

Research

Whole-genome analysis of animal A- and B-type cyclins

Conrad A Nieduszynski^{*†}, James Murray^{*} and Mark Carrington^{*}

Addresses: ^{*}Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK. [†]Current address: Cancer Research UK Chromosome Replication Group, School of Life Sciences, Dundee University, Dundee DD1 5BH, UK.

Correspondence: Mark Carrington. E-mail: mc115@cam.ac.uk

Published: 15 November 2002

Genome **Biology** 2002, **3**(12):research0070.1–0070.8

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2002/3/12/research/0070>

© 2002 Nieduszynski et al., licensee BioMed Central Ltd
(Print ISSN 1465-6906; Online ISSN 1465-6914)

Received: 30 May 2002

Revised: 2 September 2002

Accepted: 3 October 2002

Abstract

Background: Multiple A- and B-type cyclins have been identified in animals, but their study is complicated by varying degrees of functional redundancy. A non-essential phenotype may reflect redundancy with a known or as yet unknown gene. Complete sequencing of several animal genomes has allowed us to determine the size of the mitotic cyclin gene family and therefore to start to address this issue.

Results: We analyzed the *Caenorhabditis elegans*, *Drosophila melanogaster* and *Homo sapiens* genomes to identify known and novel A- and B-type cyclin genes and distinguish them from related pseudogenes. We find only a single functional A-type cyclin gene in invertebrates but two in vertebrates. In addition to the single functional cyclin A gene, the *C. elegans* genome contains numerous cyclin A pseudogenes. In contrast, the number and relationship of B-type cyclins varies considerably between organisms but all contain at least one cyclin B1-like gene and a cyclin B3 gene.

Conclusions: There are three conserved families of mitotic cyclins in animals: A-, B3- and B-type. The precise number of genes within the A- and B-type families varies in different organisms, possibly as an adaptation to their distinct developmental strategies.

Background

Progression through the eukaryotic cell cycle is controlled by the sequential activation and inactivation of cyclin-dependent kinases (CDKs). The kinase activity of CDKs is regulated by cyclin synthesis and destruction, phosphorylation of the kinase and cyclin, binding to inhibitory polypeptides and subcellular localization. In metazoa there are four cyclin types (A, B, D and E) known to regulate cell-cycle transitions. These four cyclins have multiple isotypes, the precise number varying between organisms. In mammals and amphibians, members of the A- and D-cyclin families have been shown to be expressed in distinct cell types. In contrast, B-type cyclins are ubiquitous, being co-expressed in most, if not all, proliferating cells.

Both A- and B-type cyclins have been shown to have an essential role in progression into, through and out of M-phase. In vertebrates, two A-type cyclins have been identified to date. Cyclin A1 is expressed during meiosis, in early cleaving embryos [1,2] and in some transformed cell lines [3]. Cyclin A2 is present in all proliferating cells from the beginning of S-phase until mitosis, with the exception of cells undergoing the meiotic division in males [4]. The reason for this restricted expression pattern is unclear but male mice lacking the cyclin A1 gene are sterile owing to an arrest in meiosis [5]. However, female mice lacking the cyclin A1 are fertile, possibly as a result of rescue by cyclin A2 which is expressed during female meiosis [6]. Deletion of murine cyclin A2 is lethal; embryos are able to undergo the

cleavage divisions but die immediately after implantation and appear unable to undergo proliferative growth [6]. Ablation of expression of cyclin A2 in cultured cells leads to an arrest in S-phase or at the end of G2 [7].

Cyclin B-dependent kinase activity is essential for entry into mitosis. The best characterized vertebrate B-type cyclins, B1 and B2, are co-expressed in the majority of dividing cells, but are differentially localized within the cell (reviewed in [8]). Gene-deletion experiments have shown that mice lacking the cyclin B1 gene die *in utero* whereas mice lacking the cyclin B2 gene develop normally and have no immediately apparent phenotype [9]. This suggests that loss of cyclin B2 can be compensated for by another cyclin, probably cyclin B1. Cyclin B3 is less well characterized, with a mammalian homolog only recently recognized [10]. In *Drosophila*, cyclin B3 is not required for mitosis [11].

Redundancy between members of each cyclin class is probably responsible for the non-essential phenotypes described above. This effect is best understood in budding yeast, *Saccharomyces cerevisiae*, which has six B-type cyclins that display functional redundancy, yet fulfill distinct roles [12]. To be better able to interpret the phenotypes of A- and B-type cyclin mutants in animals and the degree of redundancy that has to be taken into consideration, we have carried out a comprehensive search to identify A- and B-type cyclin genes in the completed *Caenorhabditis elegans*, *Drosophila melanogaster* and *Homo sapiens* genomes. The search has identified two novel B-type cyclin genes in *C. elegans* and has shown that there is only a single functional A-type cyclin gene in both invertebrates.

Results

C. elegans: a single cyclin A gene and multiple cyclin A pseudogenes

On the basis of hybridization of a cyclin A cDNA to a Southern blot of genomic DNA it has been reported that *C. elegans* has multiple A-type cyclin genes [13]. To investigate this, the *C. elegans* genome sequence [14] was queried with cyclin A sequences. In addition to the previously characterized *C. elegans* cyclin A gene (ZK507.6) another 15 regions were found with 80 to 95% sequence identity to the query sequence at the nucleotide level. All 15 were subsequently identified as pseudogenes. The cyclin A pseudogenes can be divided into six types based on identity to parts of the known cyclin A gene, although all could be derived from a single precursor (Figure 1). The pseudogenes are all derived from the central part of the cyclin A gene and span both exons and introns and are thus non-processed pseudogenes. However, it is worth noting that in several of the pseudogenes some of the conserved splice-site consensus sequences are missing, such that if the message was transcribed it would not be correctly spliced. In addition, all contain deletions and mutations relative to the cyclin A gene, resulting not only in the

loss of sizeable regions of coding sequence, but also in frameshifts, point mutations and premature stop codons, and thus cannot encode a functional protein (Figure 1). None is represented in the *C. elegans* expressed sequence tag (EST) database.

The pseudogenes are found on all six chromosomes, and none is present in a tandem array. This contrasts with the arrangement of the major sperm protein (MSP), another *C. elegans* gene found in multiple copies. These genes (both transcribed and pseudogenes) are found in loose clusters of between 3 and 13 copies [15].

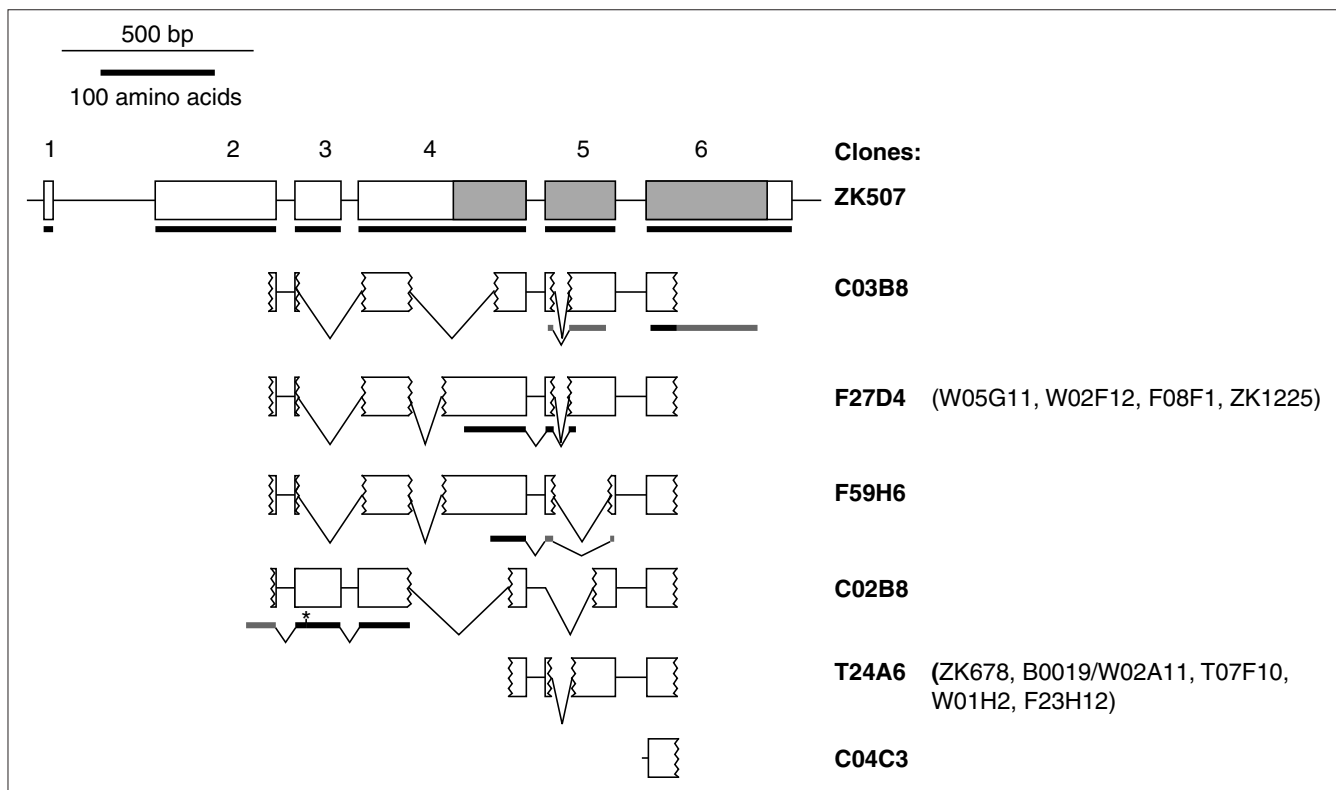
The cyclin A pseudogenes contain regions of sequence identity to each other which lie outside the region of identity to the cyclin A gene. These regions of identity extend over approximately 4.5 kilobase pairs (kbp) in total, and include direct and inverted repeats, some of which occur at other locations in the genome. This observation supports the possibility that the pseudogenes had a single origin and it is possible that the repeats have had a role in the transposition of the cyclin A pseudogenes around the genome.

C. elegans: multiple B-type cyclins

A search of the genome with several B-type cyclin query sequences identified a total of four genes: the previously characterized cyclins B (ZC168.4) and B3 (To6E6.2) [13] and two genes which appear to encode novel B-type cyclins (H31G24.4 and Y43E12A.1). The latter two encode polypeptides that have 87% identity to each other; the main difference between the two is the presence of an insert of 24 amino acids in H31G24.4 that is absent in Y43E12A.1. Both are more similar to cyclin B than to cyclin B3, and have the same exon and intron structure as the cyclin B gene. At the amino-acid level H31G24.4 has 60% identity to cyclin B and Y43E12A.1 has 65% identity. ESTs derived from both genes are present in the database: BE228087 from H31G24.4 and C44218 from Y43E12A.1. On the basis of the available EST sequences and homology to the known B-type cyclin, conceptual translations of the two newly identified genes were aligned with the known B-type cyclin (Figure 2). This alignment illustrates the high degree of sequence conservation between the B-type cyclins, particularly over the cyclin box. However, all four B-type cyclins differ in the size and sequence composition of their amino termini, and it may be this region which determines their individual cellular functions.

D. melanogaster: A- and B-type cyclins

The *D. melanogaster* genome sequence [16] was queried with the complete cyclin A protein sequence [17]. Equivalent searches were performed using *D. melanogaster* cyclin B [18] and cyclin B3 [11] protein sequences. Using all search permutations, only the three query sequences were identified. Therefore, it is highly probable that there is one A-type cyclin gene, a single cyclin B and a single cyclin B3 gene

**Figure 1**

Schematic representation of *C. elegans* cyclin A and cyclin A pseudogenes. The intron/exon structure of cyclin A (cosmid ZK507) is shown with the cyclin box shaded. Below this are shown the six different cyclin A pseudogene structures: regions with homology to exons are represented by boxes and regions homologous to introns by lines (with deletions also marked). The cosmid clones containing each of the genes are shown on the right with the genes actually represented on the schematic in bold. Note that cosmids B0019 and W02A11 overlap and represent only one pseudogene. For each schematic, potential open reading frames (ORFs) are represented by a bar. Black bars represent ORFs with high identity to cyclin A1 (> 70%) and gray bars represent ORFs with no significant identity to cyclin A1. ORFs were predicted by assuming that each pseudogene was fully transcribed and spliced in a manner analogous to the cyclin A gene. The largest ORF was identified in the pseudogene on clone C02B8 and would result in a 75- or 113-amino-acid protein depending on whether the first initiation codon within the pseudogene was used (asterisk) or an initiation codon just 5' of the pseudogene. Frameshifts and point mutations limit the extent of the identity and frequently result in termination codons.

within the available sequence. No pseudogenes derived from any of these three cyclin genes were identified.

***H. sapiens*: A-type cyclins**

The available human genome sequence (version 5.28.1) [19] was searched with cyclin A1 and A2 query sequences. In both cases the most significant matches were the two known A-type cyclins and the next most significant were the known B-type cyclins. Thus, there are only two A-type cyclin genes in the available sequence. No cyclin A-like pseudogenes were found.

***H. sapiens*: three B-type cyclins?**

Searches of the available human genome sequence using B-type cyclin queries revealed three B-type cyclins - B1, B2 and B3 - but no other genes identifiable as encoding a functional B-type cyclin. Specifically, no homologs of the *Xenopus* cyclins B4 and B5 [20] were found. Although the human genome sequence is not quite complete it seems unlikely that additional human B-type cyclins will be revealed.

Further evidence for this is provided by the lack of additional B-type cyclin sequences in the EST databases, in which cyclins B1, B2 and B3 were clearly present.

In addition to these three B-type cyclin genes, a region of high identity to cyclin B2 was found at chromosome 7p22.1. This region had the equivalent of 84% identity over 241 of the 398 amino acids of cyclin B2 and no identity outside this region. Thus the predicted open reading frame (ORF) contains only a fragment of the cyclin B2 gene and probably represents a pseudogene. There are no ESTs corresponding to this sequence in the database.

Discussion

The main findings of this study are first, that invertebrates have a single A-type cyclin whereas vertebrates have two, and second, that the number of B-type cyclins varies from two in *Drosophila* to three in humans and four in *C. elegans*,

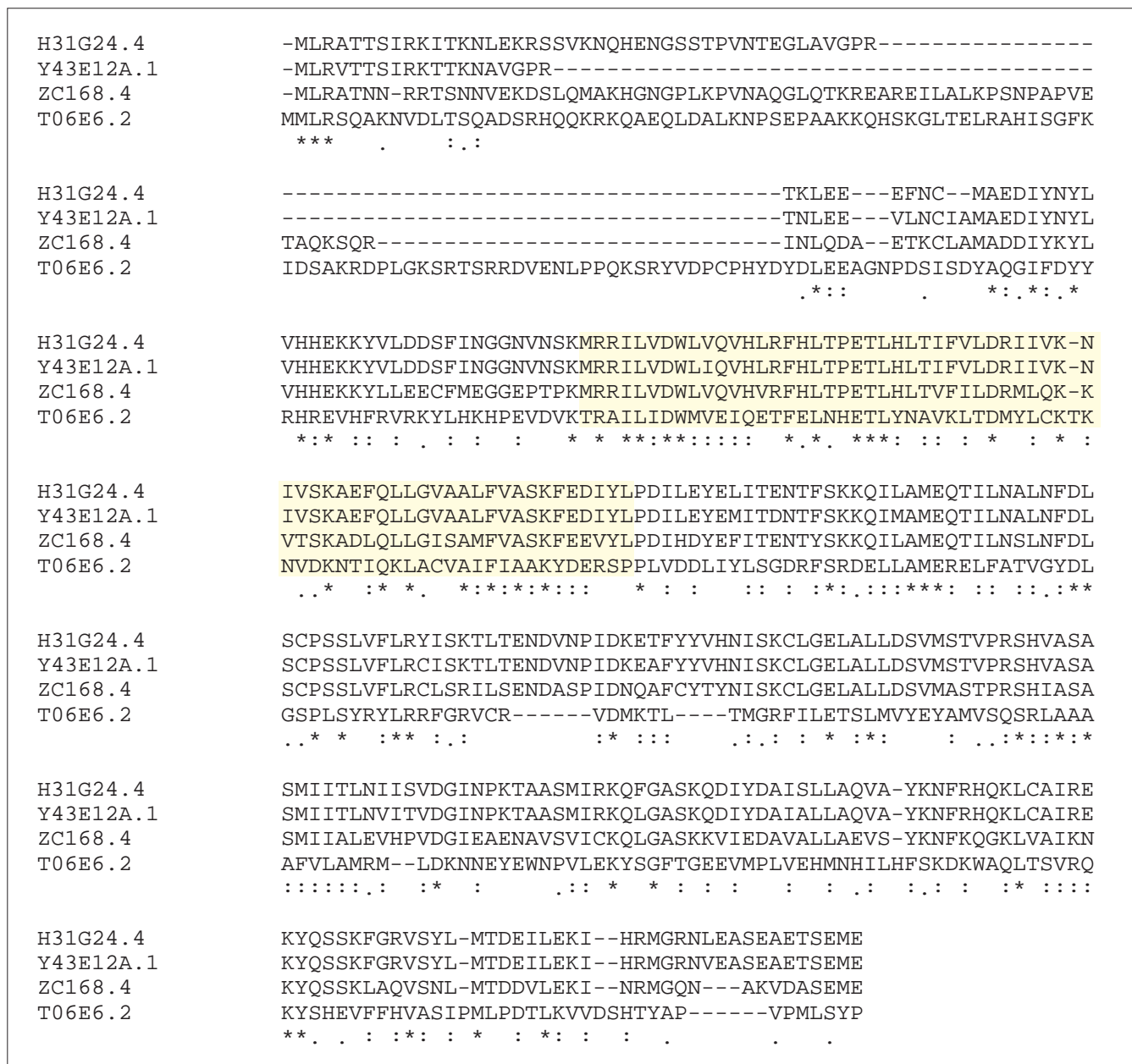


Figure 2
Alignment of *C. elegans* B-type cyclins. The known B-type and B3 cyclins (ZC168 and T06E6) were aligned to the two novel B-type cyclins (H31G24 and Y43E12A) using the GCG pileup program and refined using the CLUSTALW program. Below the sequences are marked identical (*), highly similar (:), and similar (.) amino acids. The cyclin box has been highlighted in yellow.

but all have at least one cyclin B1-like gene and a cyclin B3 gene.

A-type cyclins

In vertebrates, cyclin A1 is expressed predominantly in the germline during the meiotic divisions and early cleavage divisions of the embryo [1-3], whereas cyclin A2 is present in all proliferating cells. Near-completion of the human genome sequence has revealed that there are only two A-type cyclins.

In support of this finding an analysis of the extensive mouse and zebrafish EST databases and the draft pufferfish genome database identified cyclin A1 and A2 cDNAs and genes but no further A-type cyclins (data not shown). These sequences were used to construct a phylogenetic tree for all known A-type cyclins (Figure 3). From this tree it is clear that all known vertebrate A-type cyclins fall into one of two groups (A1 and A2) and there is no evidence for multiple invertebrate A-type cyclins. The analysis of the number of A-type cyclin

genes was initiated by the observation that mouse embryos lacking a cyclin A2 gene complete the normal number of cleavage divisions to produce an embryo of more than 100 cells before ceasing to proliferate further [6]. There are four possible explanations: first, cyclin A is not necessary for the cleavage divisions; second, there is persistence of maternal cyclin A mRNA; third, the cleavage divisions are rescued by cyclin A1 expression or (a fourth possibility) by expression of another unidentified A-type cyclin. Our findings reported here show that this last possibility can be eliminated.

The intriguing question now is why vertebrates have evolved a distinct A-type cyclin with a limited role within the germline. One possible explanation is that cyclin A1 is part of the larger phenomenon of the occurrence of germline isoforms of a large number of genes. A number of enzymes are already known to have testis-specific isoforms (reviewed in [21]), including glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase. In each case, testis-specific isoforms have been identified only in the vertebrate lineage (data not shown), suggesting that the second isoforms arose

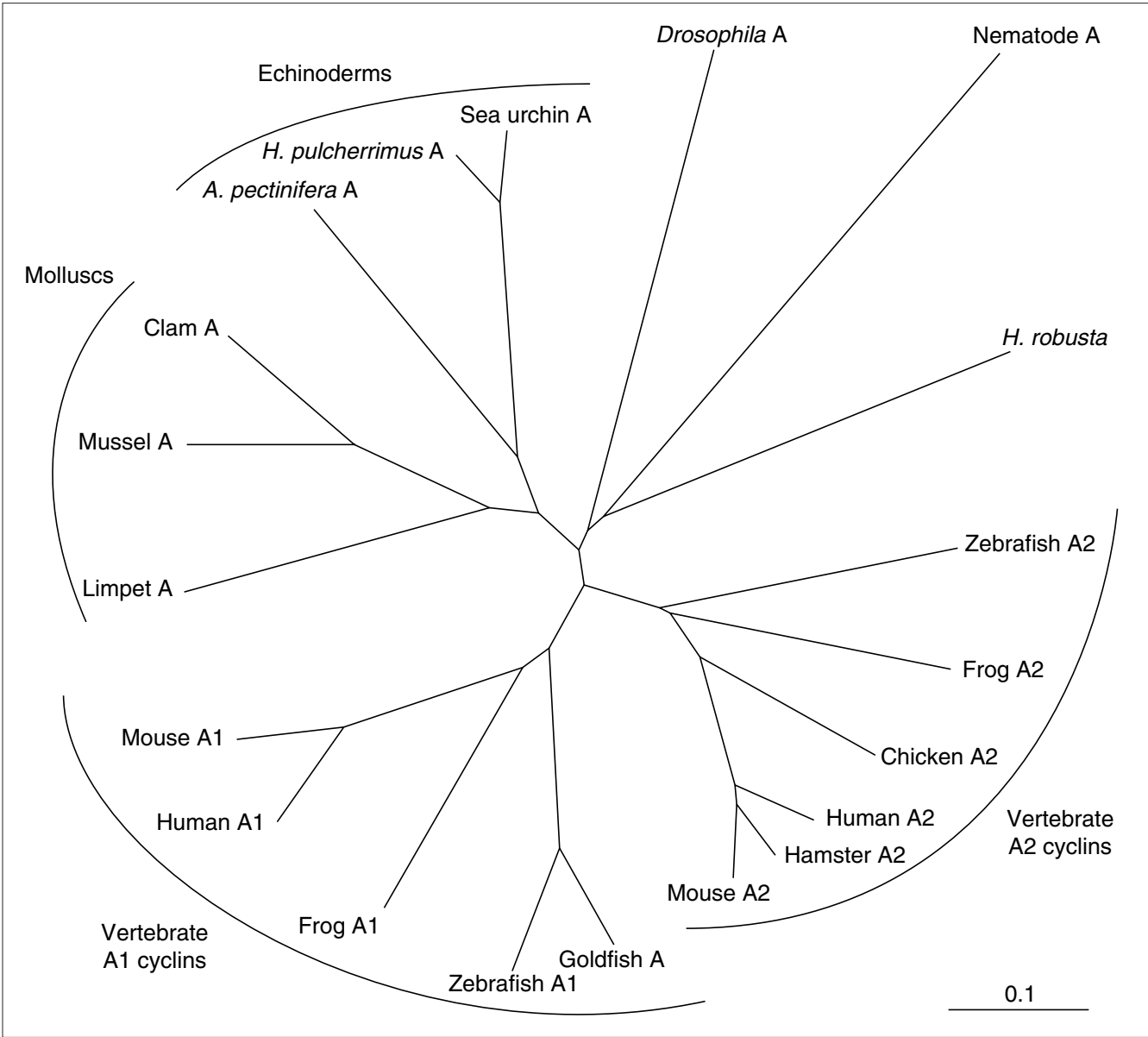
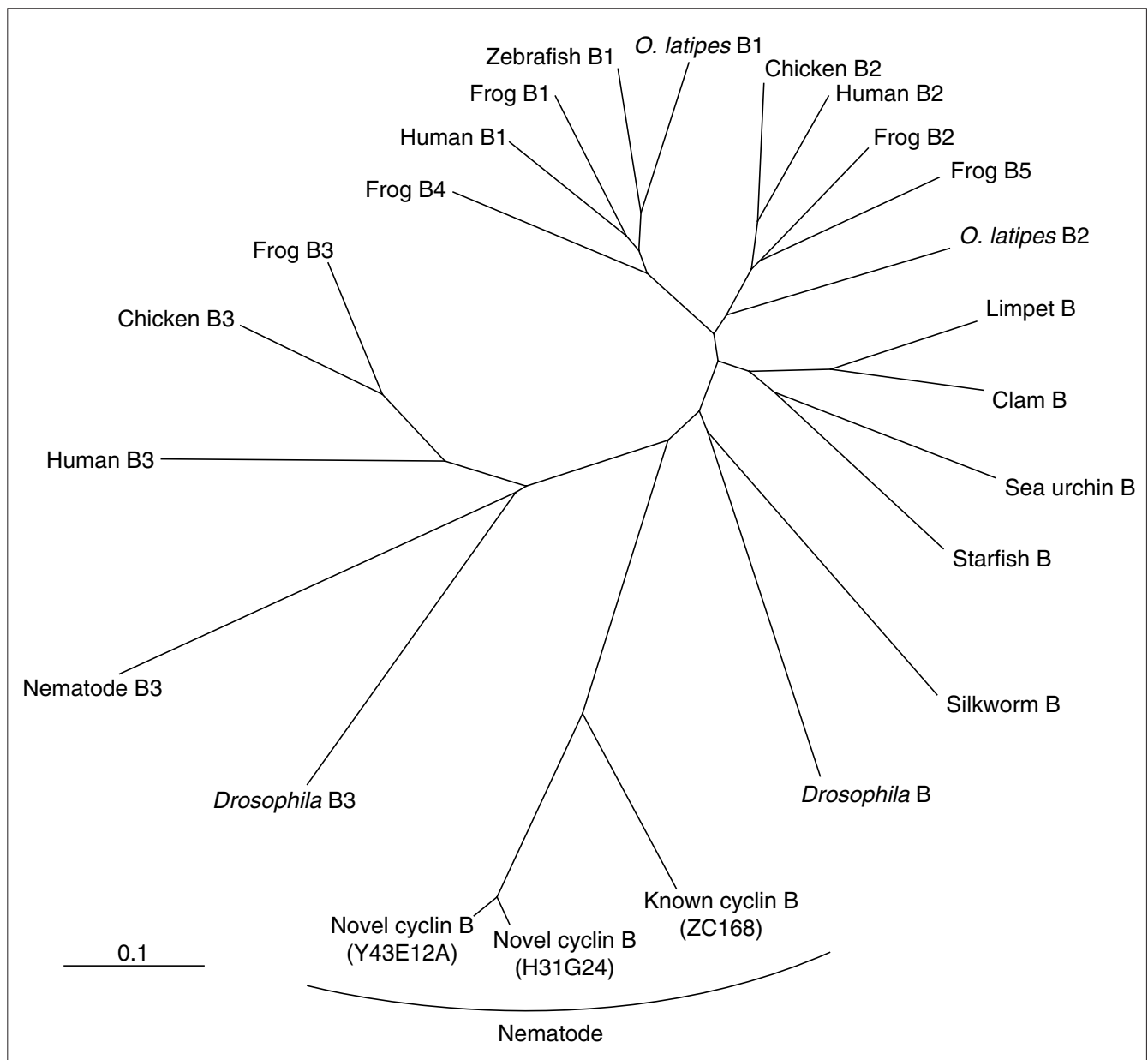


Figure 3
Phylogenetic tree for all known A-type cyclins. Vertebrate A-type cyclins can be divided into two groups: somatic (A2) and germline (A1), whereas completion of the *C. elegans* and *D. melanogaster* genome-sequencing projects shows that there is only one A-type cyclin in these organisms. The alignment used for the construction of this phylogenetic tree is given in the additional data files. *A. pectinifera*, *Asterina pectinifera*; *H. pulcherrimus*, *Hemicentrotus pulcherrimus*; *H. robusta*, *Helobdella robusta*. The scale bar represents 0.1 substitutions per site.

**Figure 4**

Phylogenetic tree for selected B-type cyclins. The B3 cyclins form an evolutionarily conserved family distinct from all the other B-type cyclins. The number of these other B-type cyclins varies between organisms. In vertebrates there are two evolutionarily divergent groups; one includes the B1 and *Xenopus* (frog) B4 cyclins and the other the B2 and *Xenopus* B5 cyclins. The alignment used for the construction of this phylogenetic tree is given in the additional data files. *O. latipes*, *Oryzias latipes*. The scale bar represents 0.1 substitutions per site.

after the divide between invertebrates and vertebrates. It is clear that in invertebrates a single A-type cyclin can carry out all the functions of the mitotic and meiotic cell cycles.

Our finding of 15 cyclin A pseudogenes was unexpected, as the *C. elegans* genome is generally considered to contain a limited number of pseudogenes. However, the observation does explain the multiple restriction enzyme fragments that hybridize with a cyclin A probe on a Southern blot [13].

B-type cyclins

In each organism we identified a single B3-type cyclin, distantly related to both the A- and B-type cyclin families [22]. The precise cellular role of cyclin B3 remains to be determined, although it appears to function late in mitosis [23]. In *Drosophila* there is functional redundancy between the B and B3 cyclins in mitosis; however, deletion of both genes results in embryonic lethality. In contrast, both cyclins B and B3 are required for female fertility and cyclin B, but not B3,

is required for male fertility. Therefore, it is possible that the mammalian B3 cyclins could also be partially functionally redundant with B-type cyclins.

While *Drosophila* has only one more B-type cyclin gene in addition to cyclin B3, *C. elegans* has a further three genes which encode near-identical B-type cyclins. It is worth noting that these multiple nematode B-type cyclins are more closely related to each other than they are to any other B-type cyclins (Figure 4). The B-type cyclin phylogenetic tree provides evidence for multiple gene-duplication events. Such an event early in animal evolution may have given rise to the B- and B3-type cyclins identified in all animals to date. More recent duplication events can explain the variation in number of B-type cyclins between organisms.

The finding of three closely related *C. elegans* B-type cyclins has some significance in the interpretation of previously published data. RNA interference (RNAi) analysis of the nematode cyclin B resulted in embryonic lethality [24,25]. However, gene product ablation using RNAi may not distinguish between such high levels of identity and consequently the function of all three B-type cyclins will be lost, so the best interpretation possible is that at least one of these B-type cyclins is required. Additionally, three *C. elegans* cyclin B mRNAs of different lengths were observed and it was reported that they may arise from the use of different polyadenylation sites [13]. An alternative interpretation, in the light of this study, is that each gene produces a single message so similar (approximately 80% nucleotide identity over the ORF) that the probes did not distinguish between them. Confirmation of this would require the generation of specific probes. One of the probes (a 3' UTR fragment) used by Kreutzer *et al.* [13] should have been specific for the previously identified cyclin B and did detect a single band. This allows us to predict that this gene gives rise to the largest (1.7 kbp) of the three transcripts. Interestingly, Kreutzer *et al.* [13] observed that the three cyclin B transcripts are differentially expressed in the maternal and paternal germlines. Therefore, it is possible that the three *C. elegans* cyclin B genes have distinct expression patterns through development.

The remaining vertebrate B-type cyclins fall into two evolutionarily divergent groups, one containing the B1 cyclins and the recently identified *Xenopus* cyclin B4 and the other the B2 cyclins and the *Xenopus* cyclin B5 [20]. The genes encoding cyclins B1, B2 and B3 were readily identifiable within the draft human genome sequence; there are no other B-type cyclin genes present. Furthermore, homologs of the mitotic cyclins A1, A2, B1, B2 and B3 were readily identified in the extensive zebrafish EST database and the draft pufferfish genome database, but no cyclin B4 or B5 homologs were found. Together, this suggests that cyclins B4 and B5 are unique to amphibians, perhaps, as suggested by Hohegger *et al.* [20], as a requirement of their large eggs.

In conclusion we suggest that there are three conserved families of mitotic cyclins in animals: the A-type cyclins with one gene in invertebrates and two genes in vertebrates; the B3-type cyclins with a single gene in all animals; and the B-type cyclins with different numbers of genes in different organisms, possibly as an adaptation to their distinct developmental strategies.

Materials and methods

BLAST searches

BLAST 2.0 was used to carry out similarity searches against the available sequence databases with protein or DNA query sequences [26]. For searches against the *C. elegans* and human genome sequences the Sanger Institute server [27] was used. For all other searches the NCBI server [28] was used.

Alignments

The GCG Wisconsin Package Version 10.0 PileUp program was used for creating multiple sequence alignments (using progressive pairwise alignments); alignments were further refined using the CLUSTAL W program version 1.8 [29] and manually. The TreeView program [30] was used for displaying and printing phylogenies produced by CLUSTAL W.

Additional data files

The alignments used for the construction of the phylogenetic trees of A- and B-type cyclins is available with the online version of this paper.

Acknowledgements

We thank Tim Hunt, Bill Colledge, Howard Baylis, Helena Webb and Catherine French for critical reading of the manuscript and/or helpful discussions. C.N. was the recipient of a BBSRC studentship.

References

1. Howe JA, Howell M, Hunt T, Newport JW: **Identification of a developmental timer regulating the stability of embryonic cyclin A and a new somatic A-type cyclin at gastrulation.** *Genes Dev* 1995, **9**:1164-1176.
2. Sweeney C, Murphy M, Kubelka M, Ravnik S, Hawkins CF, Wolgemuth D, Carrington M: **A distinct cyclin A is expressed in germ cells in the mouse.** *Development* 1996, **122**:53-64.
3. Yang R, Morosetti R, Koeffler HP: **Characterization of a second human cyclin A that is highly expressed in testis and in several leukemic cell lines.** *Cancer Res* 1997, **57**:913-920.
4. Ravnik SE, Wolgemuth DJ: **Regulation of meiosis during mammalian spermatogenesis: the A-type cyclins and their associated cyclin-dependent kinases are differentially expressed in the germ-cell lineage.** *Dev Biol* 1999, **207**:408-418.
5. Liu D, Matzuk MM, Sung WK, Guo Q, Wang P, Wolgemuth DJ: **Cyclin A1 is required for meiosis in the male mouse.** *Nat Genet* 1998, **20**:377-380.
6. Murphy M, Stinnakre M-G, Senamaud-Beaufort C, Sweeney C, Kubelka M, Carrington M, Bréchet C, Sobczak-Thépot J: **Delayed early embryonic lethality following disruption of the murine cyclin A2 gene.** *Nat Genet* 1997, **15**:83-86.
7. Pagano M, Pepperkok R, Verde F, Ansorge W, Draetta G: **Cyclin A is required at two points in the human cell cycle.** *EMBO J* 1992, **11**:961-971.
8. Pines J: **Four-dimensional control of the cell cycle.** *Nat Cell Biol* 1999, **1**:E73-E79.

9. Brandeis M, Rosewell I, Carrington M, Crompton T, Jacobs MA, Kirk J, Gannon J, Hunt T: **Cyclin B2-null mice develop normally and are fertile whereas cyclin B1-null mice die in utero.** *Proc Natl Acad Sci USA* 1998, **95**:4344-4349.
10. Lozano JC, Perret E, Schatt P, Arnould C, Peaucellier G, Picard A: **Molecular cloning, gene localization, and structure of human cyclin B3.** *Biochem Biophys Res Commun* 2002, **291**:406-413.
11. Jacobs HW, Knoblich JA, Lehner CF: **Drosophila Cyclin B3 is required for female fertility and is dispensable for mitosis like Cyclin B.** *Genes Dev* 1998, **12**:3741-3751.
12. Miller ME, Cross FR: **Cyclin specificity: how many wheels do you need on a unicycle?** *J Cell Sci* 2001, **114**:1811-1820.
13. Kreutzer MA, Richards JP, De Silva-Udawatta MN, Temenak JJ, Knoblich JA, Lehner CF, Bennett KL: **Caenorhabditis elegans cyclin A- and B-type genes: a cyclin A multigene family, an ancestral cyclin B3 and differential germline expression.** *J Cell Sci* 1995, **108**:2415-2424.
14. The *C. elegans* Sequencing Consortium: **Genome sequence of the nematode *C. elegans*: a platform for investigating biology.** *Science* 1998, **282**:2012-2018.
15. Ward S, Burke DJ, Sulston JE, Coulson AR, Albertson DG, Ammons D, Klass M, Hogan E: **Genomic organization of major sperm protein genes and pseudogenes in the nematode *Caenorhabditis elegans*.** *J Mol Biol* 1988, **199**:1-13.
16. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PV, Hoskins RA, Galle RF *et al.*: **The genome sequence of *Drosophila melanogaster*.** *Science* 2000, **287**:2185-2195.
17. Lehner CF, O'Farrell PH: **Expression and function of *Drosophila* cyclin A during embryonic cell cycle progression.** *Cell* 1989, **56**:957-968.
18. Whitfield WG, Gonzalez C, Maldonado-Codina G, Glover DM: **The A- and B-type cyclins of *Drosophila* are accumulated and destroyed in temporally distinct events that define separable phases of the G2-M transition.** *EMBO J* 1990, **9**:2563-2572.
19. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W *et al.*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.
20. Hochegeger H, Klotzbucher A, Kirk J, Howell M, le Guellec K, Fletcher K, Duncan T, Sohail M, Hunt T: **New B-type cyclin synthesis is required between meiosis I and II during *Xenopus* oocyte maturation.** *Development* 2001, **128**:3795-3807.
21. Goldberg E: **Reproductive implications of LDH-C4 and other testis-specific isozymes.** *Exp Clin Immunogenet* 1985, **2**:120-124.
22. Gallant P, Nigg EA: **Identification of a novel vertebrate cyclin: cyclin B3 shares properties with both A- and B-type cyclins.** *EMBO J* 1994, **13**:595-605.
23. Parry DH, O'Farrell PH: **The schedule of destruction of three mitotic cyclins can dictate the timing of events during exit from mitosis.** *Curr Biol* 2001, **11**:671-683.
24. Piano F, Schetter AJ, Mangone M, Stein L, Kempfues KJ: **RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*.** *Curr Biol* 2000, **10**:1619-1622.
25. **Sequence Report for: ZC168.4** [<http://www.wormbase.org/db/seq/sequence?name=ZC168.4;class=Sequence>]
26. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.
27. **The Wellcome Trust Sanger Institute** [<http://www.sanger.ac.uk/>]
28. **NCBI BLAST Home Page** [<http://www.ncbi.nlm.nih.gov/BLAST/>]
29. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.
30. Page RD: **TreeView: an application to display phylogenetic trees on personal computers.** *Comput Appl Biosci* 1996, **12**:357-358.