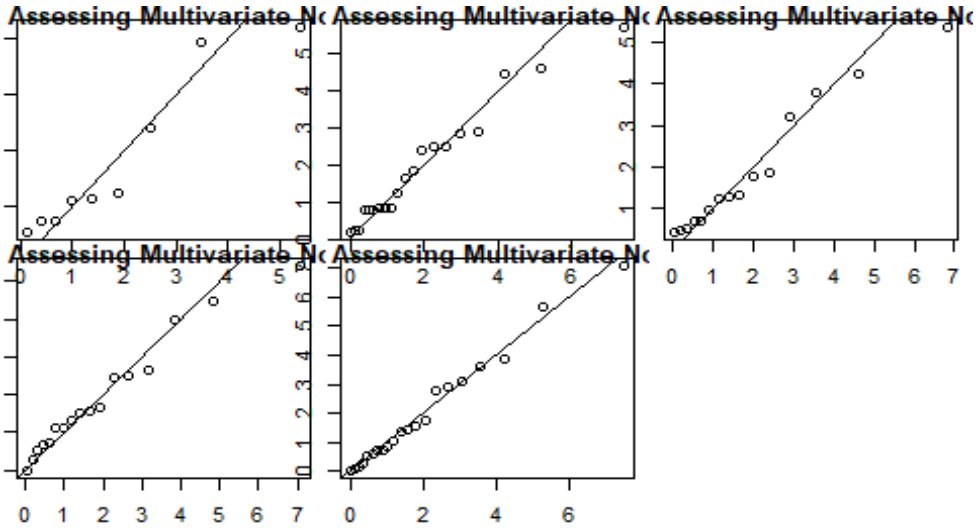
Test statistic	P value
0.9366	0.3419

Shapiro-Wilk normality test: Luhya (Port Vict.)

Test statistic	P value
0.9604	0.6381

Shapiro-Wilk normality test: Luhya (Port Vict)

Test statistic	P value
0.9573	0.4628



No group deviates significantly from multivariate normality.

APPENDIX D

On univariate homocedasticity:
Levene test for nail $\delta^{13}C$

term	df	statistic	p.value
group	4	1.028	0.3981
	77	NA	NA

Levene test for nail $\delta^{15}N$

term	df	statistic	p.value
group	4	3.337	0.0142
	77	NA	NA

For multivariate homoscedasticity.

```
## AllPopN$Tribe: El Molo
##            Nail.d13C  Nail.d15N
## Nail.d13C 0.28277778 0.05666667
## Nail.d15N 0.05666667 0.75500000
## ------------------------------------------------------------
## AllPopN$Tribe: Turkana
##            Nail.d13C Nail.d15N
## Nail.d13C 1.1764211 0.6512632
## Nail.d15N 0.6512632 1.3097632
## ------------------------------------------------------------
## AllPopN$Tribe: Luhya (Webuye)
##             Nail.d13C   Nail.d15N
## Nail.d13C  1.44714286 -0.01285714
## Nail.d15N -0.01285714  0.21257143
## ------------------------------------------------------------
## AllPopN$Tribe: Luhya (Port Vict.)
##            Nail.d13C Nail.d15N
## Nail.d13C 0.8138235 0.3740441
## Nail.d15N 0.3740441 0.7749265
## ------------------------------------------------------------
## AllPopN$Tribe: Luo (Port Vict.)
##            Nail.d13C Nail.d15N
## Nail.d13C 1.3834762 0.1517143
## Nail.d15N 0.1517143 0.1965714

##  Box's M-test for Homogeneity of Covariance Matrices
## data:  AllPopN[, c(24, 27)]
## Chi-Sq (approx.) = 14.693, df = 6, p-value = 0.02279
```

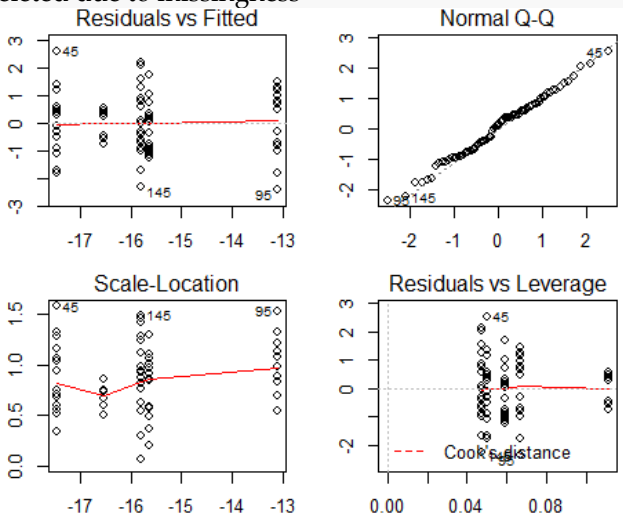
In the end, a ROBUST MANOVA was used due to lack of uni and multivariate homocedasticity.

```
##  Robust One-way MANOVA (Bartlett Chi2)
## data:  x
## Wilks' Lambda = 0.21554, Chi2-Value = 72.9880, DF = 6.9321,
## p-value = 3.368e-13
## sample estimates:
##                 cbind(Nail.d13C, Nail.d15N)Nail.d13C
```

```
## El Molo                          -16.58750
## Luhya (Port Vict.)                -15.76667
## Luhya (Webuye)                    -13.10000
## Luo (Port Vict.)                  -15.79524
## Turkana                          -17.36875
##                cbind(Nail.d13C, Nail.d15N)Nail.d15N
## El Molo                          10.37500
## Luhya (Port Vict.)                10.41333
## Luhya (Webuye)                    9.54000
## Luo (Port Vict.)                  10.25714
## Turkana                          10.83125
```

Now, considering that, for nail $\delta^{15}N$, the Turkana are non-normal and that there is no univariate heteroscedasticity, a robust ANOVA was used for this data set. For $\delta^{13}C$, a standard ANOVA was used since the data did not break the test assumptions. In addition, for the standard ANOVA, plots of residuals were drawn in order to further question the test assumptions. The interpretation of these can be found in Appendix C, p. 350-351.

```
## Call:
## t1waybt(formula = AllPopN$Nail.d15N ~ AllPopN$Tribe, tr = 0.05,
##     nboot = 999)
## Effective number of bootstrap samples was 999.
## Test statistic: 7.4855
## p-value: 0.001
## Variance explained 0.249
## Effect size 0.499

aov.C<-lm(AllPop$Nail.d13C ~ AllPop$Tribe)
summary.aov(aov.C)
##              Df Sum Sq Mean Sq F value Pr(>F)
## AllPop$Tribe  4 171.26   42.82   38.53 <2e-16 ***
## Residuals    77  85.56    1.11
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 111 observations deleted due to missingness
```



```
#effect size
omega_sq(aov.C)
## [1] 0.6467335
sqrt(omega_sq(aov.C))
## [1] 0.8041974
```

One follows robust ANOVA's using t1waybt with mcppb20 post hoc tests from the same package WRS2. On the other hand, standard ANOVA's are followed by Tukey HSD (Tukey honest signnificant difference) (Table 7.6).

```
## Call:
## mcppb20(formula = rANOVAN, nboot = 5000)
##                                  psihat ci.lower ci.upper p-value
## El Molo vs. Turkana              0.30633 -0.92036  1.31403  0.4932
## El Molo vs. Luhya (Webuye)       0.45668 -0.18610  1.11283  0.0448
## El Molo vs. Luhya (Port Vict.)  -0.59559 -1.33775  0.06422  0.0160
## El Molo vs. Luo (Port Vict.)    -0.69559 -1.36912 -0.05098  0.0024
## El Molo vs. Baka                -2.28487 -3.00420 -1.61989  0.0000
## Turkana vs. Luhya (Webuye)       0.15035 -0.79441  1.31189  0.6512
## Turkana vs. Luhya (Port Vict.)  -0.90192 -1.92500  0.29551  0.0344
## Turkana vs. Luo (Port Vict.)    -1.00192 -1.98590  0.18013  0.0186
## Turkana vs. Baka                -2.59121 -3.58205 -1.40513  0.0000
## Luhya (Webuye) vs. Luhya (Port Vict.)   -1.05227 -1.77348 -0.48864  0.0000
## Luhya (Webuye) vs. Luo (Port Vict.)     -1.15227 -1.76439 -0.65455  0.0000
## Luhya (Webuye) vs. Baka                 -2.74156 -3.43723 -2.20649  0.0000
## Luhya (Port Vict.) vs. Luo (Port Vict.) -0.10000 -0.72500  0.59167  0.6838
## Luhya (Port Vict.) vs. Baka             -1.68929 -2.37976 -0.96429  0.0000
## Luo (Port Vict.) vs. Baka               -1.58929 -2.22262 -0.95000  0.0000

posthoc <- TukeyHSD(x=aov.C, 'AllPopN$Tribe', conf.level=0.95)
posthoc
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
## Fit: aov(formula = AllPopN$Nail.d13C ~ AllPopN$Tribe)
## $`AllPopN$Tribe`
##                                  diff       lwr       upr
## Turkana-El Molo                -0.9244444 -2.1064221  0.2575332
## Luhya (Webuye)-El Molo          3.4555556  2.2139454  4.6971657
## Luhya (Port Vict.)-El Molo      0.9143791 -0.2995337  2.1282919
## Luo (Port Vict.)-El Molo        0.7603175 -0.4128939  1.9335288
## Luhya (Webuye)-Turkana          4.3800000  3.3741808  5.3858192
## Luhya (Port Vict.)-Turkana      1.8388235  0.8674015  2.8102455
## Luo (Port Vict.)-Turkana        1.6847619  0.7647069  2.6048170
## Luhya (Port Vict.)-Luhya (Webuye)   -2.5411765 -3.5843377 -1.4980153
## Luo (Port Vict.)-Luhya (Webuye)     -2.6952381 -3.6907409 -1.6997353
## Luo (Port Vict.)-Luhya (Port Vict.) -0.1540616 -1.1147980  0.8066747
##                                  p adj
## Turkana-El Molo                0.1964664
## Luhya (Webuye)-El Molo         0.0000000
## Luhya (Port Vict.)-El Molo     0.2288809
```

```
## Luo (Port Vict.)-El Molo           0.3751934
## Luhya (Webuye)-Turkana             0.0000000
## Luhya (Port Vict.)-Turkana         0.0000110
## Luo (Port Vict.)-Turkana           0.0000217
## Luhya (Port Vict.)-Luhya (Webuye)  0.0000000
## Luo (Port Vict.)-Luhya (Webuye)    0.0000000
## Luo (Port Vict.)-Luhya (Port Vict.) 0.9915108
```

For the effect sizes, this appendix shows only one of the calculations of the standard effect sizes, since the robust effect size is exemplified in the previous section, and that all results are reported in text.

```
## Loading required package: compute.es
#EMvsTK
a<-mes(-16.6,-17.5,0.5,1.1,9,20)
## Mean Differences ES:
##  Correlation ES:
##  r [ 95 %CI] = 0.4 [ 0.02 , 0.68 ]
##   var(r) = 0.02
##   p-value(r) = 0.04
```

Last, the MANOVA was followed by discriminant analysis using MVE and MCD (Figure 7.6), but this appendix only includes the details of the discriminant analysis using MCD, since all MVE results are reported in text, and the code for those analysis are included in the previous section.

```
rlda<-Linda(Tribe~Hair.d13C+Hair.d15N, data=AllPop, method = "mcdC")
predict(rlda)
## Apparent error rate 0.3507
## Classification table
##                 Predicted
## Actual            El Molo Turkana Luhya (Webuye) Luhya (Port Vict.)
##   El Molo              14      7             0                3
##   Turkana               5     13             0                2
##   Luhya (Webuye)        0      0            16                0
##   Luhya (Port Vict.)    6      4             3                1
##   Luo (Port Vict.)      2      2             1                3
##   Baka                  0      0             0                0
##                 Predicted
## Actual            Luo (Port Vict.) Baka
##   El Molo                       3    0
##   Turkana                       1    0
##   Luhya (Webuye)                1    0
##   Luhya (Port Vict.)            4    0
##   Luo (Port Vict.)             10    0
##   Baka                          0   33
##
## Confusion matrix
##                 Predicted
## Actual            El Molo Turkana Luhya (Webuye) Luhya (Port Vict.)
##   El Molo           0.519   0.259         0.000            0.111
```

```
##  Turkana            0.238  0.619        0.000           0.095
##  Luhya (Webuye)     0.000  0.000        0.941           0.000
##  Luhya (Port Vict.) 0.333  0.222        0.167           0.056
##  Luo (Port Vict.)   0.111  0.111        0.056           0.167
##  Baka               0.000  0.000        0.000           0.000
##                Predicted
## Actual           Luo (Port Vict.) Baka
##  El Molo                 0.111    0
##  Turkana                 0.048    0
##  Luhya (Webuye)          0.059    0
##  Luhya (Port Vict.)      0.222    0
##  Luo (Port Vict.)        0.556    0
##  Baka                    0.000    1
```

D.2.2.3 Breath Data

Besides the assumptions of ANOVA, the ANCOVA also has (Field et al, 2012, pp. 464-467):
a. Independece of the covariate and independent variable b. Homogeneity of regression slopes
The first assumption states that the covariate cannot differ across the groups of the dependent variable. To test, this one may use ANOVA (Field et al, 2012, p. 472).

```
mod3<-lm(Chms~Tribe, data=AllPop)
summary.aov(mod3)
##             Df    Sum Sq   Mean Sq F value   Pr(>F)
## Tribe        5 9.867e+08 197349986    7.73 1.34e-06 ***
## Residuals   177 4.519e+09  25531711
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 10 observations deleted due to missingness
#effect size
omega_sq(mod3)
## [1] 0.1553119
sqrt(omega_sq(mod3))
## [1] 0.3940963
```

This test indicates that the data breaks this assumption and therefore that it cannot be correctly interpreted. However, here it is said that one may justify the use of ANCOVA despite the lack of independece between covariate and independent variable.
Therefore, the other assumptions of ANOVA will be checked in order to decide if a standard or robust ANCOVA should be used. Note that it is easier to check homogeneity of regression slopes after the ANCOVA has been conducted.

For outliers, all graphs and tests are included in the section above E.2.1.4.
On univariate homocedasticity:
Levene test for breath $\delta^{13}C$

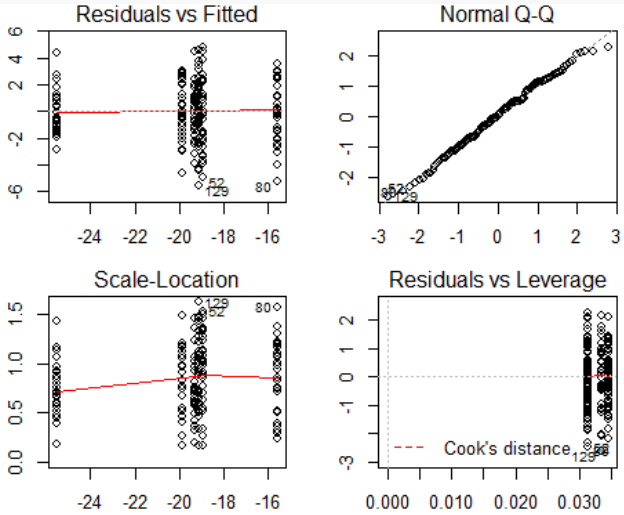
term	df	statistic	p.value
group	5	1.929	0.09172
	178	NA	NA

Next, the test itself:

```
#this the code for a factorial ANOVA, in this case, it is just meant to make sure that the tribe:chms is  not significant
## Anova Table (Type III tests)
## Response: Breath.d13C
##             Sum Sq  Df F value  Pr(>F)
## (Intercept)  15.47   1  3.3903 0.06731 .
## Tribe        42.35   5  1.8564 0.10446
## Chms         15.47   1  3.3900 0.06732 .
## Tribe:Chms   42.35   5  1.8564 0.10446
## Residuals   780.17 171
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Anova Table (Type III tests)
## Response: Breath.d13C
##             Sum Sq  Df F value Pr(>F)
## (Intercept)   2.13   1  0.4557 0.5005
## Tribe      1543.89   5 66.0713 <2e-16 ***
## Chms          2.13   1  0.4563 0.5003
## Residuals   822.52 176
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
partialOmegas(mod)
## Tribe      Chms
## 0.657431664 0.002980059
```

The first test simply aims to test the homogeneity of regression slopes. Since the tribe:chms is non significent (p<0.05), then it is reasonaly to assume this assumption is not broken.
The second test is the true ANCOVA and it indicates that time does not have a significant effect on breath carbon.
Therefore, one may simply conduct an ANOVA without taking into consideration the covariate.

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## AllPop$Tribe   5 1664.1   332.8    71.5 <2e-16 ***
## Residuals    178  828.6     4.7
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



```
omega_sq(aov.C)
## [1] 0.6570207
sqrt(omega_sq(aov.C))
## [1] 0.8105681
```

Finally, standard ANOVA's can be followed with Tukey HSD, and by the calculation of effect sizes (Table 7.7), the latter of which are not included, for the same reasons given for the effect sizes of hair and nail.

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
## Fit: aov(formula = AllPop$Breath.d13C ~ AllPop$Tribe)
## $`AllPop$Tribe`
##                                   diff        lwr       upr
## Turkana-El Molo                 0.8989224  -0.6947597  2.492604
## Luhya (Webuye)-El Molo          4.2801724   2.6864903  5.873854
## Luhya (Port Vict.)-El Molo      0.7151724  -0.9035677  2.333913
## Luo (Port Vict.)-El Molo        0.5724138  -1.0599869  2.204814
## Baka-El Molo                   -5.6729526  -7.2666347 -4.079271
## Luhya (Webuye)-Turkana          3.3812500   1.8272509  4.935249
## Luhya (Port Vict.)-Turkana     -0.1837500  -1.7634368  1.395937
## Luo (Port Vict.)-Turkana       -0.3265086  -1.9201907  1.267173
## Baka-Turkana                   -6.5718750  -8.1258741 -5.017876
## Luhya (Port Vict.)-Luhya (Webuye)   -3.5650000  -5.1446868 -1.985313
## Luo (Port Vict.)-Luhya (Webuye)     -3.7077586  -5.3014407 -2.114077
## Baka-Luhya (Webuye)             -9.9531250 -11.5071241 -8.399126
## Luo (Port Vict.)-Luhya (Port Vict.) -0.1427586  -1.7614988  1.475982
## Baka-Luhya (Port Vict.)         -6.3881250  -7.9678118 -4.808438
## Baka-Luo (Port Vict.)           -6.2453664  -7.8390485 -4.651684
##                                     p adj
## Turkana-El Molo                 0.5831246
## Luhya (Webuye)-El Molo          0.0000000
## Luhya (Port Vict.)-El Molo      0.7994210
## Luo (Port Vict.)-El Molo        0.9140467
## Baka-El Molo                    0.0000000
## Luhya (Webuye)-Turkana          0.0000000
## Luhya (Port Vict.)-Turkana      0.9994330
## Luo (Port Vict.)-Turkana        0.9915816
## Baka-Turkana                    0.0000000
## Luhya (Port Vict.)-Luhya (Webuye)   0.0000000
## Luo (Port Vict.)-Luhya (Webuye)     0.0000000
## Baka-Luhya (Webuye)             0.0000000
## Luo (Port Vict.)-Luhya (Port Vict.) 0.9998541
## Baka-Luhya (Port Vict.)         0.0000000
## Baka-Luo (Port Vict.)           0.0000000
```

D.2.3 COMPARISON TO DEMOGRAPHIC DATA

D.2.3.1 Gender

First, one must summarise and visualise the data.

```
#counting individuals by Tribe and Sex
t<-AllPop%>%
 group_by(Tribe,Sex)%>%
 summarise(b=sum(!is.na(Hair.d15N)&!is.na(Nail.d13C)),
        h=sum(!is.na(Hair.d15N)),
        n=sum(!is.na(Nail.d13C)))
pander(t)
```

Tribe	Sex	b	h	n
El Molo	M	3	12	5
El Molo	F	2	15	4
Turkana	M	10	13	15
Turkana	F	0	8	5
Luhya (Webuye)	M	2	5	13
Luhya (Webuye)	F	0	12	2
Luhya (Port Vict.)	M	1	2	13
Luhya (Port Vict.)	F	4	16	4
Luo (Port Vict.)	M	6	6	14
Luo (Port Vict.)	F	4	12	7
Baka	M	0	9	0
Baka	F	0	24	0

Counting total individuals by tissue, and with combinations of two tissues:

h	n	b	hn	hb	nb
134	82	184	32	130	78

Counting total individuals by tissue *and sex*, and with combinations of two tissues:

Sex	h	n	b	hn	hb	nb
M	47	60	84	22	46	59
F	87	22	100	10	84	19

Counting total individuals by tribe, and with combinations of two tissues:

Tribe	h	n	b	hn	hb	nb
El Molo	27	9	29	5	26	8
Turkana	21	20	32	10	20	17
Luhya (Webuye)	17	15	32	2	17	15
Luhya (Port Vict.)	18	17	30	5	17	17
Luo (Port Vict.)	18	21	29	10	18	21
Baka	33	0	32	0	32	0

Counting pregnant women

Pregnant	b	h	n
Y	2	6	2
N	8	77	20
NA	22	51	60

Counting breastfeeding women

Breastfeeding	b	h	n
Y	4	27	8
N	6	56	14
NA	22	51	60

The following is the necessary to produce Figure 7.7, 7.8, and 7.10, and Figure 7.7 is used as an example.

```
AllPop$inter <- with(AllPop, interaction(Sex,  Pregnant,Breastfeeding))
#creates dummy variable to allow coding for 3 variables
AllPop$inter<-factor(ifelse(is.na(AllPop$inter), "Male", paste(AllPop$inter)), levels = c(levels(A
llPop$inter), "Male"))
AllPop$inter<-droplevels(AllPop$inter)
levels(AllPop$inter) <- list(Breastfeeding="F.N.Y", Pregnant="F.Y.N",Female="F.N.N", Male="M
ale")

ggplot() +
 geom_point(data=AllPop, aes(x=Hair.d13C, y=Hair.d15N,colour=inter,fill=inter,
            shape=Tribe), size=2) +
 scale_shape_manual(values=c(21,24,25,22,23,8))+
 scale_color_manual(values=c("#f3b1ae","#e86f6a","#de2d26","#3182bd"),
                 name="Gender",
                 breaks=c("Male","Female", "Pregnant", "Breastfeeding"),
                 labels=c("Male","Female", "Female (Pregnant)",
                          "Female (Breastfeeding)"))+
 scale_fill_manual(values=c("#f3b1ae","#e86f6a","#de2d26","#3182bd"),
                name="Gender",
                breaks=c("Male","Female", "Pregnant", "Breastfeeding"),
                 labels=c( "Male","Female", "Female (Pregnant)",
                          "Female (Breastfeeding)"))+
 coord_fixed()+
 scale_x_continuous(name=expression(delta^13*C["Hair"]*" "("\u2030")),
                 limits=c(-25,-10),
                 breaks=round(seq(-25, -10,by=5),digits=3),
                 labels=fmt_decimals(1))+
 scale_y_continuous(name=expression(delta^15*N["Hair"]*" "("\u2030")),
                 limits=c(5,15),
                 breaks=round(seq(5,15,5),2),
                 labels=fmt_decimals(1))+
 guides(shape = guide_legend(order = 2),
       fill=guide_legend(order=1),
       color=guide_legend(order=1))+
 my_theme
```

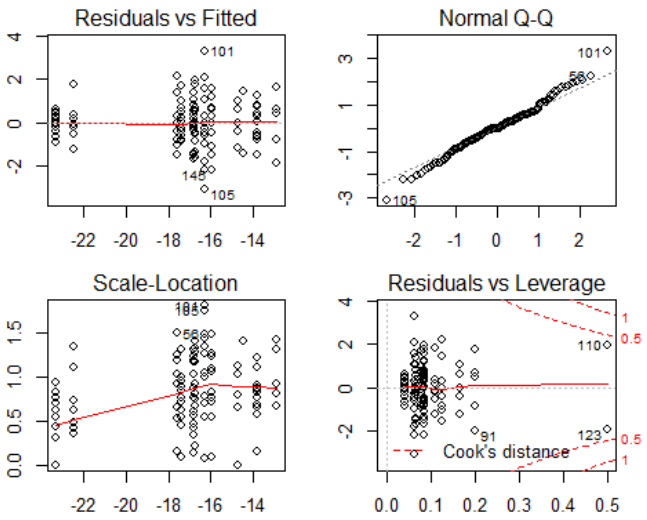
One may use a factorial ANOVA to compare the main effects of gender and tribe, as well as the interaction between the two, whereas, to understand the same for age, one may use an ANCOVA, considering age as a covariate.

Since there are a few pregnant and breastfeeding women within the sample, an initial factorial ANOVA was conducted only within women. Furthermore, significant interaction effects in

factorial ANOVAs (e.g. the effect of gender differed according to group affiliation) were followed by simple effects analysis, i.e. t-tests (robust, if following a robust factorial ANOVA) between the two genders within each population, with an FDR correction to control for the multiple comparisons. Subsequently, a significant main effect of gender (i.e. gender affected isotopic results overall) was interpreted only if no significant interaction effect was found (Field et al., 2012, p. 525). On the other hand, a significant interaction effect within an ANCOVA test breaks the assumption of homogeneity of regression slopes, which can only be addressed within a multilevel model (Field et al., 2012, p. 467). Provided this situation did not arise, the covariate was interpreted and reported (Field et al., 2012, p. 432; 480).

The code is exemplified with the carbon isotopic values of hair:

```
#since so few pregnant new variable with just breastfeeding and pregnant
AllPop$inter2 <- with(AllPop, interaction(Sex, Pregnant, Breastfeeding))
#creates dummy variable to allow coding for 3 variables
AllPop$inter2<-factor(ifelse(is.na(AllPop$inter2), "Male", paste(AllPop$inter2)), levels = c(level
s(AllPop$inter2), "Male"))
AllPop$inter2<-droplevels(AllPop$inter2)
levels(AllPop$inter2) <- list(Other=c("F.Y.N","F.N.Y"), Female="F.N.N", Male="Male")
#new database without males
female<-filter(AllPop,inter2!="Male")
female$inter2<-droplevels(female$inter2)
female2<-female%>%
  filter(!is.na(Hair.d13C))
#removed Luhya because few data points produce allias in the data, and the test fails
female3<-female2%>%
  filter(Tribe!="Luhya (Webuye)")
female3$Tribe<-droplevels(female3$Tribe)

## Call:
## t2way(formula = Hair.d13C ~ inter2 * Tribe, data = female3, tr = 0.05)
##               value p.value
## inter2       0.5584   0.468
## Tribe     1297.7205   0.001
## inter2:Tribe    1.7603   0.821
#no interaction
#since there is no interaction, one may proceed with regular factorial ANOVA
#factorial anova
```



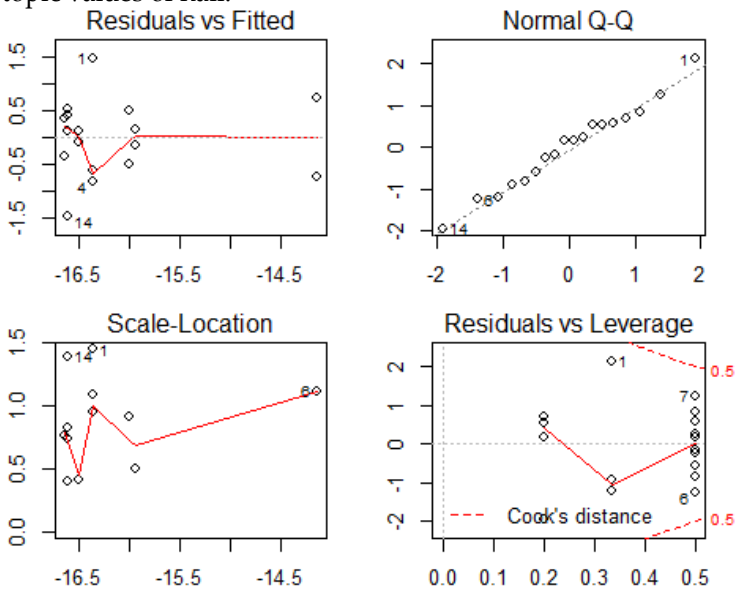

```
## Anova Table (Type III tests)
## Response: Hair.d13C
##             Sum Sq  Df   F value Pr(>F)
## (Intercept) 3366.7   1 3146.7887 <2e-16 ***
## Sex            0.0   1    0.0056 0.9404
## Tribe        397.7   5   74.3477 <2e-16 ***
## Sex:Tribe      8.3   5    1.5551 0.1779
## Residuals    130.5 122
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#no interaction

omega_factorial(22.3, 2,6,0,397.7,8.3,130.5)
## [1] "Omega-Squared A:  -0.000950751093363757"
## [1] "Omega-Squared B:  0.735827579649992"
## [1] "Omega-Squared AB:  0.0107021788592298"
```

Overall this tests says that Gender does not have a significant effect and that Sex:Tribe do not interact, in other words. the effect of Tribe on hair carbon isotopic values was not different between males and females.

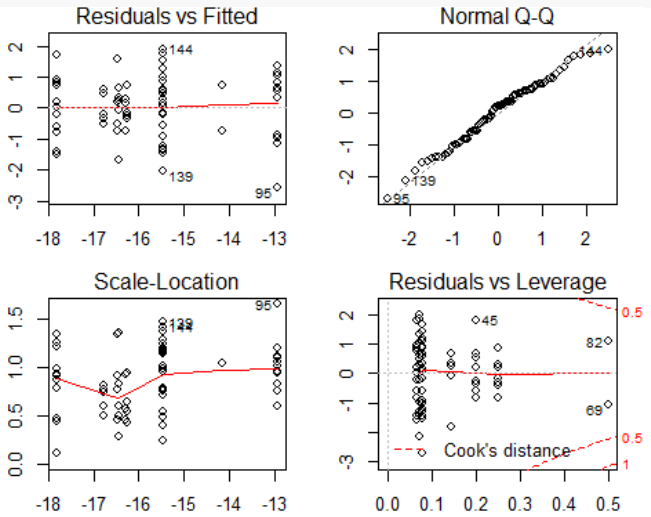
For the nitrogen isotopic values of hair:

```
## Call:
## t2way(formula = Hair.d15N ~ inter2 * Tribe, data = female3, tr = 0.05)
##               value p.value
## inter2       3.6202   0.070
## Tribe       20.8636   0.003
## inter2:Tribe  5.4467   0.209
#no interaction (can group breastfeeding and others)
# robust factorial anova
## Call:
## t2way(formula = Hair.d15N ~ Sex * Tribe, data = AllPopHN, tr = 0.05)
##             value p.value
## Sex        1.0170   0.393
## Tribe    256.5924   0.001
## Sex:Tribe 17.1918   0.060
```

For the carbon isotopic values of nail:




```
Anova(SexinF, type = "III")
## Error in Anova.III.lm(mod, error, singular.ok = singular.ok, ...): there are aliased coefficients i
n the model
#no interaction (still allias after removing Luhya..prob becaus there are very few per tribe)
#factorial anova
```




```
## Anova Table (Type III tests)
## Response: Nail.d13C
##             Sum Sq Df  F value  Pr(>F)
## (Intercept) 1407.84  1 1455.6193 < 2e-16 ***
## Sex            0.57  1    0.5860 0.44649
## Tribe        173.00  4   44.7189 < 2e-16 ***
## Sex:Tribe     15.40  4    3.9815 0.00565 **
## Residuals     69.64 72
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
omega_factorial(16.4, 2,5,0.57,173,15.4,69.64)
## [1] "Omega-Squared A:  0.000454686960309889"
## [1] "Omega-Squared B:  0.660811047238817"
## [1] "Omega-Squared AB:  0.0524629545428341"
```

Since a significant interaction effect is found in nail $\delta^{13}C$, simple effect analysis follows, as well as calculation of the effect size of the t-tests. However, the latter is not shown.

El Molo

```
EMS<-t.test(Nail.d13C~Sex,data=EM)
EMS
##  Welch Two Sample t-test
## data:  Nail.d13C by Sex
## t = -1.5341, df = 6.6382, p-value = 0.1712
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.2920665  0.2820665
## sample estimates:
## mean in group M mean in group F
##         -16.780         -16.275
```

Turkana

```
##  Welch Two Sample t-test
## data:  Nail.d13C by Sex
## t = -2.7388, df = 6.7815, p-value = 0.02988
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2.4920723 -0.1745944
## sample estimates:
## mean in group M mean in group F
##       -17.81333       -16.48000
```

Luhya (Webuye)

```
##  Welch Two Sample t-test
## data:  Nail.d13C by Sex
## t = 1.4814, df = 1.4097, p-value = 0.323
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.175793  6.598870
## sample estimates:
## mean in group M mean in group F
##       -12.93846       -14.15000
```

Luhya (Port Vict.)

```
##  Welch Two Sample t-test
## data:  Nail.d13C by Sex
## t = 2.2058, df = 9.7723, p-value = 0.05253
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01059803  1.60290572
## sample estimates:
## mean in group M mean in group F
##       -15.45385       -16.25000
```

```
Luo (Port Vict.)
```

```
##  Welch Two Sample t-test
## data:  Nail.d13C by Sex
## t = 2.161, df = 17.865, p-value = 0.04453
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.02591994 1.87408006
## sample estimates:
## mean in group M mean in group F
##       -15.47857       -16.42857
```

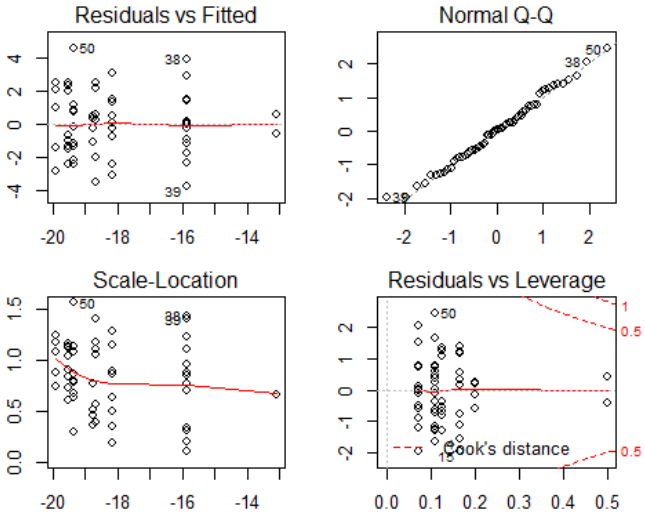
Since nail Carbon is nearly significant, possibly because of the Turkana, which are also close to significance, a graph was also produced to visualise these differences across groups (Figure 7.9).

```
#lineplot
ggplot(AllPopB, aes(x=Tribe, y=Nail.d13C, color=Sex, shape=Sex))+
 stat_summary(fun.data = mean_sdl,fun.args = list(mult = 1),
              geom="errorbar",
              position=position_dodge(width=0.60),
              width=0.3)+
 stat_summary(fun.y=mean,geom="point",
              position=position_dodge(width=0.60),
              size=2.5)+
 stat_summary(fun.data = give.n, geom = "text",size=3.5,
              position=position_dodge(width=0.60),
              colour="black")+
 scale_y_continuous(name=expression(delta^13*C["Nail"]*" "("\u2030")),
                    limits=c(-20,-10),
                    breaks=round(seq(-20, -10,by=2),digits=3),
                    labels=fmt_decimals(1),
                  expand=c(0,0))+
 my_theme+
 scale_color_manual(values=c("#3182bd","#de2d26"),
                    name="Gender",
                    labels=c( "Male","Female"))+
 scale_shape_manual(values=c(15,17),
                    name="Gender",
                    labels=c( "Male","Female"))
```

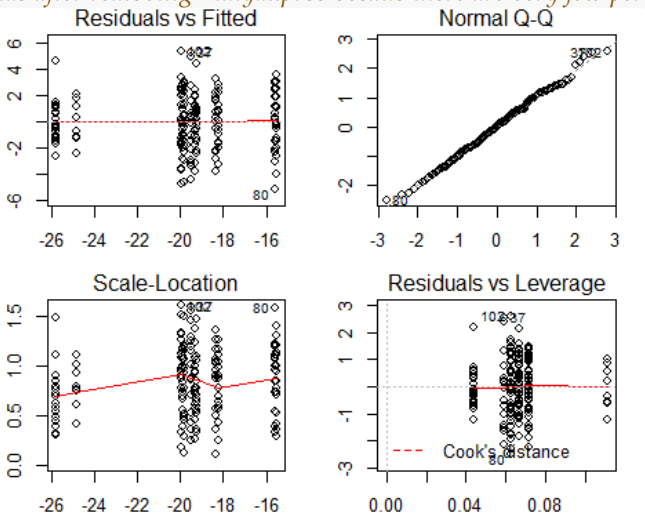
For the nitrogen isotopic values of nail:

```
## Call:
## t2way(formula = Nail.d15N ~ Sex * Tribe, data = AllPopB, tr = 0.05)
##             value p.value
## Sex        0.5195   0.483
## Tribe     48.5136   0.001
## Sex:Tribe 11.7339   0.102
```

For the carbon isotopic values of breath




```
## Anova Table (Type III tests)
## Response: Breath.d13C
##              Sum Sq Df  F value    Pr(>F)
## (Intercept)  3045.90  1 783.8668 < 2.2e-16 ***
## inter2          0.51  1   0.1325  0.717369
## Tribe          67.05  3   5.7520  0.001813 **
## inter2:Tribe    7.88  3   0.6760  0.570720
## Residuals     198.17 51
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#no interaction (still allias after removing Luhya..prob becaus there are very few per tribe)
```




```
## Anova Table (Type III tests)
## Response: Breath.d13C
##              Sum Sq  Df   F value Pr(>F)
## (Intercept) 5504.4   1 1196.6123 <2e-16 ***
## Sex            0.0   1    0.0042 0.9486
## Tribe        511.9   5   22.2572 <2e-16 ***
## Sex:Tribe     35.5   5    1.5454 0.1782
```

```
## Residuals    791.2 172
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## [1] "Omega-Squared A:  -0.0016612707562259"
## [1] "Omega-Squared B:  0.373472815126601"
## [1] "Omega-Squared AB:  0.0181698339434738"
```

D.2.3.2 Age

Moving to the influence of age, one must first check the assimptions.

```
#assumption: independence of the covariate and treatment effect
mod5<-lm(Age~Tribe, data=AllPop)
summary.aov(mod5)
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Tribe        5   3235   647.1   6.222 2.31e-05 ***
## Residuals   187  19448   104.0
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#assumption:normality of age
```

statistic	p.value	method
0.9429	6.243e-07	Shapiro-Wilk normality test

```
#assumption:normality of age by tribe
```

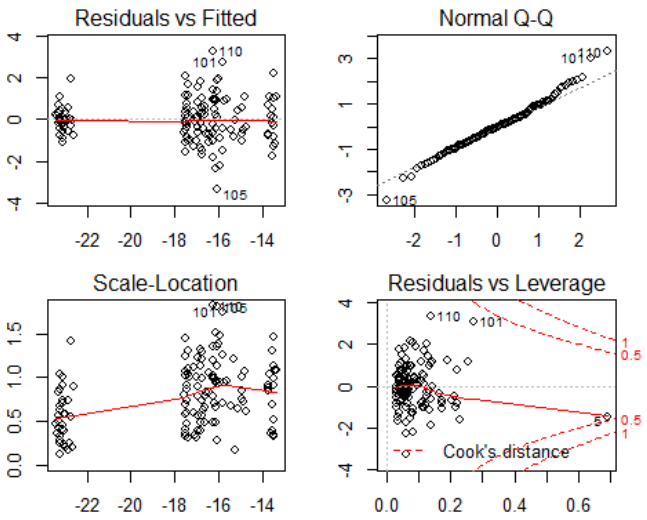
Tribe	statistic	p.value	method
El Molo	0.7451620	5.801519e-06	Shapiro-Wilk normality test
Turkana	0.9489376	8.918533e-02	Shapiro-Wilk normality test
Luhya (Webuye)	0.9334186	4.880587e-02	Shapiro-Wilk normality test
Luhya (Port Vict.)	0.9563332	2.327047e-01	Shapiro-Wilk normality test
Luo (Port Vict.)	0.9287543	5.107823e-02	Shapiro-Wilk normality test
Baka	0.9249559	2.519085e-02	Shapiro-Wilk normality test

```
p.adjust(sA_tribe$p.value, method="fdr")#El Molo non normal
## [1] 3.480912e-05 1.070224e-01 7.661734e-02 2.327047e-01 7.661734e-02
## [6] 7.557255e-02
```

Now, the actual tests, fot hair carbon isotopic values:

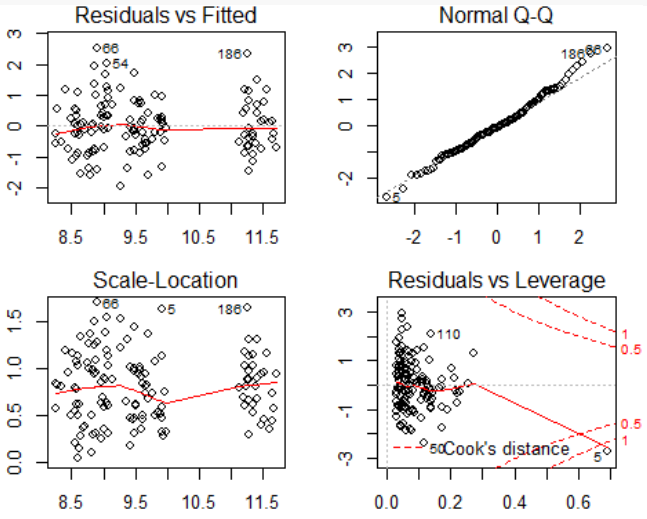
```
## Anova Table (Type III tests)
## Response: Hair.d13C
##             Sum Sq  Df  F value Pr(>F)
## (Intercept) 933.67   1 819.2150 <2e-16 ***
## Tribe       135.55   5  23.7875 <2e-16 ***
## Age           3.12   1   2.7373 0.1006
## Tribe:Age     4.45   5   0.7817 0.5647

#effect sizes
## Tribe       Age        Tribe:Age
## 0.901166256 0.038202147 0.008211235
```

APPENDIX D



For hair nitrogen isotopic results

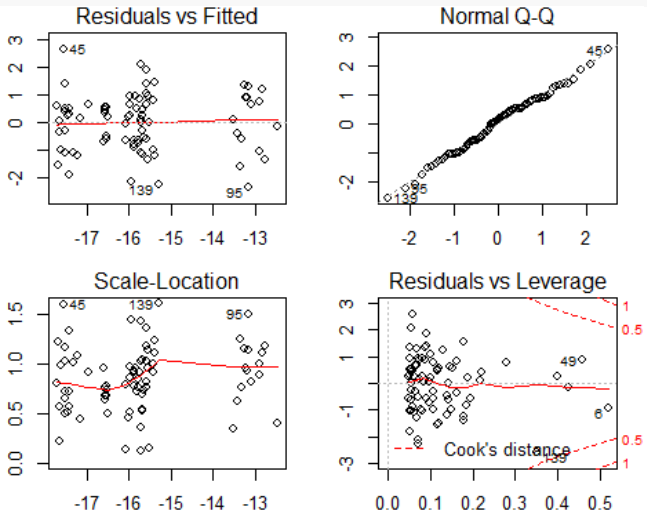
```
## Anova Table (Type III tests)
## Response: Hair.d15N
##             Sum Sq  Df  F value    Pr(>F)
## (Intercept) 208.848   1 269.3520 < 2.2e-16 ***
## Tribe        19.968   5   5.1506 0.0002538 ***
## Age           1.283   1   1.6545 0.2007768
## Tribe:Age     1.013   5   0.2613 0.9333101
## Residuals    94.595 122
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Tribe       Age        Tribe:Age
##  0.57281130  0.03573897  0.02834285
```



For nail carbon isotopic results:

```
## Anova Table (Type III tests)
## Response: Nail.d13C
##             Sum Sq Df F value    Pr(>F)
## (Intercept) 110.636  1 96.6571 6.028e-15 ***
## Tribe        14.288  4  3.1207    0.0200 *
```

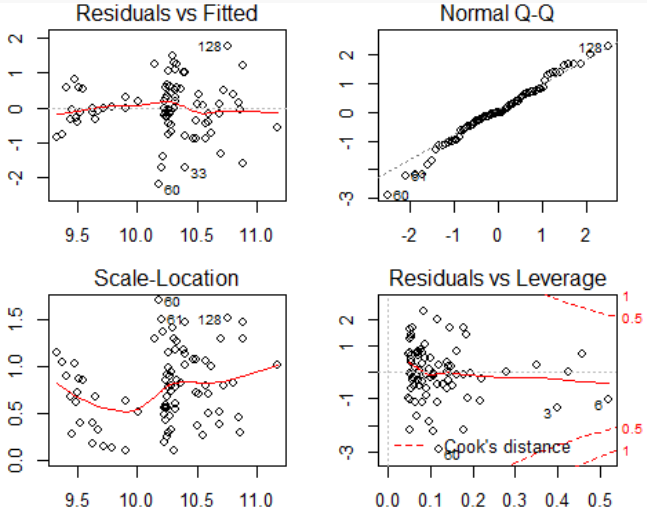
```
## Age          0.009  1  0.0082    0.9282
## Tribe:Age    2.290  4  0.5002    0.7356
## Residuals   82.413 72
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
## Tribe       Age        Tribe:Age
## 0.639755296 0.003025379 0.024988389
```

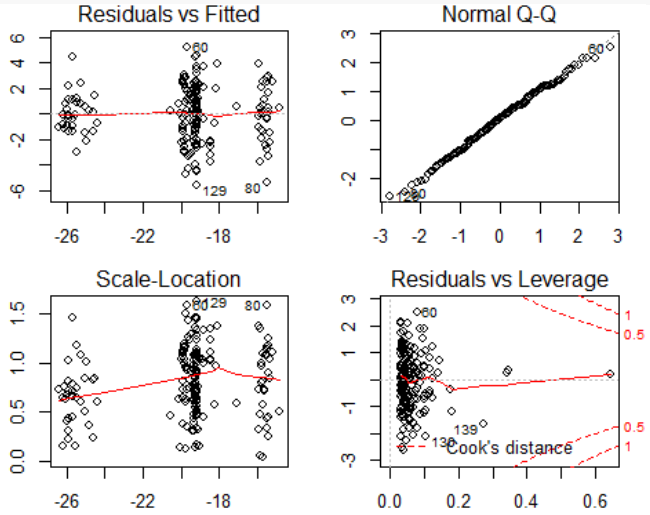
For nail nitrogen isotopic results

```
## Anova Table (Type III tests)
## Response: Nail.d15N
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 24.844  1 38.3643 3.258e-08 ***
## Tribe        4.855  4  1.8743   0.12424
## Age          2.462  1  3.8025   0.05507 .
## Tribe:Age    3.345  4  1.2914   0.28149
## Residuals   46.626 72
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Tribe       Age        Tribe:Age
## 0.127397884 0.007334235 0.014015345
```



For breath carbon isotopic results

```
## Anova Table (Type III tests)
## Response: Breath.d13C
##             Sum Sq  Df  F value    Pr(>F)
## (Intercept) 1147.31   1 244.7948 < 2.2e-16 ***
## Tribe        223.68   5   9.5450 4.713e-08 ***
## Age            0.85   1   0.1805    0.6714
## Tribe:Age     14.25   5   0.6080    0.6939
## Residuals    806.13 172
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Tribe       Age        Tribe:Age
## 0.655467743 0.004090928 0.010766259
```



D.2.4 Comparison between Body Pools

Different models were used to compare the isotopic values between body pools, but they were all variations of linear models. The assumptions of this type of model are:

a. *Variable types* All predictor variables must be quantitative or categorical (with two categories) and the outcome variable must be quantitative and continuous;
b. *Non-zero variance* Predictor variance must be different from 0;
c. *No multicollinearity* Predictor variables should not correlate too highly;
d. *Predictors uncorrelated with "external variables"* there are other variables that were not considered;
e. *Homocedasticity* Residuals should have equality of variances;
f. *Independent errors* For any two terms the residual terms should be uncorrelated;
g. *Normally distributed errors*
h. *Independence* All values of the outcome must come from separate entities;
i. *Linearity* Relationship between outcome and predictor is linear.

D.2.4.1 Hair-Nail Comparison

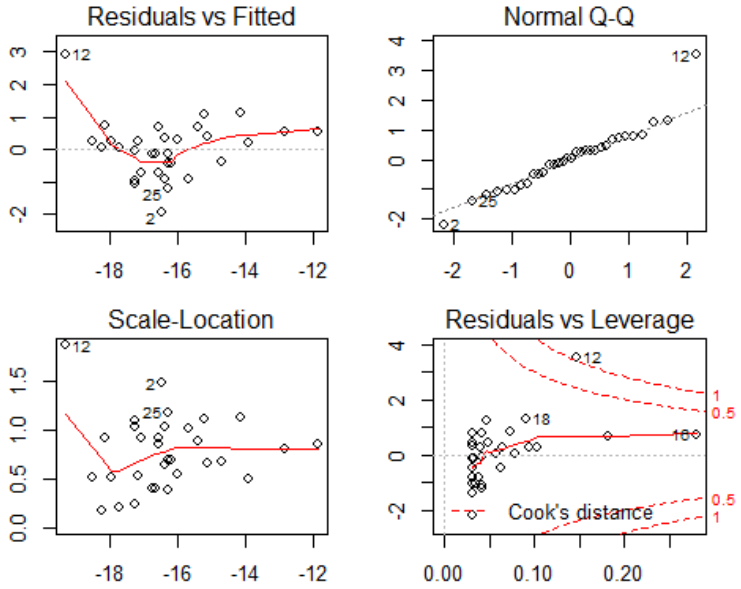
The difference in carbon and nitrogen isotopic values between hair and nail was modelled with a simple linear model.

To produce the graph for carbon (Figure 7.11):

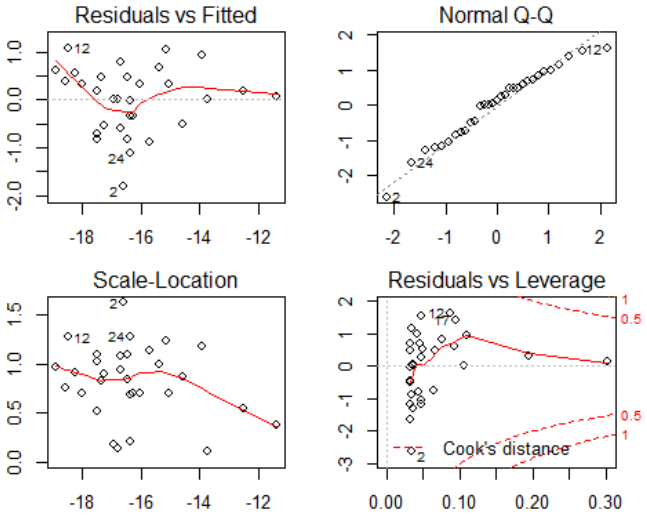
```
#lm(y~x,df)
ggplot(AllPopHN2, aes(x=Nail.d13C, y=Hair.d13C))+
 geom_point()+
 geom_smooth(method=lm,se=FALSE,size=0.7,color="black", formula = y ~ x)+
 scale_x_continuous(name=expression(delta^13*C["Nail"]*" "("\u2030")),
                    limits=c(-20,-10),
                    breaks=seq(-20, -10,5),
                    labels=fmt_decimals(1))+
 scale_y_continuous(name=expression(delta^13*C["Hair"]*" "("\u2030")),
                    limits=c(-20,-10),
                    breaks=seq(-20, -10,5),
                    labels=fmt_decimals(1))+
 coord_fixed()+
 my_theme
```

The model for carbon:

```
AllPopHN<-AllPop%>%
 filter(!is.na(Hair.d13C)&!is.na(Nail.d13C))
#linear
modCHN<-lm(Hair.d13C~Nail.d13C,AllPopHN)
summary(modCHN)
## Call:
## lm(formula = Hair.d13C ~ Nail.d13C, data = AllPopHN)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.93772 -0.51037  0.03317  0.42437  2.90782
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.46841    1.60329  -0.292    0.772
## Nail.d13C    0.98122    0.09896   9.915 5.57e-11 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.8978 on 30 degrees of freedom
## Multiple R-squared:  0.7662, Adjusted R-squared:  0.7584
## F-statistic: 98.31 on 1 and 30 DF,  p-value: 5.575e-11
#calculating confindence intervals
confint(modCHN)
##                 2.5 %   97.5 %
## (Intercept) -3.7427556 2.805941
## Nail.d13C    0.7791171 1.183321
```

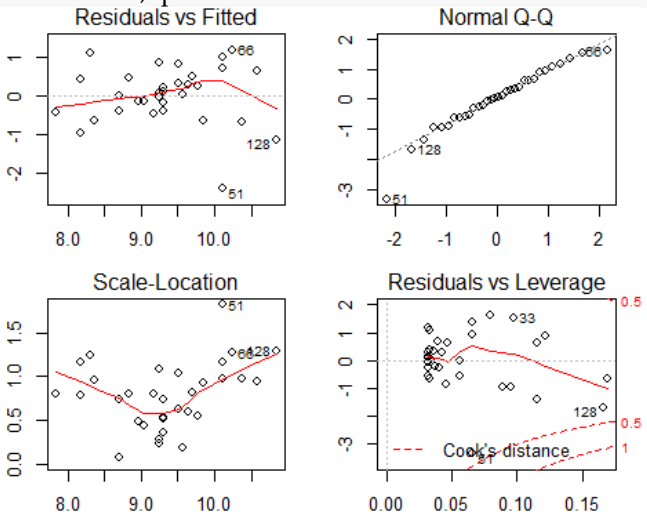



```
#12 with Cooks'above 1; remove and run model again
AllPopHN2<-AllPopHN%>%
  dplyr::filter(row_number()!=12)
## Call:
## lm(formula = Hair.d13C ~ Nail.d13C, data = AllPopHN2)
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.80858 -0.52374  0.08101  0.47400  1.07603
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.47874    1.32607   1.115    0.274
## Nail.d13C    1.10860    0.08239  13.456 5.34e-14 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.7017 on 29 degrees of freedom
## Multiple R-squared:  0.8619, Adjusted R-squared:  0.8572
## F-statistic: 181.1 on 1 and 29 DF,  p-value: 5.344e-14

##                 2.5 %   97.5 %
## (Intercept) -1.2333884 4.190861
## Nail.d13C    0.9400956 1.277101
```



The model for nitrogen:

```
## Call:
## lm(formula = Hair.d15N ~ Nail.d15N, data = AllPopHN2)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.40714 -0.41250  0.04464  0.47053  1.15919
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4878     1.2698   1.959   0.0594 .
## Nail.d15N     0.6684     0.1231   5.431 6.89e-06 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.7497 on 30 degrees of freedom
##   (161 observations deleted due to missingness)
## Multiple R-squared:  0.4958, Adjusted R-squared:  0.479
## F-statistic:  29.5 on 1 and 30 DF,  p-value: 6.89e-06
```

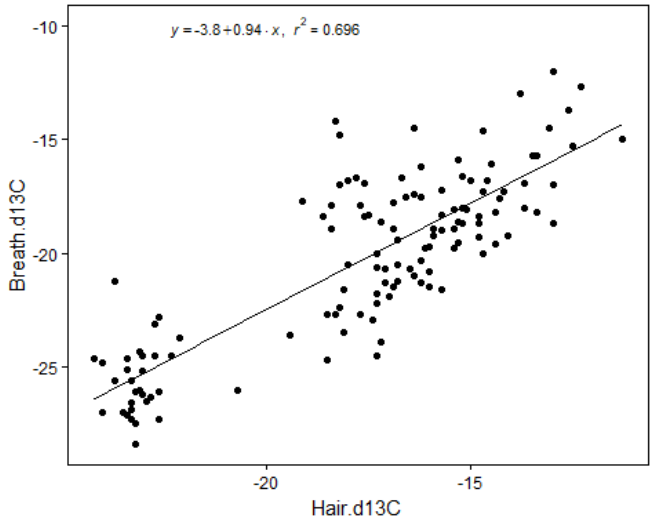



```
##                 2.5 %    97.5 %
## (Intercept) -0.1054143 5.0810373
## Nail.d15N    0.4170524 0.9196713
```

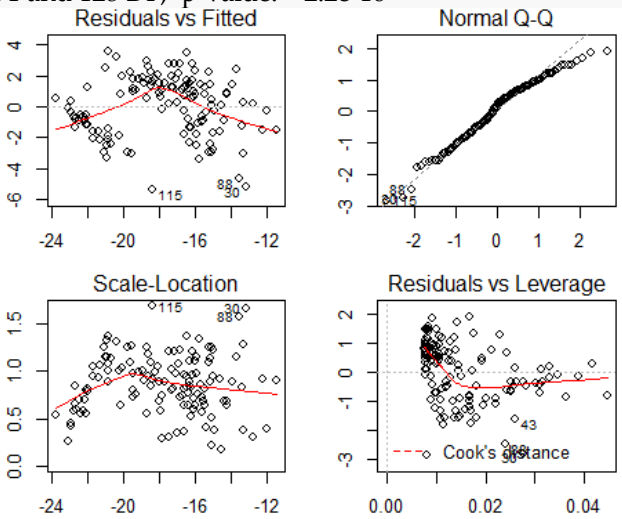
D.2.4.2 Hair-Breath Comparison

Several attempts were made to model the difference in carbon isotopic values between hair and breath.

First, a simple linear model

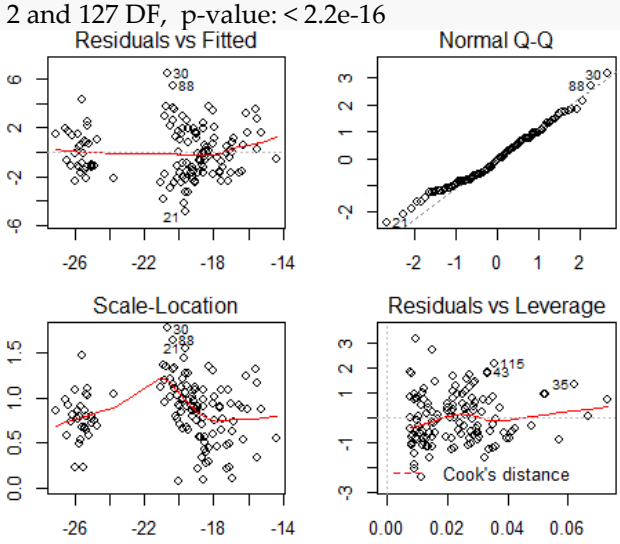



```
## Call:
## lm(formula = AllPopHB$Hair.d13C ~ AllPopHB$Breath.d13C)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.3312 -1.3166  0.2453  1.4620  3.5329
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -2.53884    0.89870  -2.825  0.00549 **
## AllPopHB$Breath.d13C  0.74670    0.04336  17.219  < 2e-16 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.889 on 128 degrees of freedom
## Multiple R-squared:  0.6985, Adjusted R-squared:  0.6961
## F-statistic: 296.5 on 1 and 128 DF,  p-value: < 2.2e-16
```



Then, a multiple linear regression, with the nitrogen isotopic signal as a measure of trophic level:

```
my.lm <- lm(AllPopHB$Breath.d13C ~ AllPopHB$Hair.d13C + AllPopHB$Hair.d15N)
summary(my.lm)
## Call:
## lm(formula = AllPopHB$Breath.d13C ~ AllPopHB$Hair.d13C + AllPopHB$Hair.d15N)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.8923 -1.6390 -0.1059  1.4720  6.4634
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -1.2760     1.3284  -0.961  0.33859
## AllPopHB$Hair.d13C  0.8322     0.0653  12.743  < 2e-16 ***
## AllPopHB$Hair.d15N -0.4472     0.1651  -2.708  0.00769 **
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 2.064 on 127 degrees of freedom
## Multiple R-squared:  0.7149, Adjusted R-squared:  0.7104
## F-statistic: 159.3 on 2 and 127 DF,  p-value: < 2.2e-16
```



This is the code to profuce the 3D plot (Figure 7.13):

```
color <- c("#f0e442","#cc79a7","#009e73",
                    "#56b4e9","#0072b2","#e69f00")
color <- color[as.numeric(as.factor(AllPopHB$Tribe))]
#lines
with(AllPopHB, {
   s3d <- scatterplot3d(Hair.d13C,Breath.d13C,Hair.d15N,
             color=color, pch=19, angle=20,
             type="h", lty.hplot=2,
             xlab=expression(delta^13*C["Hair"]*" "("\u2030")),
             ylab=expression(delta^13*C["Breath"]*" "("\u2030")),
             zlab=expression(delta^15*N["Hair"]*" "("\u2030")))
# add the legend
legend("topright", inset=-0.1,      # location and inset
 cex=.9,bg="white", xpd=T,          # suppress legend box, shrink text 10%; allows legend to be outside
plot
```

```
  title="Tribe",
  c("El Molo", "Turkana", "Luhya (Webuye)",
    "Luhya (Port Vict.)", "Luo (Port Vict.)", "Baka"), col=c("#f0e442","#cc79a7","#009e73",
                    "#56b4e9","#0072b2","#e69f00"), pch=19)
})
```

Then, a multiple linear regression with information for C4 intake (Figure 7.14):

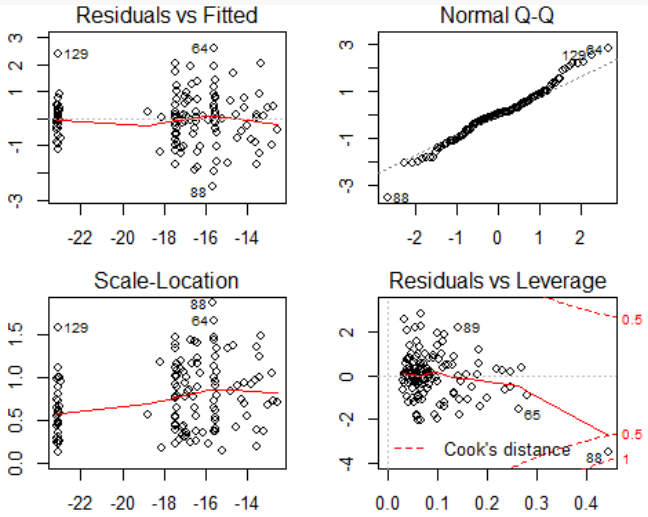
```
#make new variable with tertiles
AllDataHB<-AllDataHB %>% mutate(tertileC4W = ntile(C4_Weight, 3))
AllDataHB$tertileC4W<-as.factor(AllDataHB$tertileC4W)

#multiple regression Hair
modCHB6<-lm(Hair.d13C~Breath.d13C+tertileC4W,AllDataHB)
summary(modCHB6)
## Call:
## lm(formula = Hair.d13C ~ Breath.d13C + tertileC4W, data = AllDataHB)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9088 -0.7206  0.0610  1.0510  3.9023
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -9.82600    1.05958  -9.273 6.45e-15 ***
## Breath.d13C  0.35842    0.05518   6.496 3.90e-09 ***
## tertileC4W2  0.70872    0.35605   1.990   0.0494 *
## tertileC4W3  0.83935    0.35858   2.341   0.0214 *
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.444 on 94 degrees of freedom
## Multiple R-squared:  0.3442, Adjusted R-squared:  0.3232
## F-statistic: 16.44 on 3 and 94 DF,  p-value: 1.142e-08
#graph
ggplot(AllDataHB, aes(x= Breath.d13C, y=Hair.d13C))+
 geom_point(aes(colour=tertileC4W, shape=tertileC4W,fill=tertileC4W),
          size=2) +
 geom_smooth(method="lm", se=F, size=1.0,
           aes(group=tertileC4W, colour=tertileC4W))+
 coord_fixed()+
 scale_shape_manual(values=c(21,24,25),
                  name="C4 Weight (g/day)",
                 breaks=c("1","2","3"),
                 labels=c("tertile 1","tertile 2",
                        "tertile 3"))+
 scale_fill_manual(values=c("#56b4e9","#cc79a7","#009e73"),
                 name="C4 Weight (g/day)",
                 breaks=c("1","2","3"),
                 labels=c("tertile 1","tertile 2",
                        "tertile 3"))+
 scale_color_manual(values=c("#56b4e9","#cc79a7","#009e73"),
                  name="C4 Weight (g/day)",
                 breaks=c("1","2","3"),
                 labels=c("tertile 1","tertile 2",
```

```
                    "tertile 3"))+
  scale_x_continuous(name=expression(delta^13*C["Breath"]*" "("\u2030")),
                     limits=c(-25,-10),
                     breaks=seq(-25,-10,5),
                     labels=fmt_decimals(1))+
  scale_y_continuous(name=expression(delta^13*C["Hair"]*" "("\u2030")),
                     limits=c(-20,-10),
                     breaks=seq(-20,-10,5),
                     labels=fmt_decimals(1))+
  my_theme
```

And finally, the multiple linear regression with Tribe (Figure 7.15). This model produced the best results. The code for Figure 7.15 is the same as that of Figure 7.14.

```
modCHB7<-lm(Hair.d13C~Breath.d13C*Tribe-1,AllPopHB)
summary(modCHB7)
## Call:
## lm(formula = Hair.d13C ~ Breath.d13C * Tribe - 1, data = AllPopHB)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.50309 -0.52040  0.03127  0.47765  2.58769
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## Breath.d13C                       0.14263    0.08766   1.627  0.10638
## TribeEl Molo                    -13.97534    1.73901  -8.036 7.93e-13
## TribeTurkana                    -17.36436    1.50333 -11.551  < 2e-16
## TribeLuhya (Webuye)              -9.46540    1.73626  -5.452 2.78e-07
## TribeLuhya (Port Vict.)          -5.69059    1.70268  -3.342  0.00111
## TribeLuo (Port Vict.)            -16.20076    2.80671  -5.772 6.45e-08
## TribeBaka                       -22.39653    2.81666  -7.951 1.24e-12
## Breath.d13C:TribeTurkana          -0.13916    0.11783  -1.181  0.23997
## Breath.d13C:TribeLuhya (Webuye)    0.11654    0.13984   0.833  0.40629
## Breath.d13C:TribeLuhya (Port Vict.)  0.38663    0.12277   3.149  0.00208
## Breath.d13C:TribeLuo (Port Vict.)   -0.17667    0.17279  -1.022  0.30864
## Breath.d13C:TribeBaka             -0.11618    0.14076  -0.825  0.41085
##
## Breath.d13C
## TribeEl Molo                    ***
## TribeTurkana                    ***
## TribeLuhya (Webuye)             ***
## TribeLuhya (Port Vict.)          **
## TribeLuo (Port Vict.)           ***
## TribeBaka                       ***
## Breath.d13C:TribeTurkana
## Breath.d13C:TribeLuhya (Webuye)
## Breath.d13C:TribeLuhya (Port Vict.) **
## Breath.d13C:TribeLuo (Port Vict.)
## Breath.d13C:TribeBaka
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.9541 on 118 degrees of freedom
```

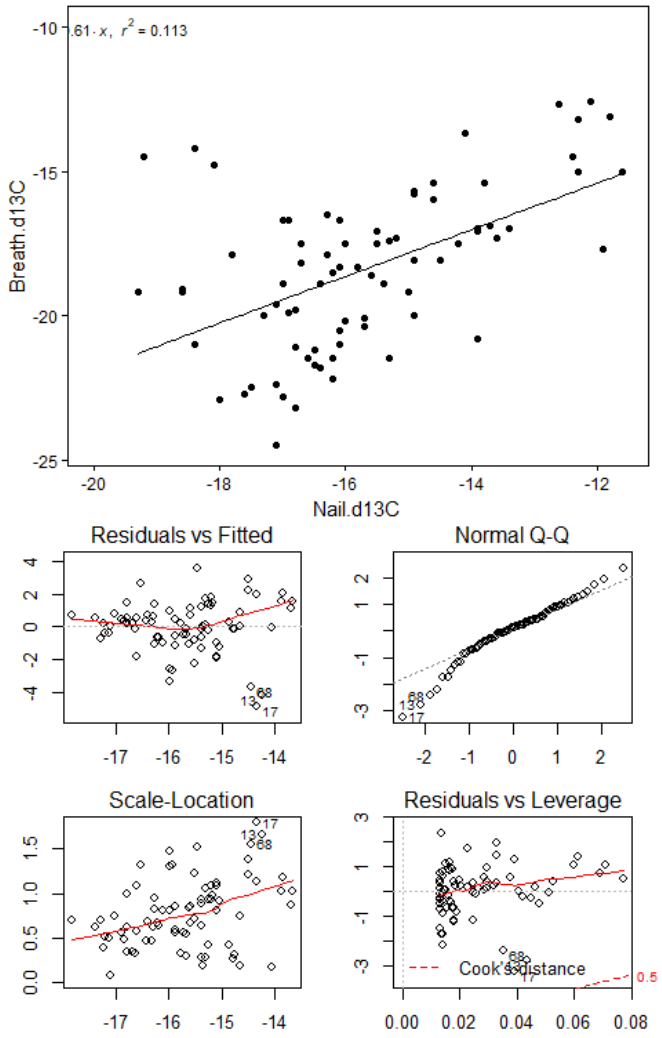
```
## Multiple R-squared:  0.9975, Adjusted R-squared:  0.9972
## F-statistic:  3878 on 12 and 118 DF,  p-value: < 2.2e-16

##                                  2.5 %      97.5 %
## Breath.d13C                   -0.0309589   0.31622052
## TribeEl Molo                 -17.4190592 -10.53162716
## TribeTurkana                 -20.3413560 -14.38735822
## TribeLuhya (Webuye)          -12.9036778  -6.02712791
## TribeLuhya (Port Vict.)       -9.0623601  -2.31882253
## TribeLuo (Port Vict.)        -21.7588207 -10.64270559
## TribeBaka                    -27.9742900 -16.81876334
## Breath.d13C:TribeTurkana      -0.3724903   0.09417498
## Breath.d13C:TribeLuhya (Webuye)   -0.1603722   0.39345811
## Breath.d13C:TribeLuhya (Port Vict.)   0.1434978   0.62975346
## Breath.d13C:TribeLuo (Port Vict.)    -0.5188428   0.16549286
## Breath.d13C:TribeBaka         -0.3949269   0.16257540
```



D.2.4.3 Nail-Breath Comparison

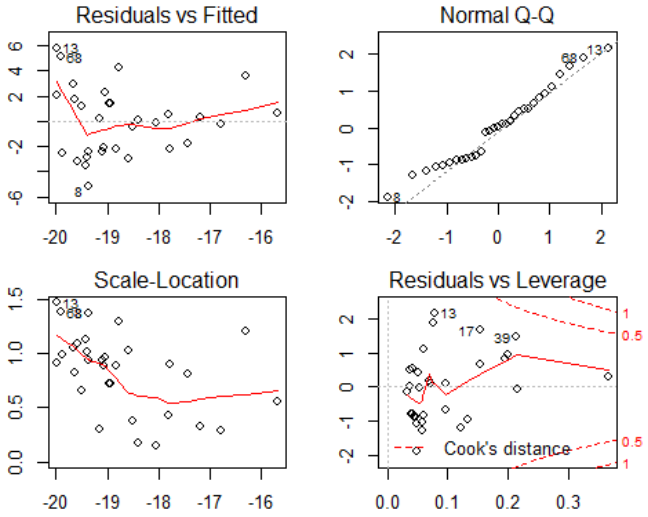
The same reasoning was applied to the difference between the carbon isotopic values of hair and nail. First, the simple linear model, and hence, only the results not reported are shown, and not the code:

```
## Call:
## lm(formula = AllPopNB$Nail.d13C ~ AllPopNB$Breath.d13C)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.8634 -0.6646  0.1696  0.8432  3.5524
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -9.28026    1.18826  -7.810 2.54e-11 ***
## AllPopNB$Breath.d13C  0.34871    0.06387   5.459 5.78e-07 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.536 on 76 degrees of freedom
## Multiple R-squared:  0.2817, Adjusted R-squared:  0.2722
## F-statistic:  29.8 on 1 and 76 DF,  p-value: 5.779e-07
```



Then, the multiple regression with nitrogen isotopic values (Figure 7.16):

```
## Call:
## lm(formula = AllPopNB$Breath.d13C ~ AllPopNB$Hair.d13C + AllPopNB$Hair.d15N)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1453 -2.3075  0.0888  1.5985  5.7766
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -8.59194    7.49239  -1.147   0.2612
## AllPopNB$Hair.d13C  0.60974    0.28854   2.113   0.0436 *
## AllPopNB$Hair.d15N -0.02435    0.52251  -0.047   0.9632
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 2.798 on 28 degrees of freedom
##   (47 observations deleted due to missingness)
## Multiple R-squared:  0.1429, Adjusted R-squared:  0.08165
## F-statistic: 2.334 on 2 and 28 DF,  p-value: 0.1155
```



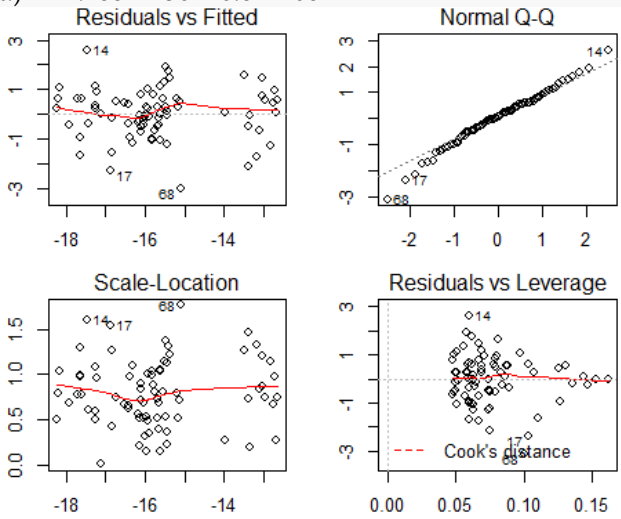
Then, a multiple linear regression with information for C4 intake (Figure 7.17):

```
## Call:
## lm(formula = Nail.d13C ~ Breath.d13C + tertileC4W, data = AllDataNB)
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6837 -0.6582  0.1118  0.8869  3.7053
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -9.5818     1.2577  -7.619 6.79e-11 ***
## Breath.d13C   0.3403     0.0663   5.133 2.22e-06 ***
## tertileC4W2   0.3394     0.4330   0.784    0.436
## tertileC4W3   0.1016     0.4329   0.235    0.815
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.55 on 74 degrees of freedom
## Multiple R-squared:  0.2878, Adjusted R-squared:  0.259
## F-statistic:  9.97 on 3 and 74 DF,  p-value: 1.347e-05
```

And finally, the multiple linear regression with Tribe (Figure 7.18). This model produced the best results.

```
##                  Df Sum Sq Mean Sq   F value Pr(>F)
## Breath.d13C       1  19146   19146 18289.851 <2e-16 ***
## Tribe             5    248      50    47.402 <2e-16 ***
## Breath.d13C:Tribe 4      4       1     0.923  0.456
## Residuals        68     71       1
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#does not break assumption of homogeneity of regression slopes and thus, a simple multiple regression ma
y be used
## Call:
## lm(formula = Nail.d13C ~ Breath.d13C + Tribe - 1, data = AllPopNB)
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.01932 -0.52057  0.04701  0.60165  2.56587
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## Breath.d13C         0.16049    0.05243   3.061   0.0031 **
## TribeEl Molo      -13.16787    1.14693 -11.481   <2e-16 ***
```

```
## TribeTurkana          -14.56104    1.00650 -14.467   <2e-16 ***
## TribeLuhya (Webuye)    -10.63170    0.84841 -12.531   <2e-16 ***
## TribeLuhya (Port Vict.) -12.63911    1.01158 -12.494   <2e-16 ***
## TribeLuo (Port Vict.)   -12.70546    1.03376 -12.291   <2e-16 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.021 on 72 degrees of freedom
## Multiple R-squared:  0.9961, Adjusted R-squared:  0.9958
## F-statistic:  3101 on 6 and 72 DF,  p-value: < 2.2e-16

##                          2.5 %     97.5 %
## Breath.d13C             0.05596413   0.2650117
## TribeEl Molo          -15.45422602 -10.8815137
## TribeTurkana          -16.56746303 -12.5546164
## TribeLuhya (Webuye)    -12.32297987  -8.9404121
## TribeLuhya (Port Vict.) -14.65566466 -10.6225525
## TribeLuo (Port Vict.)   -14.76622150 -10.6447061
```



D.2.5 Effect of Diet on Modern Human Isotopic Values

Both correlations and PCA were used to study the effect of diet on modern human isotopic values.

D.2.5.1 Correlation between Stable Isotopes and Dietary Intake

The correlations were done for measures of total dietary intake, namely weight and energy, and of macronutrient consumption, namely protein, fat, and carbohydrates (Table 7.11). As before, only the first example analysis is shown.

```
#HairCarbonEnergy no
tidy(cor.test(RawData$Hair.d13C,
              RawData$Energy,
              method="kendall",
              use="pairwise.complete.obs"))
##      estimate statistic   p.value                method
## 1 -0.08718448 -1.282794 0.1995642 Kendall's rank correlation tau
```

Other important variables include these same measurements of dietary intake, but separated by food categories, i.e. animal, fish, C3 plant, and C4 plant (Table 7.11).

```
#HairCarbonAW
tidy(cor.test(AllData$Hair.d13C,
              AllData$Animal_Weight,
              method="kendall",
              use="pairwise.complete.obs"))
##      estimate   statistic   p.value                 method
## 1 -0.002593589 -0.03816093 0.9695594 Kendall's rank correlation tau
```

D.2.5.2 Principal Component Analysis

To conduct these analysis, one must first prepare the data:
Then, one must create a correlation matrix between the variables of interest.

```
corMatrix1<-cor(PCAcor1)
#exported correlation matrix: write.csv(corMatrix1, file="corMatrixM.csv")
cortest.bartlett(corMatrix1,n=98)
## Warning in log(detR): NaNs produced
## $chisq
## [1] NaN
## $p.value
## [1] NaN
## $df
## [1] 231

KMO(corMatrix1)
## matrix is not invertible, image not found

det(corMatrix1)
## [1] -2.605781e-64
```

This correlation matrix is not ideal to conduct a PCA; most values are below 0.3; and there are a few correlations higher than 0.8. On the other hand, it's good that the Bartlett's test is significant. It means that this matrix is different from an identity matrix. Following this, the KMO does not run probably due to the strongly correlated variables. Finally, the determinant is below 0.000001, which indicates multicollinearity (Field et al., 2012, pp. 769–771).

In a second attempy, the follwing were eliminated: C4 Fat, which correlates too much with C4 Carbohydrate, Protein and Energy; Fish carbohydrate, because it's always 0 and it correlates too highly with fish fat; all Weights and Energies, because these are general measures of diet intake and correlate too highly with the macronutrient that contributes the most to that category (e.g. C3 carbohydrate very correlated with C3 energy).

```
## $chisq
## [1] 546.5177
## $p.value
## [1] 1.824853e-71
## $df
## [1] 78
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = corMatrix2)
## Overall MSA =  0.62
## MSA for each item =
##  Animal_Protein   Animal_Weight C3_Carbohydrate      C3_Energy
##            0.48           0.50           0.66           0.72
##         C3_Fat     C3_Protein C4_Carbohydrate     C4_Protein
##            0.72           0.78           0.57           0.57
##      C4_Weight     Fish_Energy      Hair.d13C      Hair.d15N
##            0.61           0.60           0.57           0.70
##     Breath.d13C
##            0.52

## [1] 0.00260274
```

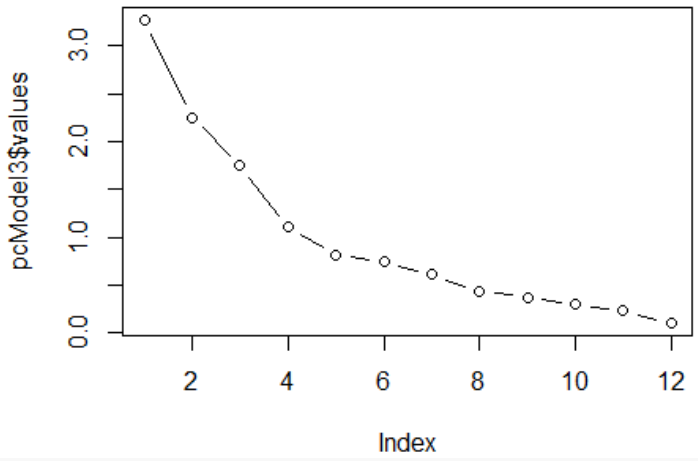
In this new correlation matrix, no correlations are higher than 0.8, but most are still below |0.3|. Nevertheless, Bartlett's test still tells me that the matrix is very different from an identity matrix (Field et al., pp. 770, 775). Following this, the KMO indicates whether the pattern of correlations is compact (desirable) or diffuse. Ideally, both the overall and individual (each variable) KMO should be minimum 0.5. The overall KMO is 0.6 and Animal Protein and Animal Weight have KMO's below 0.5 (Field et al., pp. 769, 776). On the other hand the determinant is higher that 0.00001, indicating I have solved the multicollinearity issue (Field et al., pp. 771, 777).

For the final attempt, animal protein was eliminated due to its KMO.

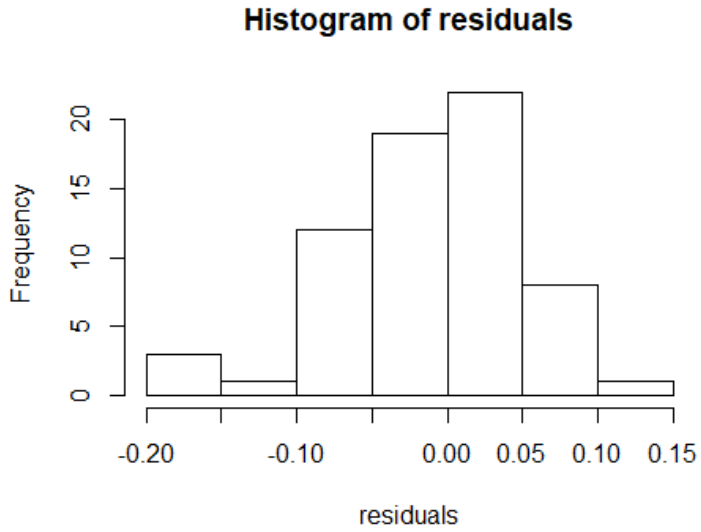
```
## $chisq
## [1] 465.8834
## $p.value
## [1] 1.699347e-61
## $df
## [1] 66

## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = corMatrix3)
## Overall MSA =  0.65
## MSA for each item =
##   Animal_Weight C3_Carbohydrate      C3_Energy         C3_Fat
##            0.74           0.71           0.69           0.72
##      C3_Protein C4_Carbohydrate     C4_Protein      C4_Weight
##            0.81           0.56           0.56           0.60
##     Fish_Energy      Hair.d13C      Hair.d15N     Breath.d13C
##            0.69           0.57           0.73           0.51

## [1] 0.006378689
```

To decide the number of factors that must kept in the PCA, one may use a scree plot (below). However, the scree plot for this data does not have a clear inflexion point, and this method is only recommended when the sample size is above 200, here the sample size is only 98. Alternatively, by Kaiser's criterion, one should keep factors with eigenvalues (SS loadings) bigger than 1 (so 5); by Joliffe's criterion we should keep factors with eigenvalues bigger than 0.7 (so 6). In the end, 6 factors were kept.

```
pcModel3<-principal(corMatrix3,nfactors=length(corMatrix3[,1]),rotate="none")
plot(pcModel3$values, type="b")
```




```
pc1<-principal(corMatrix3,nfactors=6,rotate="none")
pc1
## Principal Components Analysis
## Call: principal(r = corMatrix3, nfactors = 6, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                  PC1   PC2   PC3   PC4   PC5   PC6   h2    u2 com
## Animal_Weight    0.37  0.03 -0.51  0.03  0.62 -0.40 0.94 0.061 3.4
## C3_Carbohydrate  0.62  0.56  0.32 -0.11 -0.12 -0.11 0.83 0.169 2.7
## C3_Energy        0.49  0.62  0.37 -0.14 -0.11 -0.13 0.81 0.192 2.9
## C3_Fat           0.56  0.41 -0.07  0.38  0.22  0.35 0.81 0.193 3.9
## C3_Protein       0.76  0.43  0.09  0.00  0.05 -0.03 0.77 0.230 1.6
## C4_Carbohydrate  0.66 -0.60  0.21 -0.30  0.05  0.10 0.94 0.064 2.7
## C4_Protein       0.56 -0.59  0.06 -0.01  0.25  0.18 0.76 0.240 2.6
## C4_Weight        0.61 -0.48  0.14 -0.36 -0.12  0.08 0.77 0.230 2.9
## Fish_Energy      0.45 -0.17 -0.21  0.65 -0.22  0.25 0.80 0.195 2.9
## Hair.d13C       -0.36 -0.11  0.75  0.18  0.27  0.03 0.81 0.186 1.9
## Hair.d15N       -0.43  0.41  0.12 -0.33  0.35  0.52 0.88 0.122 4.7
## Breath.d13C     -0.07 -0.24  0.74  0.38  0.14 -0.23 0.83 0.174 2.1
##                        PC1  PC2  PC3  PC4  PC5  PC6
## SS loadings           3.27 2.26 1.75 1.11 0.81 0.75
## Proportion Var        0.27 0.19 0.15 0.09 0.07 0.06
## Cumulative Var        0.27 0.46 0.61 0.70 0.77 0.83
## Proportion Explained  0.33 0.23 0.18 0.11 0.08 0.08
## Cumulative Proportion 0.33 0.56 0.73 0.84 0.92 1.00
## Mean item complexity =  2.9
## Test of the hypothesis that 6 components are sufficient.
## The root mean square of the residuals (RMSR) is  0.06
## Fit based upon off diagonal values = 0.95

residual.stats(factor.residuals(corMatrix3,pc1$loadings))
## Roots means squared residual =  0.05927216
## Number of absolute residuals > 0.05 =  25
## Proportion of absolute residuals > 0.05 =  0.3787879
```



With 6 factors, the fit based upon off diagonal values is over 0.95, which is desired.

For the residual stats (Field et al., pp. 785-787):
Ideally the roots means squared residual should be below 0.08 (ideally lower), otherwise we might need more factors;
Regarding the proporion of absolute residuals >0.05, a value above 0.5 is reason for concern;
Finally, the residual histogram should be about normal.

In conclusion, this was the PCA used for interpretation and reported in Table 7.12.

As stated in text, a rotation maximizes the loading of each variable on one of the extracted factors while minimizing the loading of all other factors, making it clearer which variables relate to which factors (Field et al., p. 788). However, Joliffe (2002) discusses why the interpreation of rotated components is complicated, while Rencher (1992) discourages the use of rotated components."Note that rotated components are no longer principal components."

In this case, the result of rotation is that each factor is entirely explained my animal products consumption (e.g), and no other variables. However, I want to explore how the isotopes relate to the other, and so rotation does not help in achieving this goal. Nevertheless, for the record:

```
pc2<-principal(corMatrix3,nfactors=6,rotate="varimax")
pc2
## Principal Components Analysis
## Call: principal(r = corMatrix3, nfactors = 6, rotate = "varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                  RC1   RC2   RC3   RC4   RC6   RC5   h2    u2 com
## Animal_Weight    0.06  0.07 -0.20  0.05 -0.09  0.94 0.94 0.061 1.1
## C3_Carbohydrate  0.91  0.07 -0.01  0.05 -0.03 -0.02 0.83 0.169 1.0
## C3_Energy        0.89 -0.03  0.03 -0.04  0.03 -0.06 0.81 0.192 1.0
## C3_Fat           0.46  0.01 -0.06  0.71  0.20  0.22 0.81 0.193 2.1
## C3_Protein       0.78  0.18 -0.11  0.26 -0.03  0.21 0.77 0.230 1.6
## C4_Carbohydrate  0.09  0.95  0.04  0.01 -0.11  0.02 0.94 0.064 1.1
## C4_Protein      -0.08  0.79  0.10  0.26 -0.05  0.20 0.76 0.240 1.4
## C4_Weight        0.15  0.84 -0.11 -0.05 -0.15 -0.09 0.77 0.230 1.2
## Fish_Energy     -0.03  0.14 -0.08  0.80 -0.36 -0.06 0.80 0.195 1.5
## Hair.d13C       -0.09 -0.04  0.85 -0.11  0.23 -0.16 0.81 0.186 1.3
```

```
## Hair.d15N       -0.01 -0.24  0.00 -0.13  0.89 -0.10 0.88 0.122 1.2
## Breath.d13C      0.02  0.06  0.88  0.00 -0.19 -0.07 0.83 0.174 1.1
##                       RC1  RC2  RC3  RC4  RC6  RC5
## SS loadings           2.50 2.37 1.58 1.33 1.10 1.07
## Proportion Var        0.21 0.20 0.13 0.11 0.09 0.09
## Cumulative Var        0.21 0.41 0.54 0.65 0.74 0.83
## Proportion Explained  0.25 0.24 0.16 0.13 0.11 0.11
## Cumulative Proportion 0.25 0.49 0.65 0.78 0.89 1.00
## Mean item complexity =  1.3
## Test of the hypothesis that 6 components are sufficient.
## The root mean square of the residuals (RMSR) is  0.06
## Fit based upon off diagonal values = 0.95

pc3<-principal(corMatrix3,nfactors=6,rotate="oblimin")
## Loading required namespace: GPArotation
pc3
## Principal Components Analysis
## Call: principal(r = corMatrix3, nfactors = 6, rotate = "oblimin")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                  TC1   TC2   TC3   TC4   TC5   TC6   h2    u2 com
## Animal_Weight   -0.02 -0.01 -0.06 -0.06  0.96 -0.07 0.94 0.061 1.0
## C3_Carbohydrate  0.91  0.03  0.01  0.00 -0.03 -0.04 0.83 0.169 1.0
## C3_Energy        0.92 -0.06  0.04 -0.08 -0.06  0.00 0.81 0.192 1.0
## C3_Fat           0.34 -0.03 -0.01  0.71  0.16  0.27 0.81 0.193 1.9
## C3_Protein       0.73  0.13 -0.06  0.20  0.17  0.00 0.77 0.230 1.4
## C4_Carbohydrate  0.04  0.97  0.03 -0.05 -0.01 -0.02 0.94 0.064 1.0
## C4_Protein      -0.18  0.79  0.12  0.22  0.17  0.06 0.76 0.240 1.4
## C4_Weight        0.12  0.85 -0.14 -0.11 -0.13 -0.08 0.77 0.230 1.2
## Fish_Energy     -0.11  0.03 -0.06  0.82 -0.14 -0.29 0.80 0.195 1.4
## Hair.d13C       -0.05  0.00  0.83 -0.06 -0.09  0.22 0.81 0.186 1.2
## Hair.d15N       -0.04 -0.06 -0.03 -0.07 -0.10  0.89 0.88 0.122 1.1
## Breath.d13C      0.07  0.01  0.89  0.01  0.00 -0.21 0.83 0.174 1.1
##                       TC1  TC2  TC3  TC4  TC5  TC6
## SS loadings           2.44 2.35 1.56 1.37 1.11 1.11
## Proportion Var        0.20 0.20 0.13 0.11 0.09 0.09
## Cumulative Var        0.20 0.40 0.53 0.64 0.74 0.83
## Proportion Explained  0.25 0.24 0.16 0.14 0.11 0.11
## Cumulative Proportion 0.25 0.48 0.64 0.78 0.89 1.00
##  With component correlations of
##       TC1   TC2   TC3   TC4   TC5   TC6
## TC1  1.00  0.10 -0.06  0.19  0.12  0.06
## TC2  0.10  1.00  0.02  0.17  0.13 -0.29
## TC3 -0.06  0.02  1.00 -0.08 -0.24  0.03
## TC4  0.19  0.17 -0.08  1.00  0.22 -0.15
## TC5  0.12  0.13 -0.24  0.22  1.00 -0.05
## TC6  0.06 -0.29  0.03 -0.15 -0.05  1.00
## Mean item complexity =  1.2
## Test of the hypothesis that 6 components are sufficient.
```

```
## The root mean square of the residuals (RMSR) is  0.06
## Fit based upon off diagonal values = 0.95
```

D.2.6 COMPARISON WITH OTHER STUDIES

The last graph (Figure 7.21) produced, contained the carbon and nitrogen isotopic results of this study compared up to other studies.

```
Comparison$inter <- with(Comparison, interaction(Tissue,  Diet))
#creates dummy variable to allow coding for 3 variables
ggplot(Comparison, aes(x=d13C, y=d15N)) +
 geom_point(aes(shape=inter,color=Continent), size=2) +
 scale_shape_manual(name="Shape", values=c(16,21,15,22,17,24))+
 scale_color_manual(values=c("#cc79a7","#0072b2","#56b4e9",
                             "#009e73","#e69f00","#d55e00",
                             "#f0e442"))+
 scale_x_continuous(name=expression(delta^13*C*" "("\u2030")),
                    limits=c(-23.5,-13),
                    breaks=seq(-23,-13,2),
                    labels=fmt_decimals(1))+
 scale_y_continuous(name=expression(delta^15*N*" "("\u2030")),
                    limits=c(6,16.5),
                    breaks=seq(6,16,2),
                    labels=fmt_decimals(1))+
 coord_fixed()+#makes x and y proportional
 my_theme

## R version 3.3.1 (2016-06-21)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 15063)
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] tcltk     stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
##  [1] compute.es_0.2-4  WRS2_0.9-1        rrcov_1.4-3
##  [4] robustbase_0.92-7 biotools_3.0      SpatialEpi_1.2.2
##  [7] sp_1.2-4          lattice_0.20-33   tkrplot_0.0-23
## [10] rpanel_1.1-3      psych_1.7.5       multcomp_1.4-6
## [13] TH.data_1.0-7     MASS_7.3-47       survival_2.41-3
## [16] mvtnorm_1.0-5     mvnormtest_0.1-9  mvoutlier_2.0.8
## [19] sgeostat_1.0-27   pander_0.6.0      pgirmess_1.6.5
```

```
## [22] broom_0.4.1       gridExtra_2.2.1   car_2.1-3
## [25] chron_2.3-50      aplpack_1.3.0     tidyr_0.6.3
## [28] dplyr_0.5.0       ggplot2_2.2.1
##
## loaded via a namespace (and not attached):
##  [1] nlme_3.1-128         pbkrtest_0.4-6       gmodels_2.16.2
##  [4] RColorBrewer_1.1-2   prabclus_2.2-6       rprojroot_1.2
##  [7] tools_3.3.1          backports_1.0.5      rgdal_1.2-5
## [10] R6_2.2.0             rgeos_0.3-22         DBI_0.5-1
## [13] lazyeval_0.2.0       mgcv_1.8-17          colorspace_1.2-7
## [16] trimcluster_0.1-2    nnet_7.3-12          splancs_2.01-39
## [19] GGally_1.3.0         mnormt_1.5-5         quantreg_5.33
## [22] SparseM_1.72         expm_0.999-1         sROC_0.1-2
## [25] sandwich_2.3-4       labeling_0.3         diptest_0.75-7
## [28] scales_0.4.1         DEoptimR_1.0-6       lmtest_0.9-35
## [31] mc2d_0.1-18          stringr_1.2.0        digest_0.6.10
## [34] foreign_0.8-66       minqa_1.2.4          rmarkdown_1.6
## [37] htmltools_0.3.6      lme4_1.1-12          zoo_1.7-13
## [40] mclust_5.2.2         gtools_3.5.0         spdep_0.6-11
## [43] magrittr_1.5         modeltools_0.2-21    Matrix_1.2-6
## [46] Rcpp_0.12.11         munsell_0.4.3        stringi_1.1.2
## [49] yaml_2.1.14          cvTools_0.3.2        flexmix_2.3-13
## [52] plyr_1.8.4           grid_3.3.1           maptools_0.9-1
## [55] parallel_3.3.1       pls_2.6-0            gdata_2.17.0
## [58] deldir_0.1-12        splines_3.3.1        knitr_1.16
## [61] boot_1.3-18          fpc_2.1-10           codetools_0.2-14
## [64] reshape2_1.4.2       stats4_3.3.1         LearnBayes_2.15
## [67] evaluate_0.10        data.table_1.10.4    laeken_0.4.6
## [70] vcd_1.4-3            nloptr_1.0.4         MatrixModels_0.4-1
## [73] VIM_4.6.0            gtable_0.2.0         kernlab_0.9-25
## [76] reshape_0.8.6        assertthat_0.1       e1071_1.6-8
## [79] coda_0.19-1          class_7.3-14         pcaPP_1.9-61
## [82] robCompositions_2.0.3 tibble_1.2          cluster_2.0.4
```