POT1 mutations predispose to familial melanoma

Carla Daniela Robles-Espinoza[1]‡, Mark Harland[2]‡, Andrew J. Ramsay[3]‡, Lauren G. Aoude[4]‡, Víctor Quesada[3], Zhihao Ding[1], Karen A. Pooley[5], Antonia L. Pritchard[4], Jessamy C. Tiffen[1], Mia Petljak[1], Jane M. Palmer[4], Judith Symmons[4], Peter Johansson[4], Mitchell S. Stark[4], Michael G. Gartside[4], Helen Snowden[2], Grant W. Montgomery[6], Nicholas G. Martin[7], Jimmy Z. Liu[8], Jiyeon Choi[9], Matthew Makowski[9], Kevin M. Brown[9], Alison M. Dunning[10], Thomas M. Keane[1], Carlos López-Otín[3], Nelleke A. Gruis[11], Nicholas K. Hayward[4]†, D. Timothy Bishop[2]†, Julia A. Newton-Bishop[2]†, David J. Adams[1]†*

‡ These authors contributed equally

† These authors are co-senior authors

Affiliations:

[1]Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA. UK.

[2]Section of Epidemiology and Biostatistics, Institute of Cancer and Pathology, University of Leeds, Leeds, LS9 7TF, UK.

[3]Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología del Principado de Asturias (IUOPA) Universidad de Oviedo, Oviedo, Spain.

[4]Oncogenomics Laboratory, QIMR Berghofer Medical Research Institute, Herston, Brisbane, QLD, Australia.

[5]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

[6]Molecular Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Herston, Brisbane, QLD, Australia.

[7]Genetic Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Herston, Brisbane, QLD, Australia.

[8]Statistical Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA. UK.

[9]Laboratory of Translational Genomics, National Cancer Institute, Bethesda, MD, USA.

[10]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK

[11]Leiden University Medical Centre, Leiden, The Netherlands.

Correspondence to:

Dr. David Adams
Experimental Cancer Genetics
Wellcome Trust Sanger Institute
Hinxton, Cambridge, CB10 1HH
Ph: +44 1223 83496862
Email: da1@sanger.ac.uk

Abstract: We identify families in which early onset multiple primary melanoma co-segregates with inactivating mutations in the protection of telomeres 1 (*POT1*) gene. We show that missense mutations of *POT1* in three families alter key residues of the oligonucleotide-/oligosaccharide-binding (OB) domains, disrupting telomere binding, leading to increased telomere length. Our results suggest *POT1* mutations strongly predispose to melanoma formation.

Main text: Mutations in *CDKN2A* account for 40-50% of familial melanoma cases[1], with rare mutations in *CDK4, BRCA2*, *BAP1*, and the promoter of *TERT,* also linked to the disease[2-5]. Here we set out to identify high-penetrance genes accounting for the ~50% of cases of familial melanoma that cannot be attributed to mutations in known genes.

To achieve this, we sequenced 184 melanoma cases from 105 pedigrees (168 exomes and 16 whole genomes) recruited in the United Kingdom, the Netherlands, and Australia that had been screened and found negative for pathogenetic variants in *CDKN2A* and *CDK4*. The patients sequenced came from pedigrees with between two and eleven cases of melanoma (169 cases), or were single cases that presented with either multiple primary melanomas (MPM), multiple primary cancers, one of which was melanoma, and/or an early age of onset (<4th decade) (15 cases) (**Supplementary Methods**, **Supplementary Tables 1-2**). Two-case families were preferentially selected for those enriched with cases of MPM. Exome capture and sequencing resulted in on average 84% of target bases being covered ≥ 10x across the autosomes and sex chromosomes. Whole genomes were sequenced to at least 27x mapped coverage. Focusing on coding regions we called nucleotide variants and filtered these data removing positions found in Phase 1 of the 1000 Genomes Project[6], dbSNP build 135 (see URLs), and variants called from 805 in-house control exomes. No known pathogenetic variants in *BAP1* or *BRCA2* were found, while capillary sequencing of the *TERT* promoter identified a six-case pedigree carrying the -57bp mutation[2] that was excluded from further analysis. Co-segregating variants found in all members of a sequenced pedigree, that were predicted to affect protein structure or function, were taken forward for further analysis (**Supplementary Methods**).

The abovementioned analysis left 23,051 novel variants that were either private to a pedigree or individual. We therefore focused on genes with co-segregating variants from the 26 pedigrees for which we had sequence data for three or more members (314 genes) (**Supplementary Table 3**). Of particular interest were 6 genes (defined by GO/PubMed analysis) linked to biological processes such as G_1 cell cycle progression or the regulation of telomere length or integrity, since high-penetrance mutations in *CDKN2A* and *CDK4*, and genetic variation in the vicinity of *TERT, ATM* and *PARP1* has been associated with melanoma susceptibility[7-9]. Analysis of missense and disruptive variants (nonsense, essential splice site and frameshift mutations) led us to identify a 5-case pedigree (id: UF20) carrying a Tyr89Cys change (GRCh37 chr. 7: g. 124503684T>C) in the highly conserved N-terminal oligonucleotide-/oligosaccharide-binding (OB) domain of the *POT1* (protection of telomeres 1) gene product[10], and then a further two families with non-synonymous OB domain changes (id: UF31, Gln94Glu, g. 124503670G>C and id: UF23, Arg273Leu, g. 124493077C>A) (**Fig. 1, Table 1, Supplementary Figs. 1&2a-c, Supplementary Tables 1, 4-5**). Remarkably, Gln94 has been found to be recurrently somatically mutated (Gln94Arg) in chronic lymphocytic leukemia (CLL), where ~5% of cases carry *POT1* mutations that cluster in the OB domains[11]. We also identified a 6-case family (id: AF1), 3 members of which were sequenced, who carried an invariant essential splice acceptor mutation in intron 17-18 (g.124465412C>T). This variant was scored as deleterious by the MaxEntScan algorithm[12] and was shown to affect transcript splicing by RT-PCR and sequencing (**Fig. 1, Table 1, Supplementary Methods, Supplementary Figs. 2d-e&3, Supplementary Tables 4-5**). Pedigrees UF20 and AF1 are among our largest families, suggesting a strong association between *POT1* mutations and melanoma formation. Importantly, the identification of *POT1* mutations in 4/105 melanoma families represents a statistically significant enrichment (P=0.0033) of variants compared to a control dataset of 520 exomes in which we found only one missense mutation passing our filters, and this was located outside the OB domains of POT1 (**Supplementary Methods, Supplementary Fig. 4, Supplementary Table 6**). None of the four variants mentioned above (three missense and the splice acceptor mutation) were found by genotyping 2,402 additional controls (**Supplementary Methods, Supplementary Table 6**). We also genotyped these positions across 1,739 population-based melanoma cases from a case-control series. We found one case with MPM and early onset, who carried the Arg273Leu

variant, and another case who carried the essential splice site mutation, whose sibling was also a melanoma patient who was diagnosed in the 5th decade (**Supplementary Methods**).

All of the missense variants identified in *POT1* disrupt amino acids completely conserved throughout Eutherian evolution (**Fig. 1b**) and code for amino acids that are more evolutionarily conserved than the average for other OB domain residues (**Supplementary Methods**). Importantly, all 9 confirmed mutation carriers from the familial cohort developed melanoma, with their presentation ranging from 1 primary melanoma (4 cases) to 8 melanomas, from 25 to 80 years of age**.** One mutation carrier from these familial cases also developed breast cancer at 65 and another small cell lung cancer at 50 (pedigree UF20), while 4 of the 12 untested first degree relatives of mutation carriers, who were over the age of 30 years or deceased, were reported to have developed cancer. These malignancies included melanoma (pedigrees UF20 and UF31), endometrial cancer (pedigree UF20) and brain tumors (pedigrees UF20 and UF23). Intriguingly, the pedigree with the splice acceptor mutation (AF1) had a patient with a history of melanoma and CLL, in keeping with a role for *POT1* in CLL development[11]. Collectively, these data suggest a possible role for germline *POT1* mutations in susceptibility to a range of cancers in addition to melanoma.

To test whether the effect of the identified missense variants might be similar to that of the somatic mutations found in CLL in that they disrupt telomere binding, we examined the structure of POT1 bound to a telomere-like polynucleotide (dTUdAdGdGdGdTdTdAdG) (PDB 3KJP)[13,14]. According to this model, all three mutated residues (Tyr89, Gln94 and Arg273) are amongst 24 residues located in close proximity (<3.5Å) to the telomeric polynucleotide[11] (**Fig. 2a**). Arg273 interacts with the oxygen at position 2 of the telomeric dT7, whereas Gln94 and Tyr89 both interact with the G deoxynucleotide at position 4. Therefore, as described for the somatic mutations in CLL, the POT1 variants we identified are expected to weaken or abolish the interaction between POT1 and telomeres. Analysis of nucleotides coding for these 24 OB domain residues identified one non-synonymous change in 6,498 control exomes[15] compared to 3 in 105 families, a highly significant enrichment of mutations in the melanoma cohort (P=1.54x10^{-5}) (**Supplementary Methods, Supplementary Tables 6-7**).

To test formally that the OB mutations we identified disrupt POT1 function we assessed the ability of *in vitro*-translated POT1 Tyr89Cys, Gln94Glu and Arg273Leu proteins to bind to $(TTAGGG)_3$ sequences, with electromobility shift assays revealing a complete abolition of mutant POT1-DNA complex formation (**Fig. 2b, Supplementary Methods, Supplementary Fig. 5**). Importantly, the POT1 Tyr36Asn and Tyr223Cys variants recently described in CLL by Ramsay *et al*[11], which appear to be functionally analogous to the variants we describe here, promote uncapping of telomeres, telomere length extension and chromosomal aberrations and thus promote tumorigenesis[11]. Using DNA sequence from 41 patients for whom we had unmapped read data we estimated telomere length by counting TTAGGG repeats (See Ref. 16). All three members of pedigree UF20, carrying the Tyr89Cys mutation, were found to have telomeres that were as much as 3.5-fold longer than *POT1* wild-type controls (P<0.0002, **Fig. 2c**). This result was confirmed by telomere length PCRs, which also showed longer telomeres for patients carrying the Gln94Glu and Arg273Leu variants (**Fig. 2d**, **Supplementary Methods**). Thus, missense mutations in the OB domains of POT1 not only abolish telomere binding, they are also associated with increased telomere length, a key factor influencing risk of melanoma[17]. Importantly, OB domain variants that disrupt the interaction of POT1 with telomeric ssDNA are thought to function as dominant-negative alleles[11,18], yet as we show here are compatible with life, suggesting additional somatic events are required to promote tumorigenesis. Analysis of somatic variants within POT1 from the COSMIC[19] and IntOGen[20] databases (TCGA/ICGC data) across 14 cancer types show a tendency for variants to result in missense mutations (P<0.03), to hit residues in close proximity to DNA (P<0.02), and to have a high functional bias (P<0.03) (**Supplementary Methods**). This result suggests that although rare, somatic *POT1* mutations may drive tumorigenesis across multiple histologies.

Here we describe germline mutations in the gene encoding the telomere-associated protein POT1 in almost 4% of *CDKN2A*/*CDK4*-negative familial melanoma pedigrees and in 2/34 (5.8%) 5+ case pedigrees, making *POT1* the second most frequently mutated high-penetrance familial melanoma gene reported to date. In combination with the recently described *TERT* promoter mutation[2], these findings significantly extend our understanding of a novel mechanism predisposing to the development of familial melanoma. Since the dysregulation of telomere protection by POT1 has recently been identified as a target for potential therapeutic

intervention[21], in the future, it may be possible that early identification of families with *POT1* mutations may facilitate better management of their disease.

Methods

Methods and associated references are available in the online version of the paper.

Figure legends

Fig. 1. **Mutations in *POT1* predispose to familial melanoma. a)** We identified four pedigrees carrying deleterious mutations in the protection of telomeres 1 (*POT1*) gene. Shown are a 5-(UF20) and a 6-(AF1) member pedigree carrying the disruptive Tyr89Cys OB domain mutation and an essential splice acceptor mutation, respectively. Please note that pedigrees have been adjusted to protect the identity of the families without a loss of scientific integrity. CMM; cutaneous malignant melanoma. CLL; chronic lymphocytic leukemia. The patients that were sequenced have a red outline. All melanomas were confirmed by histological analysis with the exception of two cases (*). The number of primary melanomas in each patient is indicated. **b)** Highly conserved residues of POT1 are mutated in familial melanoma. Shown are the positions of the mutations identified on the POT1 protein (top panel), and on an amino acid alignment (missense mutations, bottom panel).

Fig. 2. Missense mutations in *POT1* disrupt the interaction between POT1 and single-stranded DNA and lead to elongated telomeres. a) Shown is the location of the POT1 residues Tyr89, Gln94 and Arg273 in the N-terminal two oligonucleotide-/oligosaccharide-binding (OB) domains, in green. A telomere-like polynucleotide sequence is shown in orange. All three substitutions are predicted to disrupt the association of POT1 with telomeres. **b)** Mutant Tyr89Cys, Gln94Glu and Arg273Leu POT1 proteins are unable to bind telomeric $(TTAGGG)_3$ sequences as revealed by an electromobility shift assay. The Tyr223Cys POT1 mutant was used as a positive control representing a known disruptive mutation[11]. **c)** Calculation of telomere length from exome sequence data. The method used is analogous to the one described in Ref. 16. Relative telomere lengths for the three sequenced members of pedigree UF20 are shown

alongside the mean telomere length of 38 (all) other melanoma cases who were sequenced alongside, but were wildtype for *POT1*. Error bars indicate one standard deviation. A Wilcoxon rank sum test was performed comparing the telomere length of the 3 Tyr89Cys cases to the 38 non-carrier controls. **d)** PCR-based estimate of telomere length. Negative mean ΔCt values, which correlate positively with telomere length, for POT1 missense mutation carriers and non-carrier family controls are shown against a distribution of values from 258 POT1 non-mutation carrier controls (**Supplementary Methods**). The black line represents a Gaussian kernel density estimate for this set using Silverman's "rule-of-thumb"[22] as the smoothing bandwidth. Orange dots represent members of pedigree UF20; pink, UF31; blue, UF23; red, individual CT1663 from the Leeds Melanoma Case-control study carrying the Arg273Leu variant (**Supplementary Table 5**). The number of biological replicates for each case ranged from 1 to 4, each with two technical replicates for the POT1 missense mutation carriers and non-carrier family controls. Between one and two technical replicates were performed for the 258 POT1 non-mutation carrier controls. Error bars indicate the standard error of the mean.

Table 1. *POT1* mutations identified in this study.

Pedigree	Num. cases in pedigree	Num. of carriers / num. of tested cases	Mutation (genomic coordinates)	HGVS name*	Exon	Amino acid change	Mutation type	Bioinformatic prediction tools		
								SIFT[23]	PolyPhen 2[24]	CAROL[25]
UF20	5	4/4	g.124503684T>C	c.266A>G	8	Tyr89Cys	Missense	Deleterious	Probably damaging	Deleterious
UF31	2	1/1	g.124503670G>C	c.280C>G	8	Gln94Glu	Missense	Tolerated	Probably damaging	Deleterious
UF23	2	1/2**	g.124493077C>A	c.818G>T	10	Arg273Leu†	Missense	Deleterious	Probably damaging	Deleterious
AF1	6	3/3	g.124465412C>T	c.1687-1G>A	-	-††	Splice acceptor (intronic, between ex. 17-18)	-	-	-

*The reference transcript, taken from the Ensembl database (release 70) is POT1-001 (ENST00000357628).

**A second case within this pedigree had a different clinical presentation (solitary melanoma *in situ*) in the 6th decade and did not carry the Arg273Leu variant.

†This variant was also detected in a melanoma case from population-based case-control series that presented with MPMs and an early age of onset (see **Supplementary Methods**).

††This variant was also detected in a case from population-based case-control series that presented with melanoma and who had a sibling with a previous diagnosis of the disease (see **Supplementary Methods**).

URLs

dbSNP135, http://www.ncbi.nlm.nih.gov/projects/SNP/; European Genome-phenome Archive, http://www.ebi.ac.uk/ega/.

Accession codes

Sequence data has been deposited at the European Genome-phenome Archive (EGA), hosted at the European Bioinformatics Institute, under accession EGAS00001000017.

Acknowledgments: D.J.A, C.D.R-E, J.Z.L., J.C.T, M.P and T.M.K were supported by Cancer Research UK and the Wellcome Trust. C.D.R-E was also supported by Consejo Nacional de Ciencia y Tecnología of Mexico. N.K.H. was supported by a fellowship from the National Health and Medical Research Council of Australia (NHMRC). L.G.A was supported by an ANZ Trustees PhD scholarship. A.P is supported by Cure Cancer Australia. The work was funded in part by the NHMRC and the Dutch Cancer Society (UL 2012-5489). M.H. J.A.N-B, D.T.B were supported by Cancer Research UK (Programme Awards C588/A4994, C588/A10589 and the Genomics Initative). C.L-O, A.J.R and V.Q are funded by the Spanish Ministry of Economy and Competitiveness through the Instituto de Salud Carlos III (ISCIII), Red Temática de Investigación del Cáncer (RTICC) del ISCIII and the Consolider-Ingenio RNAREG Consortium. CL-O is an Investigator with the Botín Foundation. We thank the UK10K Consortium (funded by the Wellcome Trust; WT091310) for access to control data.

Author contributions

C.D.R-E. M.H., J N-B., D.T.B., N.K.H. and D.J.A. designed the study wrote the paper. C.D.R-E., M.H., L.G.A., J.C.T., M.M., J.C., M.P., A.J.R., Z.D., V.Q., A.P., J.P., J.S., M.S., M.G., A.M.D., K.A.P., P.J., J.Z.L., K.M.B., C.L-O. and T.M.K. performed experiments or analysis. N.A.G. G.M., and N.M provided vital biological resources.

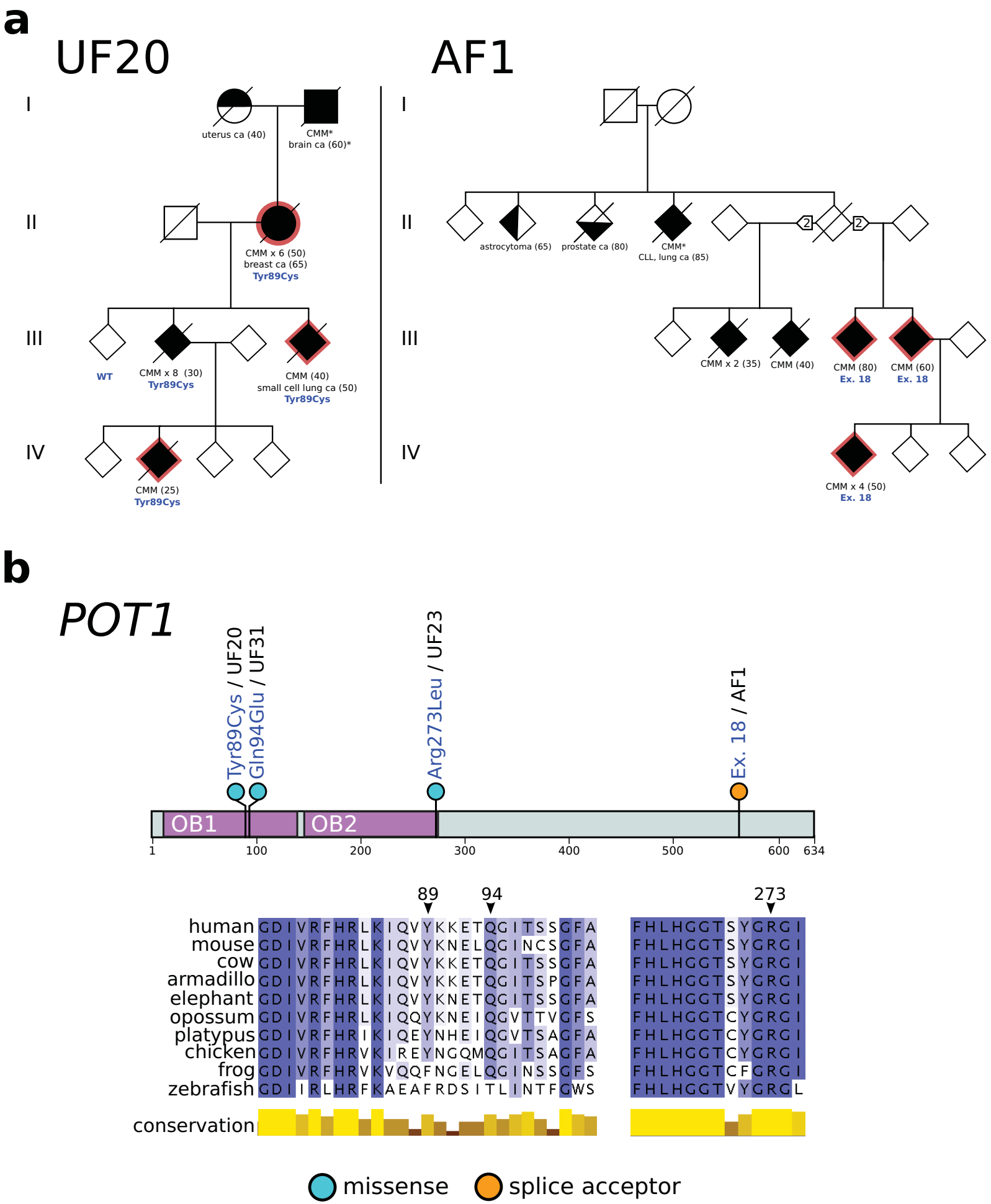
Competing financial interests

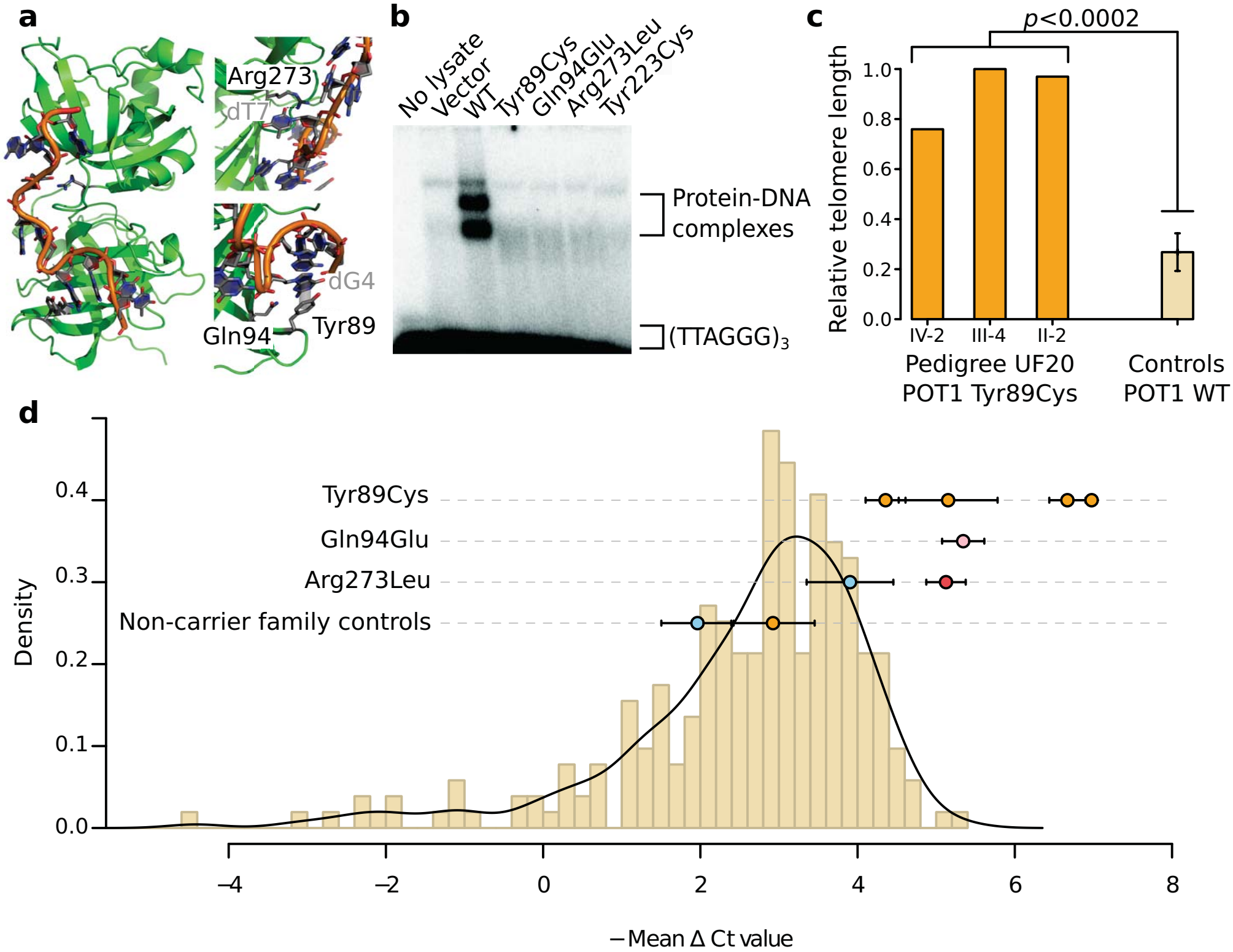
The authors declare no competing financial interests.

References

1. Goldstein, A.M. *et al. J Med Genet* **44**, 99-106 (2007).
2. Horn, S. *et al. Science* **339**, 959-61 (2013).
3. *J Natl Cancer Inst* **91**, 1310-6 (1999).
4. Wiesner, T. *et al. Nat Genet* **43**, 1018-21 (2011).
5. Zuo, L. *et al. Nat Genet* **12**, 97-9 (1996).
6. The 1000 Genomes Project Consortium *et al. Nature* **491**, 56-65 (2012).
7. Barrett, J.H. *et al. Nat Genet* **43**, 1108-13 (2011).
8. Bodelon, C. *et al. PLoS One* **7**, e52466 (2012).
9. Nan, H., Qureshi, A.A., Prescott, J., De Vivo, I. & Han, J. *Hum Genet* **129**, 247-53 (2011).
10. Baumann, P. & Cech, T.R. *Science* **292**, 1171-5 (2001).
11. Ramsay, A.J. *et al. Nat Genet* (2013).
12. Yeo, G. & Burge, C.B. *J Comput Biol* **11**, 377-94 (2004).
13. Nandakumar, J., Podell, E.R. & Cech, T.R. *Proc Natl Acad Sci U S A* **107**, 651-6 (2010).
14. Lei, M., Podell, E.R. & Cech, T.R. *Nat Struct Mol Biol* **11**, 1223-9 (2004).
15. Fu, W. *et al. Nature* **493**, 216-20 (2013).

16. Parker, M. *et al. Genome Biol* **13**, R113 (2012).
17. Burke, L.S. *et al. PLoS One* **8**, e71121 (2013).
18. Kendellen, M.F., Barrientos, K.S. & Counter, C.M. *Mol Cell Biol* **29**, 5611-9 (2009).
19. Forbes, S.A. *et al. Nucleic Acids Res* **39**, D945-50 (2011).
20. Gonzalez-Perez, A. *et al. Nat Methods* (2013).
21. Jacobs, J.J. *Front Oncol* **3**, 88 (2013).
22. Silverman, B.W. *Density Estimation*, (Chapman and Hall, London, 1986).
23. Ng, P.C. & Henikoff, S. *Nucleic Acids Res* **31**, 3812-4 (2003).
24. Adzhubei, I.A. *et al. Nat Methods* **7**, 248-9 (2010).
25. Lopes, M.C. *et al. Hum Hered* **73**, 47-51 (2012).


a
UF20
AF1
I
II
III
IV
uterus ca (40)
CMM*
brain ca (60)*
CMM x 6 (50)
breast ca (65)
Tyr89Cys
WT
CMM x 8 (30)
Tyr89Cys
CMM (40)
small cell lung ca (50)
Tyr89Cys
CMM (25)
Tyr89Cys
astrocytoma (65)
prostate ca (80)
CMM*
CLL, lung ca (85)
2
2
CMM x 2 (35)
CMM (40)
CMM (80)
Ex. 18
CMM (60)
Ex. 18
CMM x 4 (50)
Ex. 18
b
POT1
Tyr89Cys / UF20
Gln94Glu / UF31
Arg273Leu / UF23
Ex. 18 / AF1
OB1
OB2
1
100
200
300
400
500
600
634
89
94
273
human
mouse
cow
armadillo
elephant
opossum
platypus
chicken
frog
zebrafish
conservation
missense
splice acceptor


a
Arg273
dT7
dG4
Gln94
Tyr89
b
No lysate
Vector
WT
Tyr89Cys
Gln94Glu
Arg273Leu
Tyr223Cys
Protein-DNA complexes
(TTAGGG)3
c
p<0.0002
Relative telomere length
1.0
0.8
0.6
0.4
0.2
0.0
IV-2
III-4
II-2
Pedigree UF20
POT1 Tyr89Cys
Controls
POT1 WT
d
Density
0.4
0.3
0.2
0.1
0.0
Tyr89Cys
Gln94Glu
Arg273Leu
Non-carrier family controls
–4
–2
0
2
4
6
8
−Mean Δ Ct value