



OPEN

Telomeric DNA sequences in beetle taxa vary with species richness

Daniela Prušáková^{1,2}, Vratislav Peska³, Stano Pekár⁴, Michal Bubeník³, Lukáš Čížek^{1,2}, Aleš Bezděk¹ & Radmila Čapková Frydrychová^{1,2✉}

Telomeres are protective structures at the ends of eukaryotic chromosomes, and disruption of their nucleoprotein composition usually results in genome instability and cell death. Telomeric DNA sequences have generally been found to be exceptionally conserved in evolution, and the most common pattern of telomeric sequences across eukaryotes is $(T_xA_yG_z)_n$ maintained by telomerase. However, telomerase-added DNA repeats in some insect taxa frequently vary, show unusual features, and can even be absent. It has been speculated about factors that might allow frequent changes in telomere composition in Insecta. Coleoptera (beetles) is the largest of all insect orders and based on previously available data, it seemed that the telomeric sequence of beetles varies to a great extent. We performed an extensive mapping of the $(TTAGG)_n$ sequence, the ancestral telomeric sequence in Insects, across the main branches of Coleoptera. Our study indicates that the $(TTAGG)_n$ sequence has been repeatedly or completely lost in more than half of the tested beetle superfamilies. Although the exact telomeric motif in most of the $(TTAGG)_n$ -negative beetles is unknown, we found that the $(TTAGG)_n$ sequence has been replaced by two alternative telomeric motifs, the $(TCAGG)_n$ and $(TTAGGG)_n$, in at least three superfamilies of Coleoptera. The diversity of the telomeric motifs was positively related to the species richness of taxa, regardless of the age of the taxa. The presence/absence of the $(TTAGG)_n$ sequence highly varied within the Curculionoidea, Chrysomeloidea, and Staphylinoidea, which are the three most diverse superfamilies within Metazoa. Our data supports the hypothesis that telomere dysfunctions can initiate rapid genomic changes that lead to reproductive isolation and speciation.

Telomeres are essential nucleoprotein structures located at the ends of linear eukaryote chromosomes. They play an indispensable role by compensating chromosome shortening, which results from the inability of conventional DNA polymerase to replicate the very end of a DNA molecule, and maintain genome stability by preventing end-to-end fusion or degradation of chromosomes¹. Telomere length is most often maintained by a specialized RNA-dependent DNA polymerase called telomerase, which anneals to the 3' end of chromosomes and repeatedly attaches short non-coding tandem sequences, usually 5–8-bp long, using its internal RNA subunit as a template. The activity of this enzyme is highly regulated and is generally associated with proliferating cells, such as embryonic or germ cells, while activity is low in quiescent, differentiated cells^{2,3}. Telomeres distinguish natural chromosome termini from chromosomal breaks via a multiprotein structure called a telomere cap, and impairment in the telomere cap can lead to end-to-end fusions, genome instability, and cell death. Studies of ciliates, yeasts, vertebrates, and plants show that the formation of telomere caps depends on the interaction of numerous proteins, and some of the proteins bind to chromosome ends in a DNA sequence-dependent manner^{4–6}. Therefore, telomeres of a satisfactory length containing a specific DNA sequence are essential to protect cell viability⁷.

Telomeric DNA sequences have generally been found to be exceptionally conserved in evolution, and the most common sequence pattern across eukaryotes is $(T_xA_yG_z)_n$. For example, the ancestral telomere motif of Metazoa is considered to be $(TTAGGG)_n$, which is present in all basal metazoan groups, i.e., sponges, Cnidaria, Ctenophora, and Placozoa, the unicellular metazoan sister group Choanozoa, and all vertebrates⁸. In contrast, there are taxa in which telomerase-added DNA repeats frequently vary, show unusual features, or have been lost and replaced by a completely different mechanism of telomere loss compensation. For instance, telomeric repeats in closely related species of budding yeast display great diversity in length, sequence, and composition⁹. In addition, there are a number of plant families that have lost the plant-type telomeric sequence $(TTAGGG)_n$,

¹Biology Centre of the Czech Academy of Sciences, Institute of Entomology, Branišovská 31, 370 05 České Budějovice, Czech Republic. ²Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic. ³Department of Cell Biology and Radiobiology, Institute of Biophysics of the Czech Academy of Sciences, Brno, Czech Republic. ⁴Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic. ✉email: Radmila.Frydrychova@hotmail.com

which was replaced by a different one¹⁰. Regarding telomere sequence composition and elongation strategies, unusual features are seen in insects. Although the (TTAGG)_n sequence is considered the ancestral telomeric motif of arthropods, the sequence has been lost in numerous insect orders^{11,12}. However, the uniqueness of Insecta in telomere biology is seen in dipteran species, which have no telomerase and instead employ a completely different telomere elongation system. Dipteran telomere systems involve either a co-option of non-long terminal repeat (non-LTR) retrotransposons, as observed in *Drosophila*^{13,14}, or recombination of satellite sequences, as observed in chironomid midges¹⁵. The current review showed that the (TTAGG)_n telomeric sequence is either partially or completely lost in 12 of the 24 tested insect orders¹⁶, but the DNA sequences present at the telomeres of TTAGG-absent species are largely unknown. In fact, in addition to the dipteran retrotransposon and satellite telomeres, there is only one alternative sequence in Insecta that has been recorded: the (TCAGG)_n sequence in Tenebrionidae (darkling beetles)¹⁷.

Insects are considered to be the most evolutionary successful group of terrestrial organisms in life history, comprising more than one million species, which represents roughly 80% of the species of the world¹⁸. With 155,477 species¹⁸, Diptera is a highly successful order among insects, which shows that the loss of telomerase is not a limiting factor for insect evolution and leads us to a question how widespread non-telomerase systems and the insect telomeric motif are among insects⁷.

Coleoptera is the largest of all insect orders (386,500 species¹⁸), and based on available data, it seems that the telomeric sequence of beetles has evolved dynamically during their evolution. The insect telomeric motif has been lost five to six times during Coleoptera evolution^{19,20}, and it has been replaced by an alternative sequence [(TCAGG)_n] in Tenebrionoidea¹⁷. Furthermore, the occurrence of both sequences varies within families and even genera^{17,20}. The high rate of change in beetle telomere sequences remains poorly understood. We can speculate about the factors that might allow frequent changes, such as high flexibility of the telomere capping complex, which might easily adopt new changes in the telomerase-added sequences, the sequence-independent binding of the telomere capping complex, or simply the evolution history in combination with Coleoptera diversity. However, given that beetles are the most diverse and species-rich group of insects, it should be pointed out that the sampling of representatives (14 families represented by one or two species) in previous studies appears to be inadequate and, therefore, the real variability in the presence of telomeric sequences might be highly underestimated.

The goal of our study was to understand the high rate of change in beetle telomere sequences via an extensive mapping of (TTAGG)_n and (TCAGG)_n telomeric sequences across the main branches of the order, a dramatic increase in species sampling compared to existing studies, identification of novel telomeric repeats, and interpretation of our data with the recently determined phylogenetic relationships, species richness and age of the tested clades.

Results

Distribution of telomeric (TTAGG) and (TCAGG) sequences. The first part of our study was an extensive survey of the distribution of (TTAGG)_n and (TCAGG)_n telomeric sequences across Coleoptera, for which we evaluated 44 families in 15 superfamilies (Table S1), covering the suborders Adephaga and Polyphaga. The mapping was performed using dot-blot hybridization (Fig. 1), Southern hybridization (Fig. 2), and by searching the National Center for Biotechnology Information (NCBI) databases for tandem repeats (Table S3). The obtained data (Table 1) were interpreted based on the recently determined phylogenetic relationships of Coleoptera (Fig. 3) and the species richness and age of taxa (Figs. 3, 4).

In Adephaga, we targeted representatives of Carabidae (Geadephaga), which is the family covering the majority of adephagan species^{21,24–26}. Using dot-blot hybridization, we tested representatives of four subfamilies (Carabinae, Pterostichinae, Licininae, and Platyninae), and the presence of (TTAGG)_n was detected only in one representative of two tested species of Pterostichinae. A further search of Carabidae was conducted using the NCBI database. In Trechinae, the data showed (TTAGG)_n in 17 species and the absence of this sequence in six species, and (TTAGG)_n was absent in Cicindelinae and Harpalinae species. The (TCAGG)_n sequence was not found in any of the Carabidae species examined. Next, we detected (TTAGG)_n in *Rhantus* sp. (Colymbetinae, Dytiscidae) using dot blot hybridization and in Hydroporinae (Dytiscidae, Hydradephaga) in the NCBI database.

We tested one species of Scirtoidea (*Scirtes* sp., Scirtidae), which is a small superfamily of beetles with a basal position within Polyphaga, sharing archaic morphological features with Archostemata and Adephaga^{21,27,28}. Scirtoidea species displayed no hybridization signal for either of the tested sequences.

All tested Elateriformia species representing Throscidae, Eucnemidae, Cantharidae, Elateridae, and Lampyridae showed the presence of (TTAGG)_n and absence of (TCAGG)_n motif by dot-blot hybridization or the NCBI search.

Next, we investigated Staphylinoidea, which includes the majority of the Staphylinoformia species, and we tested subfamilies of Staphylinidae and Leiodidae, which are the two largest Staphylinoidea families²⁹. DNA hybridization experiments revealed the absence of the (TTAGG)_n motif in Leiodinae (Leiodidae), and the NCBI search showed its absence in Oxytelinae and Tachyporinae (both Staphylinidae). In contrast, based on the NCBI search, 19 species of Aleocharinae were found to be (TTAGG)-positive. The (TCAGG)_n sequence was not found in any of the examined species. Together with the published data²⁰, we can conclude that Staphylinoidea varies in the presence/absence of the (TTAGG)_n motif.

Scarabaeoidea is the only superfamily of Scarabaeiformia, herein represented by Geotrupidae, Lucanidae, and Scarabaeidae. None of the tested representatives displayed the presence of either (TTAGG)_n or (TCAGG)_n.

Cucujiformia is an infraorder represented by seven superfamilies (Coccinelloidea, Cucujoidea, Chrysomeloidea, Curculionoidea, Cleroidea, Lymexyloidea, and Tenebrionoidea)³⁰. We detected the (TCAGG)_n as the alternative telomeric motif in Tenebrionoidea families: Pyrochroidae, Meloidae, Salpingidae, Oedemeridae, Mordellidae, Ciidae, Mycetophagidae, Tenebrionidae, and Anthicidae. Moreover, our data showed that the (TTAGG)_n

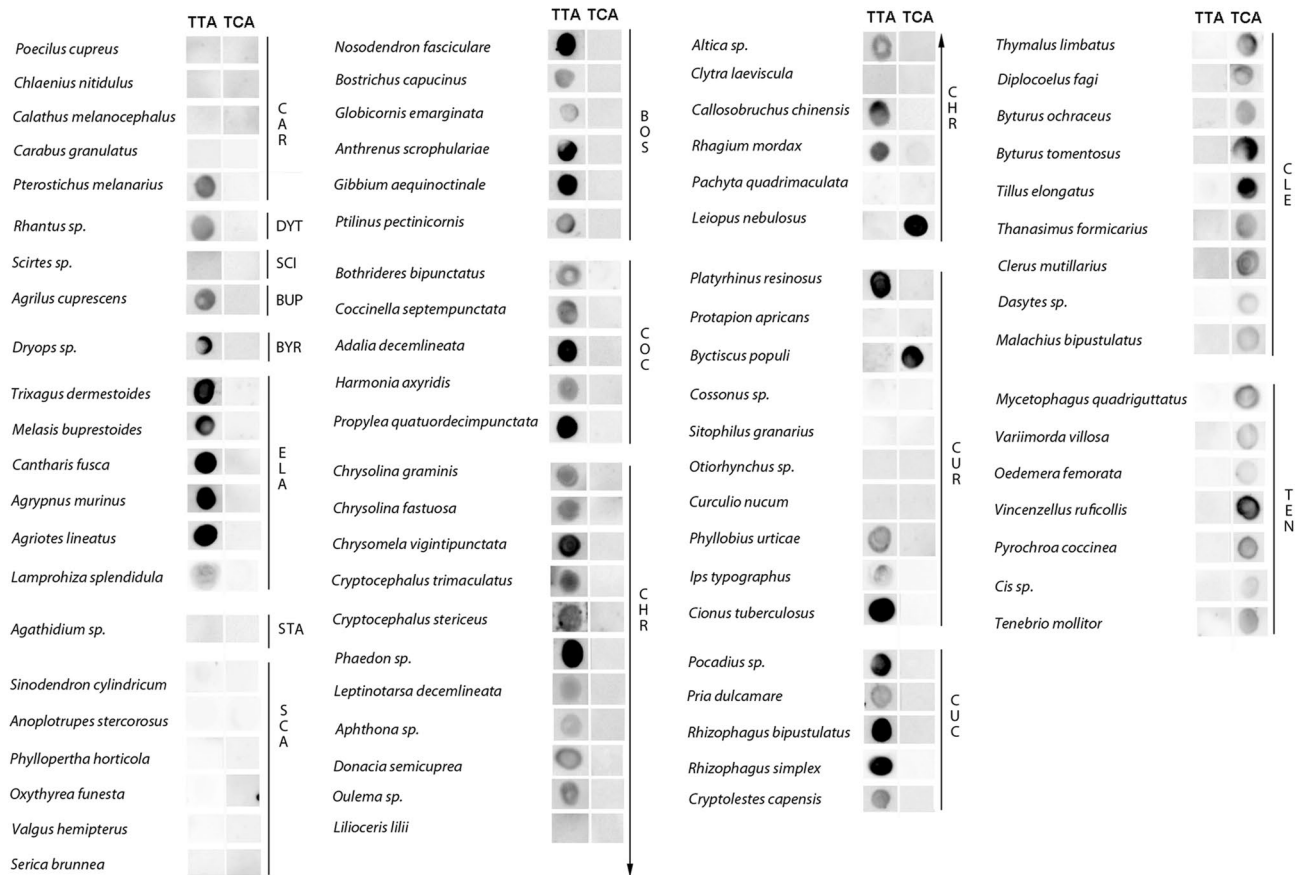


Figure 1. Dot-blot hybridizations. The presence of $(TTAGG)_n$ and $(TCAGG)_n$ sequences were examined using dot-blot hybridizations. *TTA* the $(TTAGG)_n$ sequence, *TCA* the $(TCAGG)_n$ sequence, *CAR* Caraboidea, *DYT* Dytiscoidea, *SCI* Scirtoidea, *BUP* Buprestoidea, *BYR* Byrrhoidea, *ELA* Elateroidea, *STA* Staphylinoidea, *SCA* Scarabaeoidea, *BOS* Bostrichoidea, *COC* Coccinelloidea, *CHR* Chrysomeloidea, *CUR* Curculionioidea, *CUC* Cucujoidea, *CLE* Cleroidea, *TEN* Tenebrionoidea.

sequence had been replaced with the $(TCAGG)_n$ motif in all representatives of the Cleroidea superfamily, which has been recently reported as a sister group of Tenebrionoidea²¹. The tested species of Cleroidea represented Melyridae, Cleridae, and Trogossitidae as well as Biphyllidae and Byturidae, two additional families recently transferred into Cleroidea^{28,31}. The presence of the $(TCAGG)_n$ motif at the chromosome termini of Cleroidea was confirmed using Bal31 digestion and fluorescence in situ hybridization (FISH) (Figs. 2c, 3c).

Next, we tested Cucujoidea, which is an extremely diverse and taxonomically difficult superfamily³². Despite this diversity, all tested families of Cucujoidea (Cucujidae, Monotomidae, Silvanidae, and Nitidulidae) had the $(TTAGG)_n$ sequence. The presence of $(TTAGG)_n$ motif was also detected in all tested representatives of the Coccinelloidea (Bothrideridae and Coccinellidae). None of the tested representatives displayed the presence of $(TCAGG)_n$.

In both Curculionioidea and Chrysomeloidea, which are considered sister groups³², the occurrence of the $(TTAGG)_n$ sequence highly varied. In Curculionioidea, while the sequence was detected in Anthribidae, no signal was observed in Brentidae and Attelabidae, and variability in sequence occurrence was observed at family and subfamily levels within Curculionidae. Consistent with previous findings, Curculionidae displayed the presence of the $(TTAGG)_n$ sequence in *Ips typographus* (Scolytinae; consistent with³³) and *Phyllobius urticae* (Entiminae, Phyllobiini; consistent with¹⁷). However, based on hybridization data and the NCBI database Scolytinae, Entiminae and Curculioninae showed variability in the $(TTAGG)_n$ occurrence, and the sequence was not detected in Dryophthorinae and Cossoninae. In Chrysomeloidea, instability in the $(TTAGG)_n$ presence was observed at family and subfamily levels within Chrysomelidae and Cerambycidae. Based on our data, the sequence was absent in four of the 22 tested species of Chrysomelidae within Criocerinae (in one of two tested species), Galerucinae (in two of four tested species), in one tested species of Clytrinae, and three of six Cerambycidae representatives within Lepturinae (in one of three tested species) and two tested species of Lamiinae. Interestingly, the $(TCAGG)_n$ motif was detected in two representatives of Phytophaga: *Byctiscus populi* (Curculionioidea, Attelabidae) and *Leiopus nebulosus* (Chrysomeloidea, Cerambycidae).

In the superfamily Bostrichoidea, which is a sister group of Cucujiformia³¹, all tested families (Bostrichidae, Anobiidae, Dermestidae, Nosodendridae, and Ptinidae) showed the presence of the $(TTAGG)_n$ sequence and the absence of the $(TCAGG)_n$ sequence.

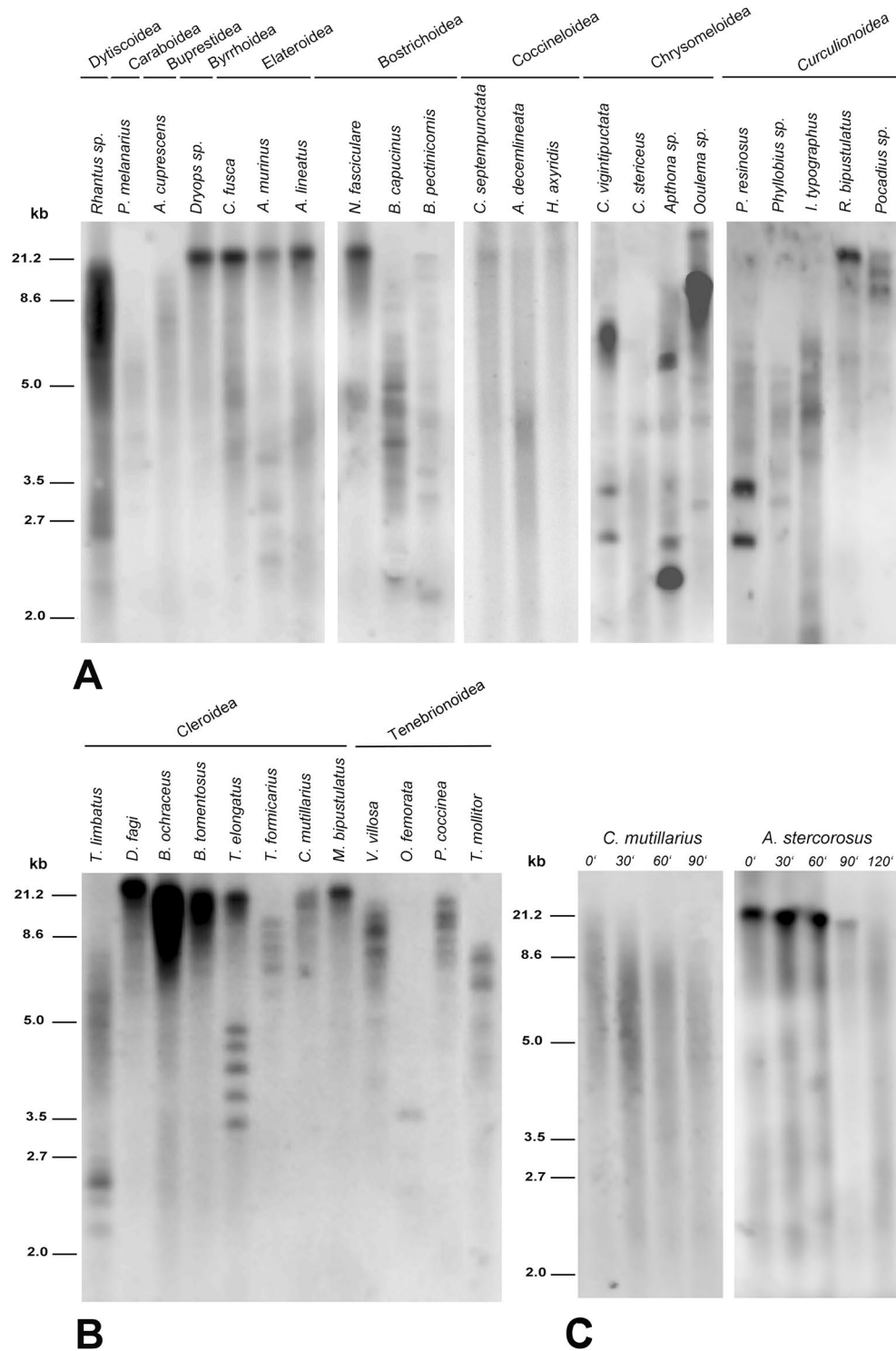


Figure 2. The character of hybridization signals revealed using Southern hybridization. The DIG-labeled probe specific to (TTAGG)_n (a) and (TCAGG)_n (b) sequence was hybridized with RsaI/HinfI-digested genomic DNAs of selected coleopteran species. (c) Bal 31 exonuclease digestion of genomic DNA of *Anoplotrupes stercorosus* and *Clerus mutillarius*. Time-course digestion intervals are indicated above the lanes.

A more detailed characterization of the hybridization signals was conducted using Southern hybridization in selected representatives of the tested taxa. The signals were formed by long smears, mostly ranging from 2 kb to more than 21 kb, and mostly revealed numbers of hybridization bands of different molecular weights (Fig. 2a,b).

Superfamily	Family/subfamily	Species	TTAGG	TCAGG	Source
Caraboidea	Carabidae				
	Harpalinae	<i>Selenophorus alternans</i>	(-) ^a	(-) ^a	
	Licinae	<i>Chlaenius nitidulus</i>	(-)	(-)	
	Cicindelinae	<i>Tetracha</i> sp.	(-) ^a	(-) ^a	
	Carabinae	<i>Carabus granulatus</i>	(-)	(-)	
	Platyninae	<i>Calathus melanocephalus</i>	(-)	(-)	
		<i>Miquihuaana rhandiniformis</i>	(+) ^a	(-) ^a	
	Trechinae	<i>Bembidarens</i> sp.	(+) ^a	(-) ^a	
		<i>Bembidion chlioperiphys</i>	(+) ^a	(-) ^a	
		<i>B. orion</i>	(+) ^a	(-) ^a	
		<i>B. castor</i>	(+) ^a	(-) ^a	
		<i>B. aeruginosum</i>	(-) ^a	(-) ^a	
		<i>B. musae</i>	(+) ^a	(-) ^a	
		<i>B. clemens</i>	(+) ^a	(-) ^a	
		<i>B. breve</i>	(-) ^a	(-) ^a	
		<i>B. laxatum</i>	(-) ^a	(-) ^a	
		<i>B. obscuripenne</i>	(-) ^a	(-) ^a	
		<i>B. lapponicum</i>	(-) ^a	(-) ^a	
		<i>B. flohri</i>	(+) ^a	(-) ^a	
		<i>B. scenicum</i>	(+) ^a	(-) ^a	
		<i>B. lachnophoroides</i>	(+) ^a	(-) ^a	
		<i>B. ulkei</i>	(+) ^a	(-) ^a	
		<i>B. subfusum</i>	(+) ^a	(-) ^a	
		<i>B. aeruginosum</i>	(+) ^a	(-) ^a	
		<i>B. eupetedromus</i>	(+) ^a	(-) ^a	
		<i>Lionepha casta</i>	(+) ^a	(-) ^a	
		<i>L. chintimini</i>	(+) ^a	(-) ^a	
		<i>L. erasa</i>	(-) ^a	(-) ^a	
		<i>L. lindrothellus</i>	(+) ^a	(-) ^a	
		<i>Pogonus chalceus</i>	(+) ^a	(-) ^a	
	Pterostichinae	<i>Pterostichus melanarius</i>	(+)	(-)	
		<i>P. oblongopunctatus</i>	(-)	Untested	^b
<i>Pseudamara arenaria</i>		(+) ^a	(-) ^a		
<i>Poecilus cupreus</i>		(-)	(-)		
Dytiscoidea	Dytiscidae				
	Colymbetinae	<i>Rhantus</i> sp.	(+)	(-)	
	Hydroporinae	<i>Stictotarsus aequinoctialis</i>	(+) ^a	(-) ^a	
		<i>Graphoderus cinereus</i>	(+)	Untested	^b
Gyrinidae:	<i>Orectochilus villosus</i>	(-)	Untested	^b	
Scirtoidea	Scirtidae: Scirtinae	<i>Scirtes</i> sp.	(-)	(-)	
Buprestoidea	Buprestidae				
	Agrilinae	<i>Agrilus cuprescens</i>	(+)	(-)	
Byrrhoidea	Dryopidae	<i>Agrilus planipennis</i>	(+) ^a	(-) ^a	
		<i>Dryops</i> sp.	(+)	(-)	
Elateroidea	Throscidae: Throscinae	<i>Trixagus dermestoides</i>	(+)	(-)	
	Eucnemidae: Melasinae	<i>Melasis buprestooides</i>	(+)	(-)	
	Cantharidae: Cantharinae	<i>Cantharis fusca</i>	(+)	(-)	
	Elateridae				
	Elaterinae	<i>Agriotes lineatus</i>	(+)	(-)	
		<i>Ampedus sanguineus</i>	(+)	Untested	^b
	Dentrometrinae	<i>Diacanthous undosus</i>	(+)	Untested	^c
	Agrypninae	<i>Agrypnus murinus</i>	(+)	(-)	
	Melanotinae	<i>Melanotus legatus</i>	(+)	Untested	^c
	Lampyridae				
Continued					

Superfamily	Family/subfamily	Species	TTAGG	TCAGG	Source	
	Lampyrinae	<i>Lamprohiza splendidula</i>	(+)	(-)		
		<i>Ellychnia corrusca</i>	(+) ^a	(-) ^a		
		<i>Lucidota atra</i>	(+) ^a	(-) ^a		
		<i>L. punctata</i>	(+) ^a	(-) ^a		
		<i>Phausis reticulata</i>	(+) ^a	(-) ^a		
		<i>Photinus australis</i>	(+) ^a	(-) ^a		
		<i>P. brimleyi</i>	(+) ^a	(-) ^a		
		<i>P. carolinus</i>	(+) ^a	(-) ^a		
		<i>P. cooki</i>	(+) ^a	(-) ^a		
		<i>P. indictus</i>	(+) ^a	(-) ^a		
		<i>P. macdermotti</i>	(+) ^a	(-) ^a		
		<i>P. obscurellus</i>	(+) ^a	(-) ^a		
		<i>P. scintillans</i>	(+) ^a	(-) ^a		
		<i>P. pyralis</i>	(+) ^a	(-) ^a		
		<i>Pyractomena borealis</i>	(+) ^a	(-) ^a		
		<i>Pyractomena angulata</i>	(+) ^a	(-) ^a		
		<i>Pyractomena. marginalis</i>	(+) ^a	(-) ^a		
		<i>Pyropyga decipiens</i>	(+) ^a	(-) ^a		
		Luciolinae	<i>Aquatica lateralis</i>	(+) ^a	(-) ^a	
Photurinae	<i>Photuris frontalis</i>	(+) ^a	(-) ^a			
Staphylinoidea	Leiodidae					
	Leiodinae	<i>Agathidium</i> sp.	(-)	(-)		
	Silphidae	<i>Silpha obscura</i>	(+)	Untested	^b	
		<i>Nicrophorus orbicollis</i>	(+) ^a	(-) ^a		
	Staphylinidae					
	Pselaphinae	<i>Adranes taylori</i>	(+) ^a	(-) ^a		
	Oxytelinae	<i>Carpelimus</i> sp.	(-) ^a	(-) ^a		
	Aleocharinae	<i>Aenictocupidus jacobsonorum</i>	(+) ^a	(-) ^a		
		<i>Mimaenictus wilsoni</i>	(+) ^a	(-) ^a		
		<i>Weissflogia rhopalogaster</i>	(+) ^a	(-) ^a		
		<i>Dalotia coriaria</i>	(+) ^a	(-) ^a		
		<i>Earota dentata</i>	(+) ^a	(-) ^a		
		<i>Diploeiton nevermanni</i>	(+) ^a	(-) ^a		
		<i>Deinopsis erosa</i>	(+) ^a	(-) ^a		
		<i>Dorylogaster longipes</i>	(+) ^a	(-) ^a		
		<i>Ecitomorpha arachnoides</i>	(+) ^a	(-) ^a		
		<i>Ecitophya simulans</i>	(+) ^a	(-) ^a		
		<i>Drusilla canaliculata</i>	(+) ^a	(-) ^a		
		<i>Ecitoglossa quadriceps</i>	(+) ^a	(-) ^a		
		<i>Platyusa sonomae</i>	(+) ^a	(-) ^a		
		<i>Tetradonia laticeps</i>	(+) ^a	(-) ^a		
		<i>Labidoglobus nevermanni</i>	(+) ^a	(-) ^a		
		<i>L. appendiculatus</i>	(+) ^a	(-) ^a		
		<i>Pseudomimeciton antennatum</i>	(+) ^a	(-) ^a		
		<i>Oxypoda opaca</i>	(+) ^a	(-) ^a		
		<i>Sceptobius lativentris</i>	(+) ^a	(-) ^a		
	Tachyporinae	<i>Coproporus ventriculus</i>	(-) ^a	(-) ^a		
	Scarabaeoidea	Geotrupidae		<i>Anoplotrupes stercorosus</i>	(-)	(-)
				<i>Geotrupes stercorarius</i>	(-)	Untested ^b
Scarabaeidae						
Rutelinae		<i>Phyllopertha horticola</i>	(-)	(-)		
		<i>Chrysina resplendens</i>	(-) ^a	(-) ^a		
		<i>Popillia japonica</i>	(-) ^a	(-) ^a		
Valginae		<i>Valgus hemipterus</i>	(-)	(-)		
Dynastinae	<i>Oryctes borbonicus</i>	(-) ^a	(-) ^a			
Continued						

Superfamily	Family/subfamily	Species	TTAGG	TCAGG	Source	
	Scarabaeinae	<i>Canthidium</i> sp.	(-) ^a	(-) ^a		
		<i>Onthophagus taurus</i>	(-) ^a	(-) ^a		
		<i>Pachysoma endroedyi</i>	(-) ^a	(-) ^a		
		<i>P. striatus</i>	(-) ^a	(-) ^a		
	Sericinae	<i>Serica brunnea</i>	(-)	(-)		
	Cetoniinae	<i>Oxythyrea funesta</i>	(-)	(-)		
	Lucanidae					
	Syndesinae	<i>Sinodendron cylindricum</i>	(-)	(-)		
Bostrichoidea	Bostrichidae: Bostrichinae	<i>Bostrichus capucinus</i>	(+)	(-)		
	Anobiidae: Anobiinae	<i>Stegobium paniceum</i>	(+)	Untested	^b	
	Dermestidae					
	Megatomiinae	<i>Globicornis emarginata</i>	(+)	(-)		
		<i>Anthrenus scrophulariae</i>	(+)	(-)		
	Ptinidae					
	Gibiinae	<i>Gibbium aequinoctinale</i>	(+)	(-)		
	Ptiliniinae	<i>Ptilinus pectinicornis</i>	(+)	(-)		
	Nosodendridae:	<i>Nosodendron fasciculare</i>	(+)	(-)		
	Tenebrionoidea	Mordellidae: Mordellinae	<i>Variimorda villosa</i>	(-)	(+)	
Oedemeridae: Oedemerinae		<i>Oedemera femorata</i>	(-)	(+)		
Salpingidae: Salpinginae		<i>Vincenzellus ruficollis</i>	(-)	(+)		
Pyrochroidae: Pyrochroinae		<i>Pyrochroa coccinea</i>	(-)	(+)		
Mycetophagidae						
Mycetophaginae		<i>Mycetophagus</i> sp.	(-)	(+)		
		<i>Typhaea stercorea</i>	(-)	(+)	^d	
Ciidae: Ciinae		<i>Cis</i> sp.	(-)	(+)		
Tenebrionidae						
Tenebrioninae		<i>Tenebrio mollitor</i>	(-)	(+)		
		<i>Phylacinus fisheri</i>	(-) ^a	(+) ^a		
		<i>Tribolium castaneum</i>	(-) ^a	(+) ^a	^d	
		<i>T. freemani</i>	(-)	(+)	^d	
		<i>T. confosum</i>	(-)	(+)	^d	
		<i>T. audax</i>	(-)	(+)	^d	
		<i>T. brevicornis</i>	(-)	(+)	^d	
		<i>T. anaphe</i>	(-)	(+)	^d	
		<i>T. destructor</i>	(-)	(+)	^d	
Pimeliinae		<i>Pimelia elevata</i>	(-)	(+)	^d	
	<i>P. cribra</i>	(-)	(+)	^d		
Meloidae						
Meloinae	<i>Hycleus cichorii</i>	(-) ^a	(+) ^a			
	<i>H. phaleratus</i>	(-) ^a	(+) ^a	^d		
Anthicidae	Anthicidae sp.	(-) ^a	(+) ^a			
Cleroidea	Melyridae					
	Malachiinae	<i>Malachius bipustulatus</i>	(-)	(+)		
	Dasytinae	<i>Dasytes</i> sp.	(-)	(+)		
	Cleridae					
	Clerinae	<i>Clerus mutillarius</i>	(-)	(+)		
		<i>Thanasimus formicarius</i>	(-)	(+)		
	Tillinae	<i>Tillus elongatus</i>	(-)	(+)		
	Trogossitidae: Peltinae	<i>Thymalus limbatus</i>	(-)	(+)		
	Biphyllidae	<i>Diplocoelus fagi</i>	(-)	(+)		
	Byturidae					
	Byturinae	<i>Byturus ochraceus</i>	(-)	(+)		
<i>B. tomentosus</i>		(-)	(+)			
Continued						

Superfamily	Family/subfamily	Species	TTAGG	TCAGG	Source
Coccinelloidea	Bothriideridae: Bothriiderinae	<i>Bothrioides bipunctatus</i>	(+)	(-)	
	Coccinellidae				
	Coccinellinae	<i>Coccinella septempunctata</i>	(+)	(-)	
		<i>Adalia decemlineata</i>	(+)	(-)	
		<i>Harmonia axyridis</i>	(+) ^e	(-) ^e	
		<i>Propylea quatuordecimpunctata</i>	(+)	(-)	
		<i>Eriopsis connexa</i>	(+) ^a	(-) ^a	
Scymninae	<i>Nephaspis</i> sp.	(+) ^a	(-) ^a		
Cucujoidea	Monotomidae				
	Rhizophaginae	<i>Rhizophagus bipustulatus</i>	(+)	(-)	
		<i>R. simplex</i>	(+)	(-)	
	Nitidulidae				
	Meligethinae	<i>Pria dulcamare</i>	(+)	(-)	
	Nitidulinae	<i>Pocadius</i> sp.	(+)	(-)	
	Silvanidae: Silvaninae	<i>Oryzaephilus surinamensis</i>	(+)	(-)	^d
	Cucujidae: Laemophloinae	<i>Cryptolestes capensis</i>	(+)	(-)	
Chrysomeloidea	Chrysomelidae				
	Chrysomelinae	<i>Chrysolina graminis</i>	(+)	(-)	
		<i>C. fastuosa</i>	(+)	(-)	
		<i>C. polita</i>	(+)	Untested	^d
		<i>C. americana</i>	(+)	Untested	^d
		<i>C. herbacea</i>	(+)	Untested	^d
		<i>Chrysomela vigintipunctata</i>	(+)	(-)	
		<i>Plagioderia versicola</i>	(+) ^a	(-) ^a	
		<i>Phaedon</i> sp.	(+)	(-)	
		<i>Gonioctena quinquepunctata</i>	(+) ^a	(-) ^a	
		<i>Leptinotarsa decemlineata</i>	(+) ^e	(-) ^e	
	Bruchinae	<i>Callosobruchus chinensis</i>	(+) ^e	(-) ^e	
		<i>Acanthoscelides obtectus</i>	(+)	Untested	^d
	Cryptocephalinae	<i>Cryptocephalus trimaculatus</i>	(+)	(-)	
		<i>C. stericus</i>	(+)	(-)	
	Eumolpinae	<i>Colaspis</i> sp.	(+) ^a	(-) ^a	
	Galerucinae	<i>Diabrotica barberi</i>	(-) ^a	(-) ^a	
		<i>D. virgifera</i>	(-) ^a	(-) ^a	
		<i>Aphthona</i> sp.	(+)	(-)	
		<i>Altica</i> sp.	(+)	(-)	
	Donaciinae	<i>Donacia semicuprea</i>	(+)	(-)	
	Criocerinae	<i>Liliocercis lili</i>	(-)	(-)	
		<i>Oulema</i> sp.	(+)	(-)	
	Clytrinae	<i>Clytra laeviscula</i>	(-)	(-)	
	Cerambycidae				
	Lepturinae	<i>Pachyta quadrimaculata</i>	(-)	(-)	
		<i>Rhagium mordax</i>	(+)	(-)	
		<i>Rhagium inquisitor</i>	(+)	Untested	^d
	Lamiinae	<i>Leiopus nebulosus</i>	(-)	(+)	
		<i>Anoplophora glabripennis</i>	(-) ^a	(-) ^a	
Cerambycinae	<i>Phymatodes lengi</i>	(+) ^a	(-) ^a		
Aseminae	<i>Arhopalus coreanus</i>	(+)	Untested	^c	
Spondyliinae	<i>Spondylis buprestoides</i>	(+)	Untested	^c	
Continued					

Superfamily	Family/subfamily	Species	TTAGG	TCAGG	Source	
Curculionoidea	Anthribidae	<i>Platyrhinus resinosus</i>	(+)	(-)		
	Attelabidae: Rhynchitinae	<i>Byctiscus populi</i>	(-)	(+)		
	Brentidae	<i>Protapion apricans</i>	(-)	(-)		
	Curculionidae					
	Entiminae		<i>Otiorynchus</i> sp.	(-)	(-)	
			<i>Phyllobius</i> sp.	(+)	(-)	^d
			<i>Phyllobius urticae</i>	(+)	(-)	
	Dryophthorinae	<i>Sitophilus granarius</i>	(-)	(-)		
	Cossoninae	<i>Cossonus</i> sp.	(-)	(-)		
	Curculioninae		<i>Curculio nucum</i>	(-)	(-)	
			<i>Cionus tuberculosus</i>	(+)	(-)	
			<i>C. nigritarsis</i>	(-)	(-)	^d
			<i>Anthonomus grandis</i>	(-) ^a	(-) ^a	
	Scolytinae		<i>Dendroctonus jeffreyi</i>	(-) ^a	(-) ^a	
			<i>D. ponderosae</i>	(-) ^a	(-) ^a	
			<i>Hypothenemus hampei</i>	(+) ^a	(-) ^a	
		<i>Ips typographus</i>	(+)	(-)		

Table 1. A list of coleopteran species tested for the presence of (TTAGG)_n and (TCAGG)_n species. (+) signal present, (-) signal absent. ^aData extracted from Database. ^bData from Frydrychova et Marec²⁰. ^cData from Okazaki et al.¹⁹. ^dData from Mravinac et al.¹⁷. ^eData based on the NCBI and dot-blot hybridization.

Collectively, our study indicates that the (TTAGG)_n sequence has been repeatedly or completely lost in Geod-ephaga and more than half of the tested polyphagan superfamilies (Fig. 3a,b), and the sequence was replaced with (TCAGG)_n motif in sister groups Tenebrionoidea and Cleroidea.

Search for telomeric sequence variants: the (TTGGG)_n sequence as a novel telomeric sequence in beetles.

A search of the NCBI databases for tandem repeats revealed that the (TTGGG)_n and (TGAGG)_n sequences are candidates for novel telomeric sequences in Coleoptera (Table S3). Using dot blot hybridization, we examined the TTAGG- and TCAGG-negative species for the presence of these sequences. Besides, we examined these species for the presence of several other sequence variants that have been reported as telomeric motifs in different organisms, including (TTTGGG)_n³⁴, (TTGGGG)_n³⁵ (from ciliate protozoans), (TTAGGG)_n (from vertebrates)³⁶, and TTTAGGG (from plants)³⁷. Except for (TTGGG)_n, no positive signals were detected for any of the other tested sequences. Consistent with the NCBI database showing the presence of the (TTGGG)_n in three scarabaeoid species (*Onthophagus taurus*, *Pachysoma striatus*, and *Canthidium* sp.), the sequence was detected in *Anoplotrupes stercorosus*, and the presence of the sequence on chromosome termini was confirmed using Bal31 digestion and FISH (Figs. 2c, 3d,e). Surprisingly, (TTGGG)_n has not been confirmed in other tested scarabaeoid species (not shown).

Repeated losses of telomeric motifs in certain taxa reflect the species richness but not the age of the taxa.

The presence/absence of the (TTAGG)_n sequence varied within the Curculionoidea, Chrysomeloidea, and Staphylinoidea, which are the three most diverse metazoan superfamilies. In these taxa, variance in the presence of (TTAGG)_n was observed even at the family or subfamily level, that is, within Staphylinidae (> 63,000 species³⁸), Chrysomelidae (> 32,000 species²¹), Cerambycidae (> 32,000 species), and Curculionidae (> 51,000 species²¹) (Fig. 3). Similarly, the occurrence of (TTAGG)_n varied in the highly diverse Carabidae (> 34,000 species²¹). In contrast, no variance was observed in, for instance, the less diverse, but highly sampled, Tenebrionoidea, Cleroidea, or Elateroidea.

To test the hypothesis that more speciose taxa have greater telomere diversity, we used the GLS method. We found that there was a significant effect of motif diversity on the species richness per taxon (GLS, $F_{1,12} = 11.2$, $P = 0.006$, Fig. 1); the diversity increased exponentially with richness (Fig. 4). We also tested the effect of age of the taxon; the effect of age was not significant (GLS, $F_{1,11} = 1.3$, $P = 0.28$).

Discussion

There are two main hypotheses to explain the number of species in a clade^{39–42}. The clade-age hypothesis proposes that species richness increases with the age of the clade. According to the diversification-rate hypothesis, the number of species depends on net diversification, which reflects speciation minus extinction over time; thus, old clades with low richness have low net diversification rates, while young clades with a high number of species have high rates of diversification. In our study, we observed a positive relationship between species number and diversity in the telomere sequence in clades, regardless of the age of the clades. Therefore based on our data, we hypothesize that the key reason for the frequently observed telomeric losses in certain coleopteran clades is the

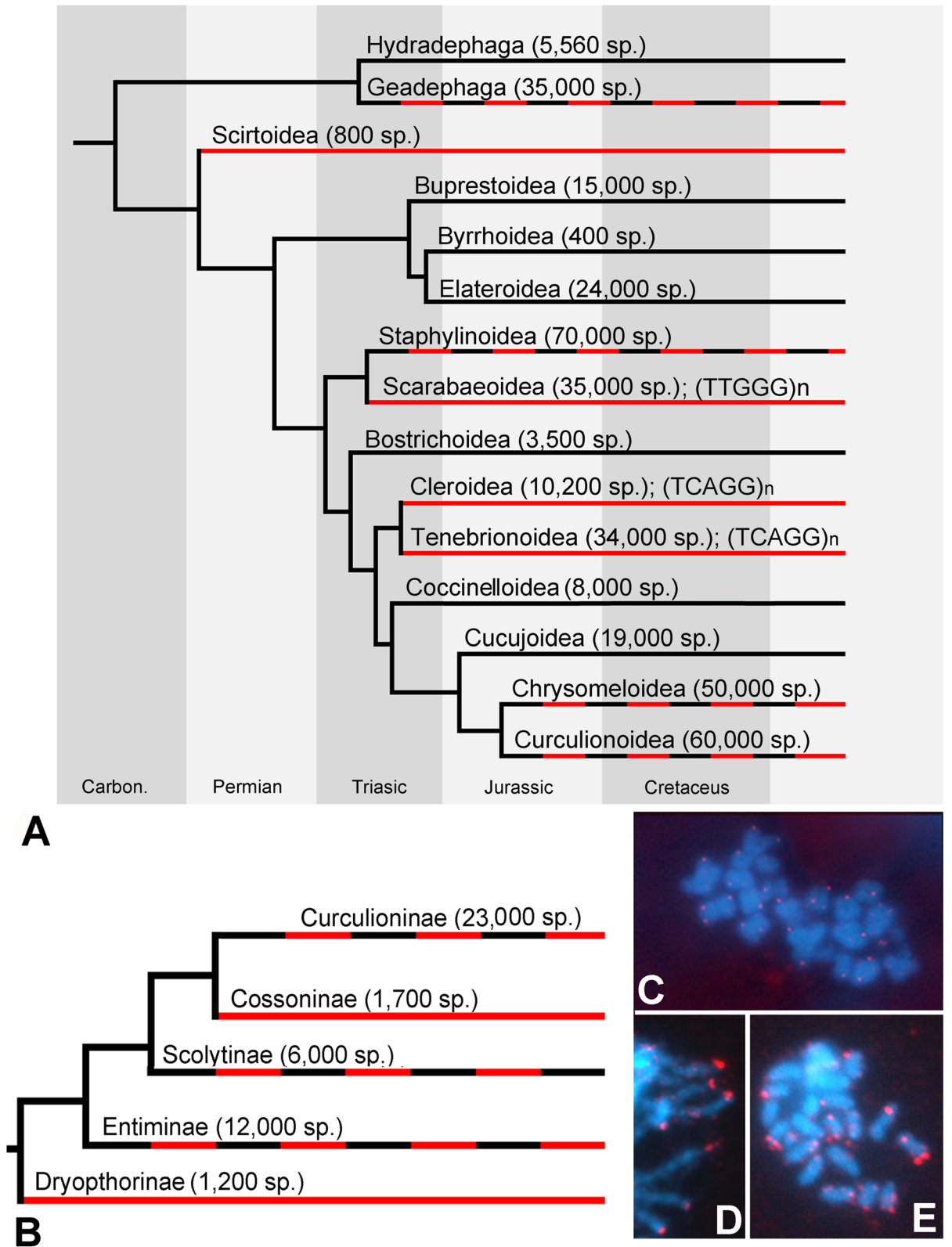


Figure 3. The presence of (TTAGG)_n telomeric sequence in Coleoptera. The (TTAGG)_n distribution in the whole order Coleoptera adopting the Coleoptera cladogram published by²¹ (a) and in the family Curculionidae based on a cladogram published by²² (b). Branch colors indicate the (TTAGG)_n presence (in black), (TTAGG)_n absence (in red), or variability in the (TTAGG)_n presence/absence (in black and red). Approximate numbers of described extant species are indicated next to the family or subfamily level taxon names. (c) Fluorescence in situ hybridization with a probe against the (TCAGG)_n in *Clerus mutillarius* (Cleroidea, Cleridae). (d, e) Fluorescence in situ hybridization with a probe against the (TTGGG)_n in *Anoplotrupes stercorosus* (Scarabaeoidea, Geotrupidae).

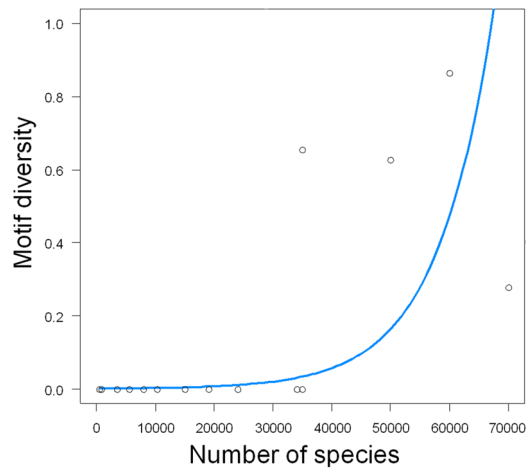


Figure 4. Relationship between motif diversity and species richness in taxa. The blue line represents the estimated model from GLS. The analysis was performed in the R environment (<http://www.r-project.org/index.html>) using R version 4.1.0²³.

high species diversity within the clades; further, it should be considered that the frequency of sequence loss might be in some cases accentuated by insufficiently clear phylogenetic relationships in clades, such as those within Chrysomeloidea³², Adephaga^{43,44} or between Scarabaeiformia and Staphyliniformia^{26,45–50}.

The positive correlation between species richness and diversity in telomeric sequences seems to be consistent with the highly frequent (TTAGG)_n losses within the Insecta, with 1,020,007 species, representing about 66% of all animals, and the variance in telomeric sequences is observed in one of the most diverse insect orders, which are, along with Coleoptera, Hemiptera (104,000 species), Hymenoptera (117,000 species), and Diptera (156,000 species)^{16,51}. In contrast, 57 tested representatives of Orthoptera (27,000 species³⁹) had the (TTAGG)_n sequence¹⁶. Certain similarities can be observed in other highly diverse classes of organisms, such as Magnoliophyta (flowering plants, 300,000 species⁵²). Loss of the plant ancestral telomere DNA sequence, (TTTAGGG)_n, has been reported in numerous flowering plants, in which the sequences were replaced by alternative motifs^{10,53–57}, and high diversity in telomere repeats was observed in two species-rich orders, Asparagales, the largest order within the monocotyledons, consisting of around 30,000 species^{58–61}, and Lamiales, consisting of 23,000 species⁵⁵. Besides, numerous plant orders possess unknown telomere sequences¹⁰. It is also debatable whether the positive correlation between species richness and telomere sequence diversity is due to the direct involvement of telomere biology in the process of speciation, or if it is simply because the diversity in species-rich clades can be detected more easily.

Telomeric DNA sequences interfere in the formation of a telomere capping structure⁶², and therefore, we can assume that a small deviation in the telomere sequence can affect whole-genome stability and have an enormous evolutionary impact. It is well known that telomere dysfunction results in chromosomal fusions, large-scale genomic rearrangements, and instability across the genome, which are all hallmarks of speciation. Chromosomal rearrangements are associated with high mortality and, due to meiotic anomalies, reduced fertility, which collectively results in a generation of isolated groups within the population; each of the groups could develop into a new species⁶³. Therefore, we can suppose that telomeres might provide a powerful mechanism for rapid genomic changes that can lead to reproductive isolation and speciation, and our data supports this premise.

We propose that chromosomal rearrangements triggered by telomere dysfunction might result in species extinction or a formation of new species, which includes either stabilization of the existing telomeric system or development of a new one. Our study showed that numerous Coleoptera species lack the ancestral insect telomere sequence, but the exact telomeric motif in most of these species remains unknown. We can only speculate whether the sequence was replaced by another short telomeric sequence or a completely different system independent of telomerase. To adopt new features, telomerase system can be very plastic as documented by the remarkable divergence in telomeric sequences in budding yeast, which shows extraordinary lengths, occasional degeneration, and a frequent absence of G/C-richness⁶⁴, or it can be documented by numerous species with backup pathways for telomere lengthening when telomerase activity is compromised^{65–68}.

Various studies identified transposable elements as modifiers of adaptive response upon exposure to a stressful environment. It has been shown that transposable elements can be activated by diverse stressors such as DNA damaging agents, thermal stress, and also telomere dysfunction^{69–72}. Transposable elements are known to induce genomic rearrangements⁷³, and we assume that the activation of transposable elements by telomere dysfunction not only contributes to speciation, but at the same time, it allows the development of the retroelement-based telomeric system during the speciation process. Although telomeric retroelements are a hallmark of *Drosophila* telomeres, they are found also in distantly related species incorporated in the telomerase-added sequences at their chromosome termini. It needs to be as well pointed out that telomeric retroelements are not universal systems of telomere elongation in the genus *Drosophila* and not at all in Diptera, as telomere maintenance in some species uses telomere-telomere recombination^{74–79}. Nevertheless, we can suppose that the retroelement and recombination systems work as the backup pathways for telomere lengthening when telomerase activity is compromised^{65–68}.

Together, these findings reveal the plasticity of chromosome ends for incorporating new features to maintain telomere integrity and functionality, perhaps pointing to the mechanisms by which telomeres contribute to the speciation and adaptation process. We believe that telomere diversity in insects provides the right opportunity to research such underexplored aspects of telomere biology.

Material and methods

Taxon sampling. In this study, we sampled 175 coleopteran species representing 14 superfamilies, 46 families, and 84 subfamilies (Table S1).

Hybridization probes. Hybridization probes were prepared using non-template PCR. The list of primers is provided in Table S2. The non-template PCR reaction contained 10 μ M forward primer, 10 μ M reverse primer, 10 mM dNTP mix, and Taq polymerase (5 U/ μ l). The PCR products were labeled by random primed labeling with digoxigenin-11-dUTP using a DIG DNA Labeling Kit (Roche diagnostics) and biotin-14-dATP using a Biotin-Nick Translation Mix (Roche diagnostics).

Southern hybridization. Genomic DNAs were isolated by a standard procedure of phenol–chloroform–isoamyl alcohol extraction. From each species, about 1 μ g DNA was digested with a mixture of RsaI/HinfI restriction endonucleases (New England BioLabs) (1:1), separated on 1% agarose gel, and blotted onto a Hybond-N + nylon membrane (Amersham Biosciences) by capillary transfer in 20 \times SSC overnight, followed by Southern or dot-blot hybridization with chemiluminescence detection as described previously⁸⁰. Briefly, the hybridization was performed for 3 h (at 50 °C) in 5 ml hybridization solution (1 M NaCl, 10 mM Tris–HCl pH 7.6, 1 mM EDTA, 40% formamide, 0.5% SDS, 5 \times Denhardt's solution) containing DIG-labeled probe (10 ng/ml), followed by washing in 2 \times SSC, 0.1% SDS (for 10 min, at room temperature) and 0.2 \times SSC, 0.1% SDS (for 20 min, at 68 °C). The chemiluminescence detection was performed using the alkaline phosphatase–CDP-Star system (Sigma–Aldrich), and the signals were detected using the CCD camera LAS-3000.

Bal 31 nuclease assay. To confirm the terminal positions of tested sequences on chromosome ends, the genomic DNA was subjected to BAL 31 exonuclease. Bal 31 nuclease degrades 3' and 5' termini of duplex DNA without generating internal scissions in the intact double helix. The assay was performed as described previously⁸¹. Briefly, the genomic DNA (15 μ g) was incubated with 0.03U BAL 31 nuclease (New England BioLabs) in a total volume of 180 μ l at 30 °C, and 60- μ l aliquots were taken before the BAL31 addition and after 30, 60, 90, and 120 min of the treatment. The reaction was immediately stopped by incubation at 65 °C for 10 min in the presence of 20 mM EGTA. Samples were purified using phenol–chloroform–isoamyl alcohol extraction, and DNA was extracted using the standard ethanol precipitation by adding 1/10th volume of sodium acetate and 2 volumes of ethanol. Then, the DNA samples (1 μ g) were digested with the restriction enzymes RsaI and Hinf I (New England Biolabs) and subjected to Southern hybridization.

Chromosomes preparations and FISH. Chromosome preparations were prepared from gonads of tested adults. The gonads were dissected in the Ringer's solution, incubated in a hypotonic solution (0.075 M KCl) for 10–20 min, and then fixed in freshly prepared Carnoy solution (ethanol:chloroform:acetic acid, 6:3:1). Using tungsten needles, tissue was ripped up in a drop of 60% acetic acid and spread on the microscope slide placed on a heating plate (45 °C). Preparations were dehydrated in an ethanol series (70%, 80%, 90%, 30 s each) and stored at –20 °C before their use.

Fluorescence in situ hybridization (FISH) with biotinylated probes was performed overnight at 37 °C in a solution of 60% formamide, 2 \times SSC, 10% dextran sulfate, 20 mM sodium phosphate, and 10 ng/ μ l of the DNA probe. Post-hybridization washes were done in 50% formamide/2 \times SSC at 37 °C. Biotin-labeled probes were detected with streptavidin–Cy3 (Jackson ImmunoResearch Laboratories) and biotinylated anti-streptavidin (Jackson ImmunoResearch Laboratories). The chromosomes were counterstained with DAPI (0.1 μ g/ μ l) mounted in 20 μ l of an anti-fade medium (25 mg/ml of 1,4-Diazobicyclo-(2,2,2) octane in 90% glycerol/10% 1 \times PBS) and analyzed with a Zeiss Axioplan 2 microscope. Images were taken separately for each fluorescent dye with a CCD camera (F-view, Soft Imaging System GmbH) in black and white resolution and subsequently stained and merged in Adobe Photoshop CS4.

Search NCBI databases. Short read archive at National Center for Biotechnology Information (SRA, NCBI) was searched for datasets from Coleoptera with restrictions to the WGS strategy, genomic source, and random selection. If available, only the first ten thousand spots were downloaded at maximum and analyzed using Tandem Repeats Finder with the options set as described previously for Bal31-NGS⁸². The presence/absence of telomeric motifs was checked manually in the Tandem Repeats Finder output.

Statistical analysis. The phylogenetic generalized least squares (GLS) from the ape package⁸³ was used to test the hypothesis that more speciose taxa have greater telomere diversity. GLS was used because the measurements were not independent (as they may have a common evolutionary history)⁸⁴ and the relationship was not linear (to be tested by correlation). We constructed a truncated phylogenetic tree using the most recent phylogenetic hypothesis of Coleoptera. As branch distances were not reported, these were computed using Grafen's method⁸⁵. Shannon (entropy) index was used to estimate the motif diversity and this response variable was then logarithmically transformed to fit an exponential relationship. Species richness was estimated at the superfamily level. The linear predictor included the number of investigated species (per taxon) to correct for different intensi-

ties of measurements and the age of the taxa. Brownian motion model of trait evolution was used in the analysis. The analysis was performed in the R environment (<http://www.r-project.org/index.html>) using R version 4.1.0²³.

Image processing. All images were processed with Adobe Photoshop CS6 (Adobe Systems; V.6.0.1), using proportionate adjustments of brightness or contrast.

Received: 10 March 2021; Accepted: 14 June 2021

Published online: 25 June 2021

References

- Blackburn, E.H. Structure and function of telomeres. *Nature*. **350**, 569–573 (1991).
- Cong, Y., Wright, W. E. & Shay, J. W. Human telomerase and its regulation. *Microbiol. Mol. Biol. Rev.* **66**, 407–425 (2002).
- Wright, W.E., Piatyszek, M.A., Rainey, W.E., Byrd, W., & Shay, J.W. Telomerase activity in human germline and embryonic tissues and cells. *Dev. Genet.* **18**(2), 173–179 (1996).
- Chan, S. W. & Blackburn, E. H. New ways not to make ends meet: Telomerase, DNA damage proteins and heterochromatin. *Oncogene* **21**, 553–563 (2002).
- Giannone, R. J. *et al.* The protein network surrounding the human telomere repeat binding factors TRF1, TRF2, and POT1. *PLoS ONE* **5**, e12407 (2010).
- Fulcher, N., Derboven, E., Valuchova, S. & Riha, K. If the cap fits, wear it: An overview of telomeric structures over evolution. *Cell. Mol. Life Sci.* **71**, 847–865 (2014).
- Mason, J. M., Randall, T. A. & Frydrychova, R. C. Telomerase lost?. *Chromosoma* **125**, 65–73 (2016).
- Traut, W. *et al.* The telomere repeat motif of basal Metazoa. *Chromosome Res.* **15**, 371–382 (2007).
- McEachern, M. J. & Blackburn, E. H. A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. *Proc. Natl. Acad. Sci. USA* **91**, 3453–3457 (1994).
- Peska, V. & Garcia, S. Origin, diversity, and evolution of telomere sequences in plants. *Front. Plant Sci.* **11**, 1–9 (2020).
- Korandová, M., Krůček, T., Vrbová, K. & Frydrychová, R. C. Distribution of TTAGG-specific telomerase activity in insects. *Chromosome Res.* **22**, 495–503 (2014).
- Frydrychová, R., Grossmann, P., Trubac, P., Vítková, M. & Marec, F. Phylogenetic distribution of TTAGG telomeric repeats in insects. *Genome* **47**, 163–178 (2004).
- Mason, J. M. & Biessmann, H. The unusual telomeres of *Drosophila*. *Trends Genet.* **11**, 58–62 (1995).
- Pardue, M.-L. & DeBaryshe, P. G. Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu. Rev. Genet.* **37**, 485–511 (2003).
- Nielsen, L. & Edström, J. E. 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 (1993).
- Kuznetsova, V., Grozeva, S. & Gokhman, V. Telomere structure in insects: A review. *J. Zool. Syst. Evol. Res.* **58**, 127–158 (2020).
- Mravinac, B., Meštrović, N., Cavrak, V. V. & Plohl, M. TCAGG, an alternative telomeric sequence in insects. *Chromosoma* **120**, 367–376 (2011).
- Stork, N. E. How many species of insects and other terrestrial arthropods are there on Earth?. *Annu. Rev.* **63**, 31–45 (2018).
- Okazaki, S., Tsuchida, K., Mackawa H., Fugiwara H. Identification of a pentanucleotide telomere sequence (TTAGG)_n in the silkworm *Bombyx mori* and in other insects. *Mol Cell Biol.* **13**, 1424–1432 (1993).
- Frydrychova, R. & Marec, F. Repeated losses of TTAGG telomere repeats in evolution of beetles (Coleoptera). *Genetica* **115**, 179–187 (2002).
- McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-beza, C., & Clarke, D.J. The evolution and genomic basis of beetle diversity. *PNAS* **116**, 24729–24737 (2019).
- Jordal, B. H., Smith, S. M. & Cognato, A. I. Classification of weevils as a data-driven science: Leaving opinion behind. *Zookeys* **18**, 1–18 (2014).
- Core Team, R. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).
- Beutel, R. G. Phylogenetic analysis of Adephaga (Coleoptera) based on characters of the larval head. *Syst. Entomol.* **18**, 127–147 (1993).
- Maddison, D. R., Baker, M. D. & Ober, K. A. Phylogeny of carabid beetles as inferred from 18S ribosomal DNA (Coleoptera: Carabidae). *Syst. Entomol.* **24**, 103–138 (1999).
- Hunt, T. *et al.* A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science (80-)*. **318**, 1913–1916 (2007).
- Friedrich, F. & Beutel, R. G. The pterothoracic skeletomuscular system of Scirtoidea (Coleoptera: Polyphaga) and its implications for the high-level phylogeny of beetles. *JZS* **44**, 290–315 (2006).
- Hunt, T. *et al.* Beetles reveals the evolutionary origins of a superradiation. *Science* **438**, 1913–1917 (2007).
- Newton, A.F. Agrytidae, Leiodidae. in *Handbook of Zoology* (ed. Beutel R. G., Leschen R. A. B.) 356–376 (De Gruyter 2016).
- Robertson, J. A. *et al.* Phylogeny and classification of Cucujoidea and the recognition of a new superfamily Coccinelloidea (Coleoptera: Cucujiformia). *Syst. Entomol.* **40**, 745–778 (2015).
- McKenna, D. D. *et al.* The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the. *Syst. Entomol.* **40**, 835–880 (2015).
- Haddad, S. & McKenna, D. Phylogeny and evolution of the superfamily Chrysomeloidea (Coleoptera: Cucujiformia). *Syst. Entomol.* **41**, 697–716 (2016).
- Sahara, K., Marec, F. & Traut, W. TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosom. Res.* **7**, 449–460 (1999).
- Klobutcher, L. A. & Swanton, M. T. All gene-sized DNA molecules in four species of hypotrachs have the same terminal sequence and an unusual 3' terminus. *PNAS* **78**, 3015–3019 (1981).
- Blackburn, E. H. & Gall, J. G. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *J. Mol. Biol.* **120**, 33–53 (1978).
- Moyzis, R. K. *et al.* A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. USA* **85**, 6622–6626 (1988).
- Richards, E. J. Isolation of a higher eukaryotic from *Arabidopsis thaliana*. *Cell* **53**, 127–136 (1988).
- Dekeirsschieter, J., Verheggen, F., Lognay, G. & Haubruge, E. Large carrion beetles (Coleoptera, Silphidae) in Western Europe: A review. *Zookeys* **15**, 435–447 (2011).
- Scholl, J. P., Wiens, J. J. & Wiens, J. J. Diversification rates and species richness across the Tree of Life. *Proc. R. Soc. B* **283**, 20161334 (2016).

40. Wiens, J.J. The causes of species richness patterns across space, time, and clades and the role of 'ecological limits'. *Q. Rev. Quarte* **86**(2), 75–96 (2011).
41. Rabosky, D. L., Slater, G. J. & Alfaro, M. E. Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biol.* **10**(8), e1001381 (2012).
42. McPeck, M. & Brown, J. M. Clade age and not diversification rate explains species richness among animal taxa. *Am. Naturalist* **169**, E97–106 (2007).
43. Bousquet, Y. Catalogue of Geadephaga (Coleoptera, Adephaga) of America, north of Mexico. *Zookeys* **1722**, 1–1722 (2012).
44. López-lópez, A. & Vogler, A. P. The mitogenome phylogeny of Adephaga (Coleoptera). *Mol. Phylogenet. Evol.* **114**, 166–174 (2017).
45. Korte, A., Ribera, I., Beutel, R. G. & Bernhard, D. Interrelationships of Staphyliniform groups inferred from 18S and 28S rDNA sequences, with special emphasis on Hydrophiloidea (Coleoptera, Staphyliniformia). *J. Zool. Syst. Evol. Res.* **42**, 281–288 (2004).
46. Kukalová-Peck, J. & Lawrence, J. Evolution of the hind wing in Coleoptera. *Can. Entomol.* **125**, 181–258 (1993).
47. Caterino, M. S., Hunt, T. & Vogler, A. P. On the constitution and phylogeny of Staphyliniformia (Insecta: Coleoptera). *Mol. Phylogenet. Evol.* **34**, 655–672 (2005).
48. Zhang, S. *et al.* Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nat. Commun.* **9**, 9–25 (2018).
49. Smith, A., Andrew, B. T., David, C. & John, M. An overview of the classification and evolution of the major scarab beetle clades (Coleoptera: Scarabaeoidea) based on preliminary molecular analyses. *Coleopt. Soc. Monogr. Number* **5**, 35–46 (2006).
50. McKenna, D. D., Farrell, B. D., Caterino, M. S. & Al, E. Phylogeny and evolution of Staphyliniformia and Scarabaeiformia: Forest litter as a stepping stone for diversification of nonphytophagous beetles. *Systematic* **40**, 35–60 (2015).
51. Zhang, Z. Animal biodiversity: An introduction to higher-level classification and taxonomic richness. *Zootaxa* **3148**, 7–12 (2011).
52. Christenhusz, M. J. M. & Byng, J. W. The number of known plants species in the world and its annual increase. *Phytotaxa* **261**, 201–217 (2016).
53. Peška, V. *et al.* Characterisation of an unusual telomere motif (TTTTTATAGGG)_n in the plant *Cestrum elegans* (Solanaceae), a species with a large genome. *Plant J.* **82**, 644–654 (2015).
54. Fajkus, P. *et al.* Allium telomeres unmasked: The unusual telomeric sequence (CTCGGTTATGGG)_n is synthesized by telomerase. *Plant J.* **85**, 337–347 (2016).
55. Tran, T. D., Cao, H. X., Jovtchev, G., Neumann, P. & Novak, P. Centromere and telomere sequence alterations reflect the rapid genome evolution within the carnivorous plant genus *Genlisea*. *Plant J.* **84**, 1087–1099 (2015).
56. Peska, V. *et al.* Human-like telomeres in *Zostera marina* reveal a mode of transition from the plant to the human telomeric sequences. *Botany* **71**, 5786–5793 (2020).
57. Weiss, H. & Scherthan, H. *Aloe* spp. plants with vertebrate-like telomeric sequences. *Chromosom. Res.* **10**, 155–164 (2002).
58. Leitch, A.R., & Fajkus, J. Asparagales telomerases which synthesize the human type of telomeres. *Plant. Mol. Biol.* **60**(5), 633–646 (2006).
59. Sýkorová, E., Leitch, A. R. & Fajkus, J. Asparagales telomerases which synthesize the human type of telomeres. *Plant Mol. Biol.* **60**, 633–646 (2006).
60. Chen, S., Kim, D., Chase, M. W. & Kim, J. Networks in a large-scale phylogenetic analysis: Reconstructing evolutionary history of Asparagales (Lilianaes) based on four plastid genes. *PLoS Genet.* **8**, 1–18 (2013).
61. Chase, M., Stevenson, D., Wilkin, P., & Rudall, P. Monocots systematics: A combined analysis. in *Monocotyledons: Systematics and Evolution* (ed. by Rudall P. J., Cribb P. J., Cutler D. F. & Humphries C. J.) 685–730 (Royal Botanic Gardens, 1995).
62. Wellinger, R. J. When the caps fall off: Responses to telomere uncapping in yeast. *FEBS Lett.* **584**, 3734–3740 (2010).
63. Stindl, R. Is telomere erosion a mechanism of species extinction?. *J. Exp. Zool. Part B Mol. Dev. Evol.* **302**, 111–120 (2004).
64. Steinberg-Neifach, O. & Lue, N. F. Telomere DNA recognition in *Saccharomycotina* yeast: Potential lessons for the co-evolution of ssDNA and dsDNA-binding proteins and their target sites. *Front. Genet.* **6**, 1–10 (2015).
65. Fujiwara, H., Osanai, M., Matsumoto, T., & Kojima, K.K. Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori*. *Chromosome Res.* **13**(5), 455–467 (2005).
66. Teng, S. C. & Zakian, V. A. Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 8083–8093 (1999).
67. Mikhailovsky, S., Belenkaya, T. & Georgiev, P. Broken chromosomal ends can be elongated by conversion in *Drosophila melanogaster*. *Chromosoma* **108**, 114–120 (1999).
68. Tacconi, E. M. C. & Tarsounas, M. How homologous recombination maintains telomere integrity. *Chromosoma* **124**, 119–130 (2015).
69. Mourier, T., Nielsen, L. P., Hansen, A. J. & Willerslev, E. Transposable elements in cancer as a by-product of stress-induced evolvability. *Front. Genet.* **5**, 1–8 (2014).
70. Korandová, M. *et al.* Chronic low-dose pro-oxidant treatment stimulates transcriptional activity of telomeric retroelements and increases telomere length in *Drosophila*. *J. Insect Physiol.* **104**, 1–8 (2018).
71. Scholes, D. T., Kenny, A. E., Gamache, E. R., Mou, Z. & Curcio, M. J. Activation of a LTR-retrotransposon by telomere erosion. *Proc. Natl. Acad. Sci. USA* **100**, 15736–15741 (2003).
72. McClintock, B. The significance of responses of the genome to challenge. *Science* **226**, 792–801 (1984).
73. Lonng, W.-E. & Saedler, H. Chromosome rearrangements and transposable elements. *Annu. Rev. Genet.* **36**, 389–410 (2002).
74. Nielsen, L. & Edstrom, J. E. 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 (1993).
75. Roth, C. W., Kobeski, F., Walter, M. F. & Biessmann, H. Chromosome end elongation by recombination in the mosquito *Anopheles gambiae*. *Mol. Cell. Biol.* **17**, 5176–5183 (1997).
76. Capkova Frydrychova, R., Biessmann, H. & Mason, J. M. Regulation of telomere length in *Drosophila*. *Cytogenet. Genome Res.* **122**, 356–364 (2009).
77. Mason, J. M., Reddy, H. M. & Capkova Frydrychova, R. Telomere maintenance in organisms without telomerase. In *DNA Replication-Current Advances* (ed. Seligman, H.) 323–346 (InTech, 2011).
78. Saint-Leandre, B., Nguyen, S. C. & Levine, M. T. Diversification and collapse of a telomere elongation mechanism. *Genome Res.* **29**, 920–931 (2019).
79. Compton, A. *et al.* The beginning of the end: A chromosomal assembly of the new world malaria mosquito ends with a novel telomere. *G3* **10**, 3811–3819 (2020).
80. Korandová, M. & Frydrychová, R. Č. Activity of telomerase and telomeric length in *Apis mellifera*. *Chromosoma* **125**, 405–411 (2016).
81. Koubová, J. *et al.* Telomerase activity is upregulated in the fat bodies of pre-diapause bumblebee queens (*Bombus terrestris*). *Insect Biochem. Mol. Biol.* **115**, 103241 (2019).
82. Peška, V., Sitová, Z., Fajkus, P. & Fajkus, J. BAL31-NGS approach for identification of telomeres de novo in large genomes. *Methods* **114**, 16–27 (2017).
83. Paradis, E. *Analysis of Phylogenetics and Evolution with R* (Springer, 2006).
84. Pekár, S. & Brabec, M. Marginal models via GLS: A convenient yet neglected tool for analysis of correlated data in behavioural sciences. *Ethology* **122**, 621–631 (2016).
85. Grafen, A. The phylogenetic regression. *Phil. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **326**, 119–157 (1989).

Acknowledgements

We thank Radek Aulický (the Crop Research Institute, Prague) for providing specimens from laboratory cultures. We express special gratitude to Dana Prušáková, Mikuláš Čapek and Matyáš Čapek, and Tomáš Stára for collecting numerous beetle specimens. Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA LM2018140) provided within the program Projects of Large Research, Development, and Innovations Infrastructures.

Author contributions

D.P.: performed all laboratory experiments; Specimen collection. V.P., M.B.: NCBI data search; Funding. S.P.: statistical analysis, revision of the text. L.Č., A.B.: Specimen collection and identification. R.C.F.: Conceptualization; Funding; manuscript writing, reviewing and editing; Supervision; Resources; Specimen collecting.

Funding

This work was supported by ERDF [project SYMBIT, reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477].

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-92705-y>.

Correspondence and requests for materials should be addressed to R.Č.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021