



HHS Public Access

Author manuscript

Chromosoma. Author manuscript; available in PMC 2019 May 13.

Published in final edited form as:

Chromosoma. 2016 March ; 125(1): 65–73. doi:10.1007/s00412-015-0528-7.

Telomerase lost?

James M. Mason¹, Thomas A. Randall², and Radmila Capkova Frydrychova^{3,*}

¹Laboratory of Genome Integrity and Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

²Integrative Bioinformatics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

³Biology Centre AS CR, Institute of Entomology, Czech Republic

Abstract

Telomerase and telomerase-generated telomeric DNA sequences are widespread throughout eukaryotes, yet they are not universal. Neither telomerase nor the simple DNA repeats associated with telomerase have been found in some plant and animal species. Telomerase was likely lost from Diptera before the divergence of Diptera and Siphonaptera, some 260 million years ago. Even so, Diptera is one of the most successful animal orders, making up 11% of known animal species. In addition, many species of Coleoptera and Hemiptera seem to lack canonical telomeric repeats at their chromosome ends. These and other insects that appear to lack canonical terminal repeat sequences account for another 10–15% of animal species. Conversely, the silk moth *Bombyx mori* maintains canonical telomeric sequences at its chromosome ends but seems to lack a functional telomerase. We speculate that a telomere-specific capping complex that recognizes the telomeric repeats and protects chromosome ends is the determining factor in maintaining canonical telomeric sequences, and that telomerase is an early and efficacious mechanism for satisfying the needs of capping complex. There are alternate mechanisms for maintaining chromosome ends that do not depend on telomerase, such as recombination found in some human cancer cells and yeast mutants. These mechanisms may maintain the canonical telomeric repeats or allow the terminal sequence to evolve when specificity of the capping complex for terminal repeat sequences is weak.

Keywords

Telomere; Telomerase; Shelterin; Evolution; Insecta

Introduction

Telomeres are protective, nucleoprotein structures at the ends of eukaryotic chromosomes. Regardless of some differences in their DNA sequences or protein compositions, telomeres

*Corresponding author: Radmila.Frydrychova@hotmail.com, phone: + 420 387 775 054, fax: +420 385 310 354.

Conflict of interest. Authors declare that they have no conflict of interest.

Ethical approval. This article does not contain any studies with human participants or animals performed by any of the authors.

across phyla cope with the same issues arising from the nature of linear chromosomes. First, they distinguish natural chromosome ends from chromosomal breaks through a telomere specific, multiprotein structure called a telomere cap (Muller 1938). Second, they compensate for chromosome shortening due to incomplete DNA replication at chromosome ends (Greider and Blackburn 1996). As formation of the telomere cap in many species is telomere sequence-dependent and requires satisfactory sequence length, the gradual loss of telomeric DNA ultimately leads to replicative senescence or apoptosis (Sfeir and de Lange 2012). Telomeres thus work as cellular timekeepers. However, in some cells, such as germ or stem cells, telomere maintenance mechanisms are activated (Shay and Wright 2010), which by extending telomere length compensate the telomere loss and enable further cell proliferation.

Several mechanisms of telomere maintenance have been identified. The most common mechanism is telomerase, a specialized reverse transcriptase that according to its RNA template repeatedly synthesizes a short telomeric sequence onto chromosome ends. Telomeres maintained by telomerase thus contain tandem arrays of a short, 5–8 bp repeat (Greider 1996). Another telomere maintenance mechanism is homologous recombination, which has been found to extend long satellite sequences at telomeres of several organisms, such as representatives of lower Diptera (Mason et al. 2011) and is also found in human cells as an alternative to telomerase (Cesare and Reddel 2010). A third method to extend telomeres has been found in *Drosophila* species, in which telomeres consist of arrays of retrotransposons that maintain telomere length by transposition specifically to chromosome ends (Casacuberta and Pardue 2002; Pardue and DeBaryshe 2003; Biessmann and Mason 2003; Mason et al. 2008; Capkova Frydrychova et al. 2009).

Although canonical telomeric repeat sequences generated by telomerase follow certain rules, such as being relatively G-rich in the 3' strand that extends to the chromosome end, these sequences are not uniform throughout eukaryotes. Within Chromalveolata, for example, the terminal repeat sequence varies considerably (Mason et al. 2011); there may also be variation within smaller taxa as well, such as green algae (Fulneková et al. 2012). In contrast, telomeric sequence is highly conserved in Unikonta, in which (TTAGGG)_n, sometimes referred to as the vertebrate telomeric sequence, is the predominant terminal repeat sequence for fungi, animals and amoebozoa (Fulneckova et al. 2013). Yet even among the unikonts the DNA sequence at chromosome ends is not completely uniform. While most filamentous fungi use (TTAGGG)_n at their chromosome ends, yeasts in the classes Saccharomycetes and Schizosaccharomycetes use a short repeat that varies from one copy to the next along a single telomere (McEachern and Blackburn 1994). Among animals, the (TTAGGG)_n repeat in deuterostomes is stable (Gomes et al. 2011), while the canonical repeat varies somewhat from one phylum to the next in protostomes (Mason et al. 2011). Thus, for example, the (TTAGGC)_n terminal repeat has been found in Nematoda, (TGTGGG)_n in Rotifera, and (TTAGGG)_n in Onychophora, Mollusca and Annelida. Extensive studies of telomere sequence composition in arthropods revealed that the predominant telomeric repeat in this phylum is (TTAGG)_n, but there is some variation, especially in insects. (TTAGG)_n has been lost at least 15 times during insect evolution (Frydrychova et al. 2004), including eight times in Coleoptera (Frydrychova and Marec 2002).

The (TTAGG)_n telomeric sequence has been detected using Southern hybridization and fluorescence *in situ* hybridization, which may not provide the resolution necessary to show that the (TTAGG)_n sequence in the tested species is the true, canonical telomeric sequence. It could be speculated that the (TTAGG)_n sequence found at chromosome ends might be, in fact, only a non-functional relict of the original (TTAGG)_n or a part of more complex sequence, such as satellite sequence, or even of subtelomeric sequence. Such suspicion, however, has been refuted recently by Korandova et al. (2014), who demonstrated TTAGG-specific telomerase activity in phylogenetically distant insect species. Nevertheless, one problem that still persists is that in many cases a single species has been tested for the canonical repeat and telomerase, and this species is assumed to be representative of the order or the family in which it is found. More tests are required to verify this assumption.

Telomeres without terminal repeat sequences

Loss of canonical telomeric sequence may not necessarily lead to the death of the organism or even the cell in which it occurred. The typical telomeric sequence in the plant kingdom is (TTTAGGG)_n, the so-called *Arabidopsis-type* motif (Riha and Shippen 2003). Two groups of flowering plants are known in which a replacement of the plant telomere sequence has occurred. The first is the family Solanaceae, in which the telomeric motif (TTTTTTAGGG)_n maintained by telomerase, was discovered (Peška et al. 2015). The other group is the order Asparagales, where two switch-points in the evolution of telomeres were documented. During the first switch-point, the *Arabidopsis-type* telomeric motif was replaced with the vertebrate-type sequence; at the second switch-point, the canonical telomeric sequence in an ancestor to *Allium* was lost and substituted by a so far undiscovered sequence proposed to be elongated independently of telomerase (Fajkus et al. 2005). It has been suggested that telomeres in *Allium* and some related liliaceous species might have been replaced by a satellite sequence, ribosomal DNA, Ty1-copia retroelements or an En/Spm-transposable element-like sequence (Pearce et al. 1996; Pich et al. 1996; Pich and Schubert 1998).

The proposed sequences in *Allium* are reminiscent of telomeric repeats found in some flies. Nematocera species, the so-called lower Diptera, have repeated sequences at their chromosome ends that may elongate telomeres by gene conversion (Nielsen and Edstrom 1993), while *Drosophila* species carry non-long terminal repeat (LTR) retrotransposons that elongate telomeres by transposition specifically to chromosome termini. More than 40 dipteran genomes have been sequenced, and none of them shows evidence of a telomerase reverse transcriptase gene (TAR and JMM unpublished data). However, a reverse transcriptase-related protein of unknown function has been detected by immunohistochemistry in telomere regions of salivary gland polytene chromosomes in *Chironomus* and *Rhynchosciara* species (López et al. 1999; Gorab 2003). A chronogram for Diptera (Wiegmann et al. 2011) shows that *Rhynchosciara* (Bibionomorpha, Sciaridae) diverged from *Drosophila* (Schizophora, Drosophilidae) some 230 million years ago (Mya). As *Rhynchosciara americana* has a complex repeat with a unit length of 14, 16, and 22 bp (Rossato et al. 2007; Madalena et al. 2010; Fernandes et al. 2012), and all tested *Drosophila* species carry retrotransposons at their chromosome ends, it appears that these telomere-specific retrotransposons arose between 65 and 230 Mya. Thus, the lineage leading to *Drosophila* may have existed for some 40–200 million years with neither telomerase nor

telomeric retrotransposons to maintain telomere-specific DNA. Further, Siphonaptera and Mecoptera, sister orders to Diptera (Fig. 1), also seem to lack (TTAGG)_n repeats at their chromosome ends, which suggests that telomerase and the terminal sequence generated by telomerase were lost around 270 Mya, before the divergence of Diptera and Siphonaptera and after the separation of the lineages leading to Diptera and Lepidoptera (Fig. 1). Despite the loss of short canonical telomeric repeats and telomerase, Diptera, which includes 152,000 species, is a highly successful order, even for insects, suggesting that loss of telomerase together with canonical telomeric repeats is not strongly deleterious.

While the evidence for replacement of canonical telomeric repeats with alternative telomeric sequences is not as strong in Coleoptera as it is in Diptera, there are some indications that it might have happened here as well. The arthropod (TTAGG)_n telomeric repeats have not been found in 8 of 19 tested coleopteran families (Fig. 2). In three closely related families within the superfamily Tenebrionoidea the (TTAGG)_n sequence has been replaced by another, similar sequence, (TCAGG)_n (Mravinac et al. 2011). In three families, Chrysomelidae, Cucujidae and Curculionidae, the results were mixed, with some species showing evidence for canonical telomeres and others not (Okazaki et al. 1993; Sahara et al. 1999; Frydrychova and Marec 2002). Several other variations of the telomeric sequence were tested in these families, but a substitute for the (TTAGG)_n or (TCAGG)_n sequence could not be found (Fig. 2; RCF unpublished data). The phylogenetic tree of Curculionidae provided by Hunt et al. (2007) suggests a single event of telomere sequence loss involving 33,000 species, while a more recent tree (McKenna et al. 2009) suggests two loss events, together involving 26,000 species. The number of species in each family does not seem to be related to the presence or absence of identifiable (TTAGG)_n or a related substitute.

The situation in Hemiptera (Fig. 3a) is simpler than in Coleoptera. The phylogenetic tree suggests that a single event occurred that resulted in the loss of (TTAGG)_n from chromosome ends sometime before the divergence of Cimicomorpha and Pentatomomorpha and after the separation of these taxa from Nepomorpha. Nine genera have been tested in the former two infraorders with consistent results, suggesting that 37,000 heteropteran species lack the canonical arthropod telomeric repeat. It also appears that heteropteran species without the (TTAGG)_n terminal repeat are more successful, or at least more diverse as defined by the species number, than those that have this sequence on their chromosome ends, even though Cimicomorpha and Pentatomomorpha have arisen relatively recently.

A survey of sequenced insect genomes identified a candidate telomerase (TERT) gene, where TERT is defined as a protein containing a ribonucleotide binding domain that is N-terminal to a reverse transcriptase domain, suggests that there may have been two separate events, one in Cimicomorpha and one in Pentatomomorpha. Telomerases in *Cimex lectularis* and *Rhodnius prolixus* (Cimicomorpha) have a truncated version of the T domain, which is a solvent exposed portion of the telomerase thought to be involved in binding the single stranded RNA template (Mitchell et al. 2010), while telomerase in *Oncopeltus fasciatus* (Pentatomomorpha) has lost much of motifs B and C (Fig. 3b). Telomeric DNA sequence was not assayed in the latter two species, but the TERT sequence results are consistent with the possible loss of canonical repeats.

As stated above, there is good evidence that 156,000 species of Diptera, Siphonaptera and Mecoptera lack both telomerase and the associated telomeric repeats; there is some evidence - the inability to find any of several possible telomeric repeats that resemble the canonical sequence - that 33,000 coleopterans lack canonical telomeric sequence; and there is a suggestion, based on FISH and/or Southern blot assays using (TTAGG)_n as a probe, that as many as 75,000 coleopterans, 37,000 hemipterans and 17,000 other insects may also lack the arthropod telomeric repeat. Based on a comparison of the number of known insect species (Fig 1) with the number of other animal species (www.globalchange.umich.edu), insects comprise 77% of known animal species, while Diptera comprises 11%. Considering the insects enumerated here, possibly as many as a quarter of animal species use a DNA sequence on their chromosome ends that is unrelated to sequences normally maintained by telomerase, although these unusual terminal sequences may be added by an as yet undiscovered divergent telomerase protein in some cases.

Telomeres of *Bombyx mori* and *Giardia lamblia*

The telomeres in *Bombyx mori* provide an interesting counterpoint to the apparent loss of canonical telomeric repeats in other insects. The TERT gene has an unusual promoter and is expressed very little in this species; most of the expressed message is too short to produce an active protein, and no TERT activity could be found in any tissue (Osanai et al. 2006). In spite of this, (TTAGG)_n is found on chromosome ends (Sahara et al. 1999; Sasaki and Fujiwara 2000; Fujiwara et al. 2005; Osanai et al. 2006). To explain the (TTAGG)_n presence at *B. mori* telomeres, based on the questionable telomerase activity, we can hypothesize that either the level of telomerase activity is still able to replenish the telomere length when necessary, or the (TTAGG)_n telomeric repeat is maintained by a mechanism other than telomerase. More than 1000 copies of non-LTR retrotransposons, TRAS and SART, are also found near chromosome ends in this species, although these transposable elements target the (TTAGG)_n sequence and insert into the middle of the array, not at the actual chromosome end (Okazaki et al. 1995; Kubo et al. 2001; Fujiwara et al. 2005). The canonical telomeric sequence in *B. mori* extends to the actual chromosome end, with an average 6–8 kb of (TTAGG)_n distal to the most distal retrotransposon. Although, these retrotransposons should be considered subtelomeric (Edward and Becker 2014), rather than telomeric in nature, the insertions of the retroelements into telomere regions might resemble a putative evolutionary intermediate between canonical and retrotransposon based telomeres.

Some similarities to telomeres in *B. mori* can be seen in one of the earliest branching eukaryotes, the protozoan *Giardia lamblia*, where terminal arrays of (TAGGG)_n maintained by telomerase are adjacent to arrays of two families of non-LTR retroelements (Arkhipova and Morrison 2001). In agreement with the suggestion that asexual organisms cannot maintain deleterious transposable elements, in the genome of *G. lamblia*, which is not known to reproduce sexually, only retrotransposons that are either subtelomeric, or completely nonfunctional have been identified. This suggests that subtelomeric retroelements in *G. lamblia* may benefit their host either by buffering the zone between the chromosome terminus and protein coding genes, or by another unknown telomeric function (Pardue et al. 2001).

The role of the telomere capping complex

Telomeres are critical for cell viability; therefore if species lose canonical telomeric DNA sequences the question arises, what keeps them viable? We speculate that a crucial player is a capping complex that binds to chromosome ends and distinguishes them from double stranded chromosome breaks (de Lange 2009; Fulcher et al. 2014). As long as the capping complex is stringent in its DNA sequence binding preference, the telomeric sequence would not be allowed to vary, regardless of the maintenance mechanism. This might explain why *B. mori* still retains the canonical telomeric sequence, even in the apparent absence of strong telomerase activity. On the other hand, if the capping complex loses strict sequence specificity, the terminal DNA sequence may be allowed to vary - possibly with some constraints - even when telomerase is active. This may be an explanation for the observed plasticity of canonical telomeric DNA sequence among Chromalveolata (Mason et al. 2011) and green algae (Fulneckova et al. 2012) or the unusual terminal sequence in yeast (McEachern and Blackburn 1994). It is also possible that the capping complex may lose all sequence specificity, as seen for *Drosophila* (Raffa et al. 2011; Capkova Frydrychova and Mason 2013). Loss of sequence specificity, however, does not necessarily mean that binding of broken chromosome ends by the capping complex becomes frequent. Even in *Drosophila* capping of broken chromosome ends is rare (Muller 1938; Muller and Herskowitz 1954) but may occur in the presence of a specific mutant gene (Mason et al. 1984), or when chromosome breaks occur very close to previous telomeres (Levis 1989; Tower et al. 1993).

According to our hypothesis telomerase provides an effective mechanism for maintaining a capping complex-preferred DNA sequence but becomes superfluous when capping complex sequence specificity breaks down, as we speculate happened in the lineage leading to Diptera. The model proposes that disruption of sequence specificity precedes the loss or modification of the canonical telomeric sequence, because changes in telomeric DNA sequence in the presence of strong sequence specificity of the capping complex would be lethal events. Thus, the relative stability of canonical telomeric repeats in deuterostomes may be due to stronger sequence specificity of the capping complex in this lineage than in protostomes.

Although telomeres play the same vital roles in chromosome length regulation and end protection, the canonical telomeric DNA sequence is stable in most unikonts, including amoebozoa, fungi and the Deuterostome branch of Metazoa, and the TERT gene is well conserved across eukaryotes, telomere capping proteins show less conservation than might be expected (Linger and Price 2009). In mammals the telomere-specific protein complex shelterin contains six proteins, three of which bind specifically to telomeric repeats, where they serve as a platform for dynamic binding of numerous telomere-associated factors participating in biological processes, such as DNA damage repair, chromosome cohesion, chromatin remodeling, cell cycle and transcription regulation (Giannone et al. 2010). Based on the involvement of shelterin proteins in numerous vital functions, it is surprising that the shelterin complex is very fast evolving in mammals and has not been found in some organisms, such as *Saccharomyces cerevisiae* (Nandakumar and Cech 2013) or any tested arthropod genomes (Mandrioli et al. 2014). While telomeric DNA sequences among protostomes are more variable, an examination of co-evolution of these sequences with their

protective protein cap is not possible, because these proteins have only been identified in *Drosophila* species (Fulcher et al. 2014).

Summary

Although our simple calculation indicating that canonical telomeric DNA sequences and possibly the telomerase system may have been lost in a substantial number of animal species, telomerase is still the most wide-spread mechanism to elongate telomeres. Given its wide distribution, telomerase obviously arose very early in the evolution of eukaryotes; it may also provide greater stringency or efficiency in the maintenance of the canonical sequence required by a sequence-specific capping complex than alternatives, such as recombination. These alternative mechanisms may occur readily, as seen in yeast (Teng and Zakian 1999) and human tumors (Cesare and Reddel 2010) even in the presence of telomerase (Dlaska et al. 2013). In addition, short telomeres or loss of telomerase from human cancer cells may stimulate alternative lengthening of telomeres (ALT) mechanisms (Morrish and Greider 2009; Queisser et al. 2013). Thus, it is perhaps not surprising that over evolutionary time telomerase may occasionally have been lost. One example, Diptera, is well documented, comprising more than 10% of animal species, and several others are suspected. With this in mind, the conservation of canonical telomeric sequence may provide a clue as to the stringency of sequence specificity of the capping complex.

Acknowledgements

We thank Walther Traut for critical reading of the manuscript and Ales Bezdek for providing information about Coleoptera phylogeny.

Funding. This work was supported by Grant No. 14–07172S from the Grant Agency of the Czech Republic (to RCF). We acknowledge the use of research infrastructure that has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 316304 (to RCF). Additional support was provided from Grant no. 052/2013/P and no. 038/2014/P of the Grant Agency of the University of South Bohemia (to RCF) and by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (to JMM and TAR).

References

- Arkhipova IR, Morrison HG (2001) Three retrotransposon families in the genome of *Giardia lamblia*: two telomeric, one dead. *Proc Natl Acad Sci U S A* 98:14497–14502. [PubMed: 11734649]
- Barnard PC (2011) *The Royal Entomological Society Book of British Insects*. Wiley-Blackwell, Oxford.
- Biessmann H, Mason JM (2003) Telomerase-independent mechanisms of telomere elongation. *Cell Mol Life Sci* 60:2325–33. [PubMed: 14625679]
- Capkova Frydrychova R, Mason JM (2013) Telomeres: Their Structure and Maintenance In: Stuart (ed) *The mechanisms of DNA replication*. In Tech, Rijeka, pp 423–444.
- Capkova Frydrychova R, Mason JM, Biessmann H (2009) Regulation of telomere length in *Drosophila*. *Cytogenet Genome Res* 122:356–364.
- Casacuberta E, Pardue M-L (2002) Coevolution of the telomeric retrotransposons across *Drosophila* species. *Genetics* 161:1113–24. [PubMed: 12136015]
- Cesare AJ, Reddel RR (2010b) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* 11:319–330. [PubMed: 20351727]
- Dlaska M, Schoffski P, Bechter OE (2013) Inter-telomeric recombination is present in telomerase-positive human cells. *Cell cycle* 12:2084–99. [PubMed: 23759591]

- Edward LJ, Becker MM (2014) Subtelomeres. Springer-Verlag Berlin Heidelberg.
- Fajkus J, Šykorová E, Leitch AR (2005) Telomeres in Evolution and Development. *Chromosome Res* 13:469–479. [PubMed: 16132812]
- Fernandes T, Madalena CRG, Gorab E (2012) Cloning and characterisation of a novel chromosome end repeat enriched with homopolymeric (dA)/(dT) DNA in *Rhynchosciara americana* (Diptera: Sciaridae). *Chromosome Res* 20:435–45. [PubMed: 22729403]
- Frydrychova R, Grossmann P, Trubac P, et al. (2004) Phylogenetic distribution of TTAGG telomeric repeats in insects. *Genome* 47:163–178. [PubMed: 15060613]
- Frydrychova R, Marec F (2002) Repeated losses of TTAGG telomere repeats in evolution of beetles (Coleoptera). *Genetica* 115:179–87. [PubMed: 12403172]
- Fujiwara H, Osanai M, Matsumoto T, Kojima KK (2005) Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori*. *Chromosome Res* 455–467. [PubMed: 16132811]
- Fulcher N, Derboven E, Valuchova S, Riha K (2014) If the cap fits, wear it: An overview of telomeric structures over evolution. *Cell Mol Life Sci* 71:847–865. doi: 10.1007/s00018-013-1469-z [PubMed: 24042202]
- Fulneková J, Hasíková T, Fajkus J, et al. (2012) Dynamic evolution of telomeric sequences in the green algal order Chlamydomonadales. *Genome Biol Evol* 4:248–64. [PubMed: 22247428]
- Fulneková J, Ševčíková T, Fajkus J, et al. (2013) A broad phylogenetic survey unveils the diversity and evolution of telomeres in eukaryotes. *Genome Biol Evol* 5:468–483. [PubMed: 23395982]
- Giannone RJ, McDonald HW, Hurst GB, et al. (2010) The protein network surrounding the human telomere repeat binding factors TRF1, TRF2, and POT1. *PLoS One* 5:e12407. [PubMed: 20811636]
- Gokhman VE, Anokhin BA, Kuznetsova VG (2014) Distribution of 18S rDNA sites and absence of the canonical TTAGG insect telomeric repeat in parasitoid Hymenoptera. *Genetica* 142:317–22. [PubMed: 24992984]
- Gomes NMV, Shay JW, Wright WE (2011) Telomere Biology in Metazoa. *Fed Eur Biochem Soc* 584:3741–3751.
- Gorab E (2003) Reverse transcriptase-related proteins in telomeres and in certain chromosomal loci of *Rhynchosciara* (Diptera: Sciaridae). *Chromosoma* 111:445–454. [PubMed: 12707782]
- Greider CW (1996) Telomere length regulation. *Annu Rev Biochem* 65:337–65. [PubMed: 8811183]
- Greider CW, Blackburn EH (1996) Telomeres, telomerase and cancer. *Sci Am* 274:92–7. [PubMed: 8560215]
- Grozeva S, Kuznetsova VG, Anokhin BA (2011) Karyotypes, male meiosis and comparative FISH mapping of 18S ribosomal DNA and telomeric (TTAGG)_n repeat in eight species of true bugs (Hemiptera, Heteroptera). *Comp Cytogenet* 5:355–74. [PubMed: 24260641]
- Grzywacz B, Chobanov DP, Maryska-Nadachowska A, Karamysheva T, Heller K-G, Warchałowska-liwa E (2014) A comparative study of genome organization and inferences for the systematics of two large bushcricket genera of the tribe Barbitistini (Orthoptera: Tettigoniidae: Phaneropterinae). *BMC Evol Biol* 14:48. [PubMed: 24625118]
- Hunt T, Bergsten J, Levkanicova Z, et al. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* (80-) 318:1913–16.
- Jetybayev IE, Bugrov AG, Karamysheva TV, Camacho JPM, Rubtsov NB (2012) Chromosomal Localization of Ribosomal and Telomeric DNA Provides New Insights on the Evolution of Gomphocerinae Grasshoppers. *Cytogene* 138:36–45.
- Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. [PubMed: 12136088]
- Kubo Y, Okazaki S, Anzai T, Fujiwara H (2001) Structural and phylogenetic analysis of TRAS, telomeric repeat-specific non-LTR retrotransposon families in Lepidopteran insects. *Mol Biol Evol* 18:848–57. [PubMed: 11319268]
- Kuznetsova VG, Grozeva SM, Anokhin BA (2012) The first finding of (TTAGG)_n telomeric repeat in chromosomes of true bugs (Heteroptera, Belostomatidae). *Comp Cytogenet* 6:341–6. [PubMed: 24260674]

- De Lange T (2009) How telomeres solve the end-protection problem. *Science* 326:948–952. [PubMed: 19965504]
- Levis RW (1989) Viable deletions of a telomere from a *Drosophila* chromosome. *Cell* 58:791–801. [PubMed: 2548737]
- Linger BR, Price CM (2009) NIH Public Access. 44:434–446.
- López CC, Rodrigues E, Díez JL, et al. (1999) Histochemical localization of reverse transcriptase in polytene chromosomes of chironomids. *Chromosoma* 108:302–7. [PubMed: 10525966]
- Lorite P, Carrillo JA, Palomeque T (2002) Conservation of (TTAGG)_n Telomeric Sequences Among Ants (Hymenoptera, Formicidae). *J Hered* 93:381–384.
- Madalena CRG, Amabis JM, Gorab E (2010) Unusually short tandem repeats appear to reach chromosome ends of *Rhynchosciara americana* (Diptera: Sciaridae). *Chromosoma* 119:613–623. [PubMed: 20614221]
- Mandrioli M, Zanasi F, Manicardi GC (2014) Karyotype rearrangements and telomere analysis in *Myzus persicae* (Hemiptera, Aphididae) strains collected on *Lavandula* sp. plants. *Comp Cytogenet* 8:259–274. [PubMed: 25610541]
- Maryanska-Nadachowska A, Kuznetsova VG, Karamysheva TV (2013) Chromosomal location of rDNA clusters and TTAGG telomeric repeats in eight species of the spittlebug genus *Philaenus* (Hemiptera: Auchenorrhyncha: Aphrophoridae). *Eur J Entomol* 110:411–418.
- Mason JM, Frydrychova RC, Biessmann H (2008) *Drosophila* telomeres: an exception providing new insights. *Bioessays* 30:25–37. [PubMed: 18081009]
- Mason JM, Reddy HM, Capkova Frydrychova R (2011) Telomere maintenance in organisms without telomerase. *DNA Replication-Current Adv*.
- Mason JM, Strobel E, Green MM (1984) mu-2: mutator gene in *Drosophila* that potentiates the induction of terminal deficiencies. *Proc Natl Acad Sci U S A* 81:6090–4. [PubMed: 6435123]
- McEachern MJ, Blackburn EH (1994) A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. *Proc Natl Acad Sci U S A* 91:3453–3457. [PubMed: 8159768]
- McKenna DD, Sequeira AS, Marvaldi AE, B.D. F (2009) Temporal lags and overlap in the diversification of weevil and flowering plants. *PNAS* 106:7083–7088. [PubMed: 19365072]
- Menezes RST, Silva TM, Carvalho AF, Andrade-Souza V, Silva JG, Costa MA (2013) Numerical and structural chromosome variation in the swarm-founding wasp *Metapolybia decorata* Gribodo 1896 (Hymenoptera, Vespidae). *Genetica* 141:273–80. [PubMed: 23813236]
- Misof B, Liu S, Meusemann K et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science* (80-) 346:763–768.
- Mitchell M, Gillis A, Futahashi M, et al. (2010) Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat Struct Mol Biol* 17:513–518. [PubMed: 20357774]
- Mohan KN, Rani BS, Kulashreshta PS, Kadandale JS (2011) Characterization of TTAGG telomeric repeats, their interstitial occurrence and constitutively active telomerase in the mealybug *Planococcus lilacinus* (Homoptera; Coccoidea). *Chromosoma* 120:165–75. [PubMed: 21088846]
- Monti V, Giusti M, Bizzaro D, et al. (2011a) Presence of a functional (TTAGG)_n telomere-telomerase system in aphids. *Chromosome Res* 19:625–33. [PubMed: 21667174]
- Monti V, Manicardi GC, Mandrioli M (2011b) Cytogenetic and molecular analysis of the holocentric chromosomes of the potato aphid *Macrosiphum euphorbiae* (Thomas, 1878). *Comp Cytogenet* 5:163–72. [PubMed: 24260627]
- Mravinac B, Mestrovic N, Cavrak VV, Plohl M (2011) TCAGG, an alternative telomeric sequence in insects. *Chromosoma* 120:367–76. [PubMed: 21499744]
- Muller H (1938) The remaking of chromosomes. *Collect Net* 13:181–198.
- Muller H, Herskowitz I (1954) Concerning the healing of chromosome ends produced by breakage in *Drosophila melanogaster*. *Am Naturlist* 88:177–208.
- Nandakumar J, Cech TR (2013) Finding the end: recruitment of telomerase to the telomere. *Nat Rev Mol Cell Biol* 14:69–82. [PubMed: 23299958]

- Nielsen L, Edstrom JE (1993) Complex telomere-associated repeat units in members of the genus *Chironomus* evolve from sequences similar to simple telomeric repeats. *Mol Cell Biol* 13:1583–1589. [PubMed: 8441399]
- Novotná J, Havelka J, Starý P, Koutecký P, Vitková M (2011) Karyotype analysis of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Hemiptera: Aphididae) reveals a large X chromosome with rRNA and histone gene families. *Genetica* 139:281–9. [PubMed: 21336570]
- Okazaki S, Ishikawa H, Fujiwara H (1995) Structural analysis of TRAS1, a novel family of telomeric repeat-associated retrotransposons in the silkworm, *Bombyx mori*. *Mol Cell Biol* 15:4545–4552. [PubMed: 7623845]
- Okazaki S, Tsuchida K, Maekawa H, et al. (1993) Identification of a Pentanucleotide Telomeric Sequence, (TTAGG)_n, in the Silkworm *Bombyx mori* and in Other Insects. *Mol Cell Biol* 13:1424–32. [PubMed: 8441388]
- Osanai M, Kojima KK, Futahashi R, et al. (2006) Identification and characterization of the telomerase reverse transcriptase of *Bombyx mori* (silkworm) and *Tribolium castaneum* (flour beetle). *Gene* 376:281–9. [PubMed: 16793225]
- Pardue M-L, DeBaryshe PG (2003) Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu Rev Genet* 37:485–511. [PubMed: 14616071]
- Pardue ML, DeBaryshe PG, Lowenhaupt K (2001) Another protozoan contributes to understanding telomeres and transposable elements. *Proc Natl Acad Sci U S A* 98:14195–7. [PubMed: 11734638]
- Pearce S, Pich U, Harrison G, et al. (1996) The Ty1-copia group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chromosome Res* 4:357–64. [PubMed: 8871824]
- Peška V, Fajkus P, Fojtová M, et al. (2015) Characterisation of an unusual telomere motif (TTTTTTAGGG)_n in the plant *Cestrum elegans* (Solanaceae), a species with a large genome. *Plant J* 82:644–654. doi: 10.1111/tpj.12839 [PubMed: 25828846]
- Pich U, Fuchs J, Schubert I (1996) How do Alliaceae stabilize their chromosome ends in the absence of TTTAGGG sequences? *Chromosome Res* 4:207–13. [PubMed: 8793205]
- Pich U, Schubert I (1998) Terminal heterochromatin and alternative telomeric sequences in *Allium cepa*. *Chromosome Res* 6:315–21. [PubMed: 9688522]
- Queisser A, Heeg S, Thaler M, et al. (2013) Inhibition of telomerase induces alternative lengthening of telomeres during human esophageal carcinogenesis. *Cancer Genet* 206:374–86. [PubMed: 24331919]
- Raffa GD, Ciapponi L, Cenci G, et al. (2011) Terminin. A protein complex that mediates epigenetic maintenance of *Drosophila* telomeres. *Nucleus* 2:5:383–391. [PubMed: 21989238]
- Riha K, Shippen DE (2003) Telomere structure, function and maintenance in *Arabidopsis*. *Chromosome Res* 11:263–275. doi: 10.1023/A:1022892010878
- Rossato RM, Madalena CRG, Gorab E (2007) Unusually short tandem repeats in the chromosome end structure of *Rhynchosciara* (Diptera: Sciaridae). *Genetica* 131:109–116. [PubMed: 17380273]
- Sahara K, Marec F, Traut W (1999) TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res* 7:449–60. [PubMed: 10560968]
- Sasaki T, Fujiwara H (2000) Detection and distribution patterns of telomerase activity in insects. *Eur J Biochem* 267:3025–31. [PubMed: 10806402]
- Sfeir A, de Lange T (2012) Removal of shelterin reveals the telomere end-protection problem. *Science* 336:593–7. [PubMed: 22556254]
- Shay JW, Wright WE (2010) Telomeres and telomerase in normal and cancer stem cells. *FEBS Lett* 584:3819–3825. [PubMed: 20493857]
- Teng SC, Zakian VA (1999) Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol Cell Biol* 19:8083–8093. [PubMed: 10567534]
- Tower J, Karpen GH, Craig N, Spradling AC (1993) Preferential Transposition of *Drosophila*. *Genetics* 133:347–359. [PubMed: 8382177]
- Vitturi R, Lannino A, Mansueto C, et al. (2008) Silver-negative NORs in *Pamphagus ortolaniae* (Orthoptera: Pamphagidae). *Eur J Entomol* 5759:35–39.

- Warchalowska-Sliwa E, Grzywacz B, Maryanska-Nadachowska A, Karamysheva TV, Chobanov DP, Heller K-G (2013) Cytogenetic variability among Bradyporinae species (Orthoptera: Tettigoniidae). *Eur J Entomol* 110:1–12.
- Warchalowska-Sliwa E, Grzywacz B, Maryanska-Nadachowska A, Karamysheva T, Rubtsov NB, Chobanov DP (2009) Chromosomal differentiation among bisexual European species of *Saga* (Orthoptera: Tettigoniidae: Saginae) detected by both classical and molecular methods. *Eur J Biochem* 106:1–9.
- Wiegmann BM, Trautwein MD, Winkler IS, et al. (2011) Episodic radiations in the fly tree of life. *Proc Natl Acad Sci U S A* 108:5690–5. [PubMed: 21402926]
- Wurm Y, Wang J, Riba-Grognuz O, et al. (2011) The genome of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci U S A* 108:5679–84. [PubMed: 21282665]

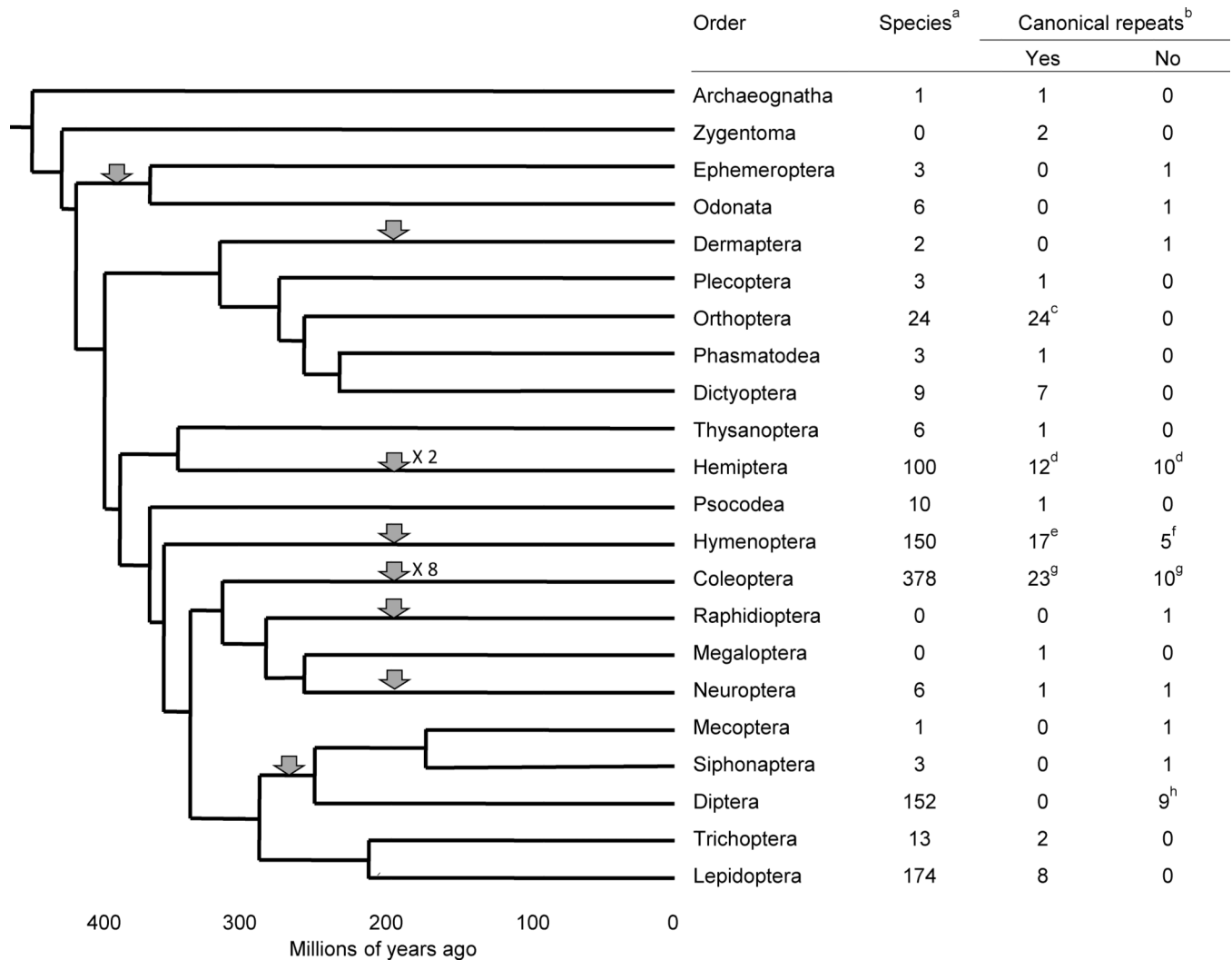


Fig. 1. Chronogram of insect evolution. The chronogram is adapted from Misof et al. (2014) with the following modifications: Mantodea, Blattodea and Isoptera have been combined as Dictyoptera, according to Barnard (2011); and Zoraptera, Mantophasmodea, Grylloblattodea, Embioptera, and Strepsiptera have been omitted, because they are very small and have not been tested for the presence of telomeric repeats. Gray arrowheads indicate points at which a possible loss of terminal (TTAGG)_n repeats occurred. **α** The numbers of insect species are taken from Barnard (2011), except that the number of coleopteran species is from Hunt et al. (2007), and the number of lepidopteran species is from the Lepidoptera Taxome Project (<http://www.ucl.ac.uk/taxome/lepnos.html>). All species numbers are in thousands. **β** When several species from a single genus have been tested all the results have been consistent. Therefore, the numbers in these two columns represent the number of genera tested. Test data are taken from Frydrychova et al. (2004), except where otherwise noted. **γ** Results from Frydrychova et al. (2004), Vitturi et al. (2008), Warchalowska-Sliwa et al. (2009), Jetybayev et al. (2012), Grzywacz et al. (2014), and Warchalowska-Sliwa et al. (2013). **δ** See Figure 3 for references. **ε** Results from

Frydrychova et al. (2004), Lorite et al. (2002), and Wurm et al.(2011). ζ Data from Gokhmant et al. (2014) and Menezes et al. (2013). η See Figure 2 for references. θ Results from Frydrychova et al. (2004) and Madelena et al. (2010).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

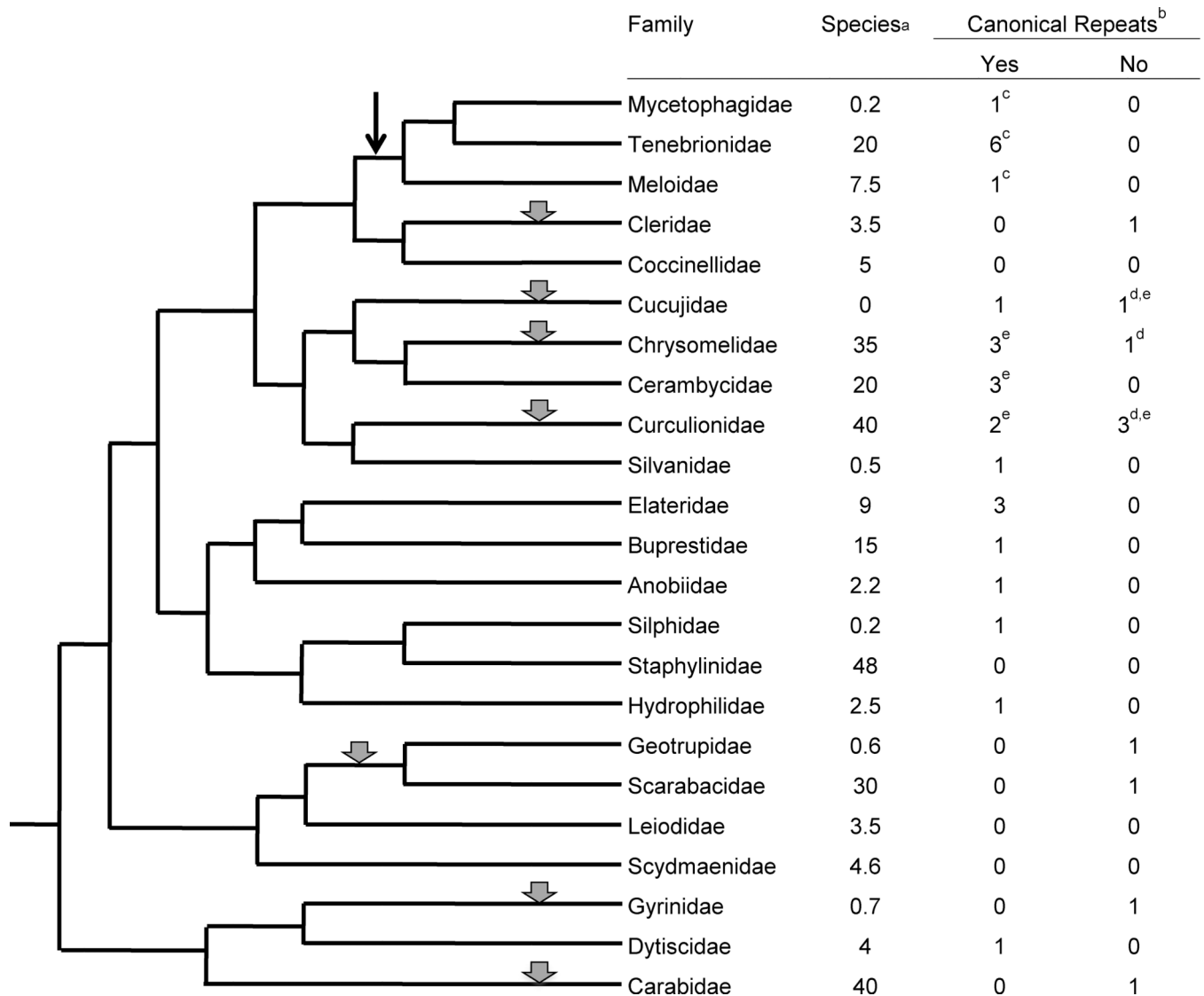


Fig. 2. Phylogenetic tree for Coleoptera. The phylogenetic tree and species numbers in families are from Hunt et al. (2007). Gray arrowheads indicate points at which a possible loss of terminal $(TTAGG)_n$ repeats occurred. Arrow indicates the point at which $(TTAGG)_n$ terminal repeats were replaced by $(TCAGG)_n$ terminal repeats. **a** Species numbers are in thousands. To save space only larger families and families tested for the presence of canonical telomeric DNA sequences are shown. Together, these comprise 290,000 species or 77% of known beetle species. Data on the presence of canonical telomeric repeats are taken from Frydrychova et al. (2004) unless specified otherwise. **b** The numbers of tests in the two columns for canonical repeats indicate the number of genera tested. **c** The canonical arthropod $(TTAGG)_n$ has been replaced in the superfamily Tenebrionoidea (families Mycetophagidae, Tenebrionidae and Meloidae) by $(TCAGG)_n$ (Mravinac et al. 2011). **d** Species in the families Cucujidae, Chrysomelidae and Curculionidae that did not show evidence for $(TTAGG)_n$ at chromosome ends were also tested for the presence of $(TCAGG)_n$, $(TTAGGG)_n$, $(CTGGG)_n$,

(TTGGG)_n, and (CTAGG)_n repeats, all with negative results, suggesting that these species may have unconventional telomeres (Mravinac et al. 2011). e Results from Frydrychova et al. (2004) and Mravinac et al. (2011).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

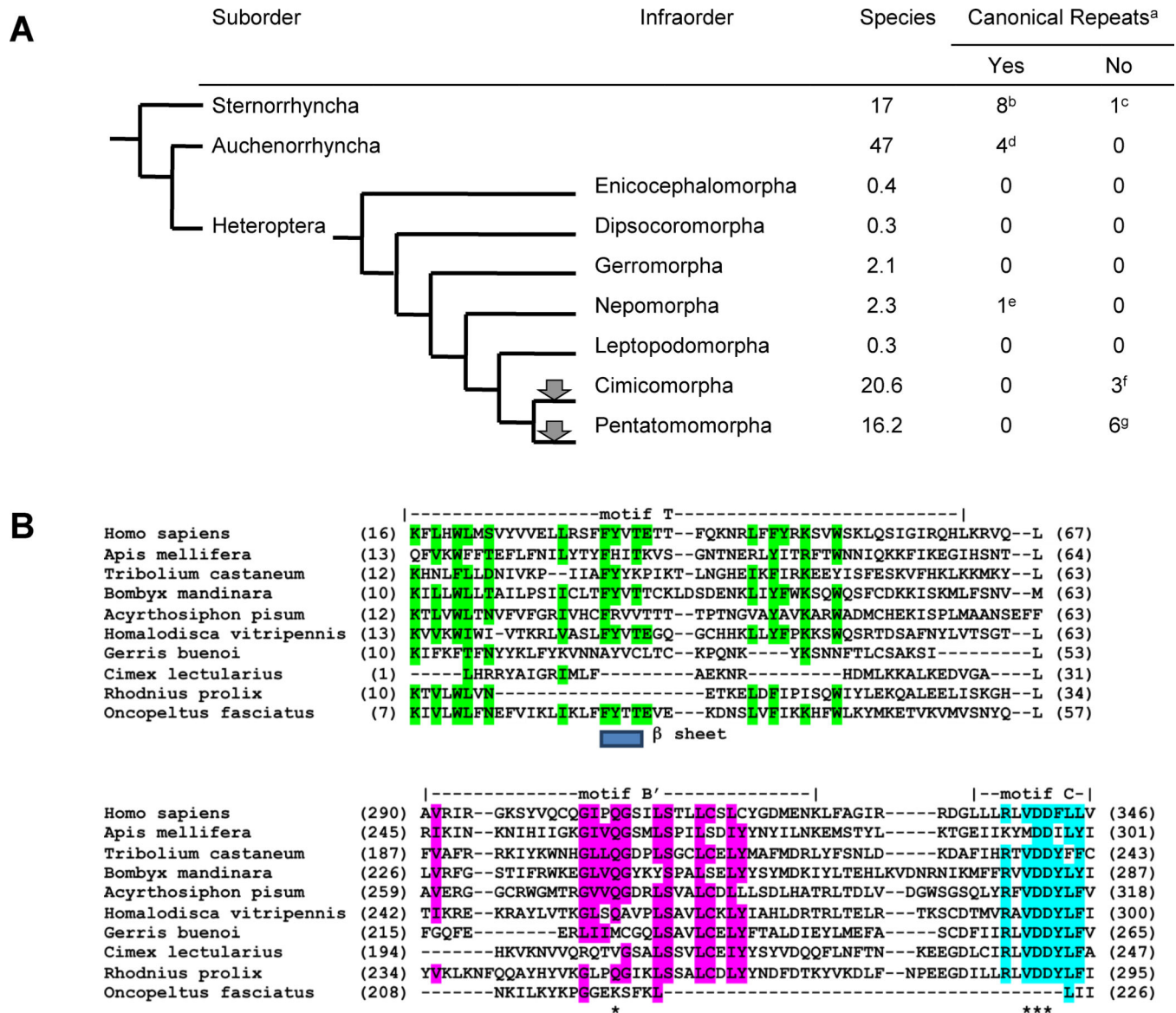


Fig. 3.

Phylogenetic tree for Hemiptera. a The phylogenetic tree is taken from tolweb.com. Gray arrowheads indicate points at which a possible loss of terminal (TTAGG)_n repeats occurred. **α** Species numbers are in thousands. **β** The numbers of tests in the two columns for canonical repeats indicate the numbers of genera tested. **γ** Results from Frydrychova et al. (2004), Mandrioli et al. (2011), Mohan et al. (2011), Monti et al. (2011a), and Monti et al. (2011b). **δ** Results from Novotná et al. (2011). **ε** Results from Frydrychova et al. (2004) and Maryanska-Nadachowska et al. (2013). **ζ** Results from Kuznetsova et al. (2012). **η** Results from Grozeva et al. (2011). **θ** Results from Frydrychova et al. (2004); Grozeva et al. (2011). b Motif T and motifs B and C of the telomerase (TERT) gene are shown from several insect species. *Homo sapiens* is an outgroup. *Apis mellifera* (Hymenoptera), *Tribolium castaneum* (Coleoptera) and *Bombyx mandarina* (Lepidoptera) are distantly related insects with terminal (TTAGG)_n repeats, and thus thought to have an active telomerase gene.

Acyrtosiphon pisum (Hemiptera; Sternorrhyncha), *Homalodisca vitripennis* (Hemiptera: Auchenorrhyncha) and *Gerris buenoi* (Hemiptera; Heteroptera; Gerromorpha) are in taxa shown to have canonical (TTAGG)_n repeats and are thus thought to have an active telomerase gene. *Cimex lectularis* (Hemiptera; Heteroptera; Cimicomorpha), *Rhodnius prolixus* (Hemiptera; Heteroptera; Cimicomorpha) and *Oncopeltus fasciatus* (Hemiptera; Heteroptera; Pentatomomorpha) are in infraorders with members that appear to lack terminal (TTAGG)_n repeats and are inferred to have defective TERT activity. The positions of three relevant motifs as defined in the TERT alignment from <http://telomerase.asu.edu/alignments.html> are shown. Motif T, with highly conserved residues shown in green, contains a conserved β sheet (location of the beta sheet from *H. sapiens* is shown below the alignment). Motif B is shown with highly conserved residues in red. In motif C, highly conserved residues are shown in blue. Asterisks denote important residues discussed in the text. The multiple sequence alignment of the predicted TERT proteins was done with MAFFT 6.849 (Kato et al. 2002). A β sheet in motif T, followed by an a helix, is predicted in all of the proteins except *C. lectularius* and *R. prolixus* suggesting that the deletion of the T domain seen in each could result in a loss of function of TERT. In contrast, *O. fasciatus* has a well conserved T domain but has a large deletion spanning most of the B and C domains. Motif B interacts with the template strand and the glutamine (bold with asterisk) is thought to be involved in substrate specificity (motif B). In motif C, the two asparagines (*) are invariant and the preceding valine is very highly conserved. Alanine mutagenesis has shown these asparagines to be essential for TERT activity.