Chromosome Research **13**: 145–156, 2005.

The evolutionary origin of insect telomeric repeats, $(TTAGG)_n$

Magda Vítková[1,2], Jiří Král[3], Walther Traut[4], Jan Zrzavý[1,2] & František Marec[1,2]*
[1]*Faculty of Biological Sciences, University of South Bohemia, CZ-370 05 České Budějovice, Czech Republic;* [2]*Institute of Entomology ASCR, Branišovská 31, CZ-370 05 České Budějovice, Czech Republic; Tel: +420 387 775 250; Fax: +420 385 310 354; E-mail: marec@entu.cas.cz;* [3]*Faculty of Sciences, Charles University, CZ-128 44 Prague 2, Czech Republic;* [4]*Institut für Biologie, Universität zu Lübeck, D-23538 Lübeck, Germany*
**Correspondence*

Received 17 November 2004. Received in revised form and accepted for publication by Hans Lipps 13 December 2004

Key words: Arthropoda, chromosomes, Onychophora, phylogeny, Tardigrada, telomere

Abstract

The $(TTAGG)_n$ sequence is supposed to be an ancestral DNA motif of telomeres in insects. Here we examined the occurrence of TTAGG telomeric repeats in other arthropods and their close relatives by Southern hybridization of genomic DNAs and fluorescence *in-situ* hybridization (FISH) of chromosomes with $(TTAGG)_n$ probes or, alternatively, with the 'vertebrate' telomeric probe, $(TTAGGG)_n$. Our results show that the $(TTAGG)_n$ motif is conserved in entognathous hexapods (Diplura and Collembola), crustaceans (Malacostraca, Branchiura, Pentastomida, and Branchiopoda), myriapods (Diplopoda and Chilopoda), pycnogonids, and most chelicerates (Palpigradi, Amblypygi, Acari, Opiliones, Scorpiones, Pseudoscorpiones, and Solifugae) but not in spiders (Araneae). The presence of TTAGG repeats in these groups suggests that the sequence is an ancestral motif of telomeres not only in insects but in Arthropoda. We failed, however, to detect the TTAGG repeats in close relatives of the arthropods, Tardigrada and Onychophora. But while Onychophora had the 'vertebrate' $(TTAGGG)_n$ motif instead, the Tardigrada did not. The $(TTAGG)_n$ motif probably evolved from the $(TTAGGG)_n$ motif. Based on our and compiled data, we presume that the 'vertebrate' motif $(TTAGGG)_n$ is an ancestral motif of telomeres in bilaterian animals and possibly also in the superclade including animals, fungi and amoebozoans.

Introduction

A special DNA–protein complex, the telomere, is a requirement of the eukaryotic linear chromosomes to protect their ends against erosion by exonucleases, against the 5′ end gap resulting from incomplete replication, and against stickiness (Blackburn 1991). Besides that, it mediates attachment of chromosomes to the nuclear envelope at least during some stages of meiosis and has several other vital functions (Zakian 1995). In most eukaryotes, telomeric DNA is composed of lengthy stretches of a simple sequence repeat with a consensus of $(T_nG_n)_n$ at the 3′ end. These stretches are maintained at the DNA ends by telomerase, a special ribonucleoprotein enzyme that adds the simple repeats to the telomere by reverse transcription (Greider 1996). The G-rich 3' end, forming an overhang, provides the basis for formation of non-Watson-Crick structures

such as G-quartets and t-loops (Krupp *et al.* 2000, de Lange 2004).

The telomerase-mediated elongation of chromosome ends is highly conserved across eukaryotes, from unicellular organisms to flowering plants and vertebrates (Krupp *et al.* 2000). However, the absence of short telomeric repeats in some flowering plants (Pich *et al.* 1996, Sykorova *et al.* 2003b) and several insect orders (Sahara *et al.* 1999, Frydrychová & Marec 2002, Frydrychová *et al.* 2004) suggest that some organisms can avoid that scheme and provide compensation for erosion of telomeric DNA by some other means. This has been demonstrated in the insect order Diptera, where two diverse mechanisms of telomere maintenance have evolved, namely, insertion of mobile elements and unequal recombination between long tandem repeats (Biessmann & Mason 2003, Pardue & DeBaryshe 2003).

Within some large phylogenetic groups, the conservation of telomeric repeats is rather precise. Thus, all vertebrates have a hexameric repeat, $(TTAGGG)_n$ (Meyne *et al.* 1989), and telomeres of most higher plants are composed of a heptameric motif, $(TTTAGGG)_n$ (Fajkus & Zentgraf 2002). In insects, a pentameric sequence, $(TTAGG)_n$, is a widespread, though not the only, telomere motif (Okazaki *et al.* 1993, Sahara *et al.* 1999). A comprehensive study by Frydrychová *et al.* (2004) revealed that the $(TTAGG)_n$ motif is conserved in the main lineages of insects with only occasional exceptions outside Endopterygota. The exceptions are Ephemeroptera, Odonata, Dermaptera, and Heteroptera. In the Endopterygota, the most successful and diversified lineage of insects, there is apparent heterogeneity in the TTAGG occurrence. The motif is conserved in Hymenoptera, Lepidoptera, Trichoptera, and Megaloptera, but absent in the clade Antliophora (Diptera + Siphonaptera + Mecoptera) and in Raphidioptera. In addition, the Coleoptera and Neuroptera show heterogeneity with respect to the TTAGG presence even within the orders. Nevertheless, the overall pattern of phylogenetic distribution of the TTAGG repeats clearly supports the hypothesis that the sequence is an ancestral motif of telomeres in insects but was lost repeatedly during evolution, being replaced with another motif or an alternative mechanism of telomere elongation.

In the present paper, we set out to establish the phylogenetic distribution of the $(TTAGG)_n$ telomere sequence in non-insect arthropods and close relatives of the arthropods, with the aim to map its evolution on the present-day phylogenetic tree of that part of the animal kingdom. For this purpose, genomic DNAs of selected species were subjected to Southern hybridization with telomeric probes, and data obtained were confirmed by fluorescence *in-situ* hybridization (FISH) on chromosome preparations.

Material and methods

Specimens

Species examined along with the origin of specimens are given in Table 1. We used 23 species representing Tardigrada (two species), Onychophora (one species), and five main clades of Arthropoda (two species of entognathous Hexapoda, five of Crustacea, two of Myriapoda, ten of Chelicerata, and one of Pycnogonida).

Telomere probes

We prepared two types of telomeric DNA probes, 'insect' type $(TTAGG)_n$ and 'vertebrate' type $(TTAGGG)_n$. Unlabelled probes were generated by the non-template polymerase chain reaction (PCR) method, following the procedure described in Sahara *et al.* (1999). For PCR, we used twenty-mer primers custom-made by Generi Biotech (Hradec Králové, Czech Republic) and *Taq* DNA polymerase of GibcoBRL (Life Technologies Inc., Karlsruhe, Germany). For non-radioactive Southern hybridization, probes were labelled by random-primed labelling with digoxigenin 11-dUTP using the DIG DNA Labeling and Detection Kit (Roche Diagnostics GmbH, Mannheim, Germany). For FISH, probes were labelled by nick translation with biotin-14-dATP using the BioNick Labeling System (GibcoBRL).

DNA isolation

Total genomic DNA of each species was extracted using three different techniques, depending on the body/tissue size. DNAs of tiny species

Table 1. List of species studied, higher taxonomic group, source of specimens, karyotype when known, and results of Southern hybridization and FISH with the 'insect' telomeric probe, $(TTAGG)_n$, or 'vertebrate' telomeric probes, $(TTAGGG)_n$ or $(CCCTAA)_3$ PNA.

				*Southern hybridization**		*FISH**	
Higher taxon	*Species*	*Source and/or locality (donor)*	*Karyotype*	$(TTAGG)_n$	$(TTAGGG)_n$	$(TTAGG)_n$	$(TTAGGG)_n$
Hexapoda							
Diplura	*Campodea* sp.	Davle Village, Prague, Czech Rep.		pos.		pos.	
Collembola	*Tetrodontophora bielanensis*	Žilina Village, Beskydy Mts., Czech Rep.		pos.		pos.	
Crustacea							
Malacostraca	*Trachelipus ratzeburgii*	České Budějovice, Czech Rep.		pos.		pos.	
Branchiopoda	*Daphnia galeata*	lab. cultures (R. Pichlová and L. Stará)	$2n=20$[a]	pos.		pos.	
	Artemia franciscana	commercial cyst lot, Great Salt Lake, Utah	$2n=42$[b]	pos.		pos.	
Branchiura	*Argulus foliaceus*	Slapy Dam, Czech Rep. (J. Matěna)		pos.		pos.	
Pentastomida	*Raillietiella* sp.	from geckos *Hemidactylus frenatus* (W. Böckeler)		—		pos.	
Myriapoda							
Diplopoda	*Julus scandinavius*	České Budějovice, Czech Rep.		pos.		pos.	
Chilopoda	*Lithobius forficatus*	České Budějovice, Czech Rep.	$2n=46$[c]	pos.		pos.	
Chelicerata							
Opiliones	*Phalangium opilio*	České Budějovice, Czech Rep.		pos.		—	
	Leiobunum limbatum	Vyšehrad Hill in Prague, Czech Rep.	♂$2n=22$[d]	—		pos.	
Scorpiones	*Heterometrus spinifer*	lab. cultures (F. Kovařík)	$2n=56$[d]	pos.		pos.	
Pseudoscorpiones	*Chernes hahnii*	Prague, Czech Rep. (F. Šťáhlavský)	♂$2n=49$[e]	pos.		pos.	
Solifugae	*Rhagodes* sp.	Ashalim Village, Sde Boker, Israel (S.Pekár)	$2n=18$[d]	pos.		pos.	
Acari	*Ixodes ricinus*	lab. cultures (J. Erhart)	♂$2n=28$[f]	pos.		pos.	
	Tetranychus urticae	lab. cultures (R. Zemek)	♀$2n=6$[g]	pos.		pos.	
Palpigradi	*Eukoenenia spelaea*	Ardovská cave, Slovak Karst, Slovakia (L. Kováč and P. Ľuptáčik)		—		pos.	
Amblypygi	*Damon medius*	lab. cultures (P. Weygoldt)	♂$2n=70$[d]	pos.		pos.	
Araneae	*Brachypelma albopilosa*	lab. cultures (B. Hlivka)	♂$2n=74$[d]	neg.	neg.	neg.	neg.
Pycnogonida	*Ammothella biunguiculata*	sea intertidal near Seto, Japan (K. Myiazaki)		pos.		pos.	
Tardigrada	*Macrobiotus hufelandi*	Dvůr Králové, Czech Rep. (K.Voženílková)		neg.	neg.	inc.	neg.
	Milnesium tardigradum			neg.	neg.	inc.	neg.
Onychophora	*Peripatopsis stelliporata*[h]	Cape Town, South Africa (M.H. Walker)		neg.	pos.	neg.	pos.

* pos./neg./inc., positive/negative/inconclusive results of Southern hybridization or FISH; [a] Beaton & Hebert (1994); [b] Colihueque & Gajardo (1996); [c] Woznicki *et al.* (2003); [d] this study; [e] Šťáhlavský (2000); [f] Kahn (1964); [g] Helle & Bolland (1967); [h] Sherborn & Walker (2004).

with amounts of tissue lower than 10 mg were extracted using the AquaPure Genomic DNA Kit (Bio-Rad Laboratories, Hercules, California). In species with a body weight of 10–25 mg, we used the DNeasy Tissue Kit (Qiagen GmbH, Hilden, Germany). Finally, standard phenol–chloroform–isoamylalcohol extraction was used in relatively large species with a minimum amount of 25 mg of tissues.

Southern hybridization

From each species, about 600 ng of genomic DNA was digested with a Hinf I and Rsa I restriction endonuclease mixture (1 : 1) overnight and separated on 1% agarose gel. After electrophoresis, the gel was weakly stained with ethidium bromide and the quality of DNA digestion and loading checked under UV light. The gel was blotted onto a positively charged nylon membrane (Hybond-N^+; Amersham Biosciences UK Ltd., Buckinghamshire, England) for 16–20 h. Hybridization of DNAs with DIG-labelled telomere probe and subsequent chemiluminescent detection were performed using the Telo TAGGG Telomere Length Assay (Roche Diagnostic GmbH) according to the protocol of the manufacturer with several modifications described in Frydrychová & Marec (2002).

Chromosome preparations

Spread chromosome preparations were made from testes of subadult or adult males except when especially mentioned. In tiny species (the dipluran, two-spotted spider mite, water bears, and palpigrade), the entire body was used for spreading. In the fresh-water flea, mitotic chromosomes were prepared from embryos carried by females. Cysts of the brine shrimp were cultivated in salt water and, after 24 h (i.e. shortly before hatching), used for preparation of mitotic chromosomes. In the scorpion and solifuge, we prepared mitotic chromosomes from the abdominal digestive glands of larvae and post-embryos, respectively. Finally, polytene chromosomes were obtained from an unidentified tissue of the springtail.

The spreading procedure was essentially the same as given in Frydrychová & Marec (2002). Briefly, dissected tissues (or entire bodies) were swollen for 10–15 min in a hypotonic solution (0.075 mol/L KCl) and fixed in Carnoy's fluid (ethanol, chloroform, acetic acid, 6 : 3 : 1) or in methanol–acetic acid (3 : 1). In the palpigrade, hypotonic treatment was omitted because it led to decay of tissues. After fixation, tissues were dissociated in 60% acetic acid and spread on the slide using a heating plate at 45°C. The preparations were dehydrated in an ethanol series (70%, 80%, 100%, 30 s each) and stored at −20°C until further use.

Fluorescence in-situ *hybridization (FISH)*

We used the procedure described in Sahara *et al.* (1999). Hybridization signals were detected with Cy3-conjugated streptavidin (Jackson ImmunoRes. Labs. Inc., West Grove, PA) and one round of amplification with biotinylated anti-streptavidin (Vector Labs. Inc., Burlingame, CA) and Cy3-conjugated streptavidin. The preparations were counterstained with DAPI (Sigma-Aldrich, Inc., St. Louis, Missouri) at a concentration of 0.5 μg/ml for pachytene bivalents or 0.05 μg/ml for metaphase chromosomes, and mounted in antifade based on DABCO (Sigma-Aldrich).

The preparations were inspected in a Zeiss Axioplan 2 microscope. Black-and-white images were recorded with a CCD camera (F-view, Soft Imaging System GmbH, Münster, Germany) separately for each fluorescent dye. Images were pseudocoloured (light blue for DAPI and red for Cy3) and superimposed with the aid of Adobe Photoshop Version 5.0.

Chromosomes of species that did not show hybridization signals with $(TTAGG)_n$ were probed for the presence of the $(TTAGGG)_n$ telomere motif. For this, we used either standard FISH procedure (see above) or the Telomere PNA FISH Kit/Cy3 (Dako, Glostrup, Denmark), according to the protocol of the manufacturer.

Results

Southern hybridization

The $(TTAGG)_n$ probe was hybridized to digested genomic DNAs of 20 species (19 shown in

Figure 1a). For each species, the experiments were repeated two to four times, always with essentially identical results. Two species were not probed by Southern hybridization, the tongue worm *Raillietiella* sp. because of a low amount of extracted DNA and the palpigrade *Eukoenenia spelaea* because the entire bodies of all specimens obtained were used for preparation of chromosomes. In Opiliones, *Phalangium opilio* was used for Southern hybridization and the other species, *Leiobunum limbatum*, for FISH.

In four species, no hybridization signals were observed indicating the absence of TTAGG repeats in their genomes (Figure 1a). These were three arthropod-related species, *Peripatopsis stelliporata* (Onychophora), *Macrobiotus hufelandi* (not shown) and *Milnesium tardigradum* (both Tardigrada), and *Brachypelma albopilosa* (Araneae). DNAs of all other species hybridized with the probe. In most TTAGG-positive species, hybridization signals formed a bold band in the high-molecular-weight region with a large fraction longer than 21 kb, and a diffuse smear in the molecular weight range of 5–21 kb. A diffuse smear in the range of about 3–15 kb without a particular band was found in *Daphnia galeata* (Branchiopoda) and *Campodea* sp. (Diplura). *Ammothella biunguiculata* (Pycnogonida), *Heterometrus spinifer*, (Scorpiones), and *Chernes hahnii* (Pseudoscorpiones) displayed a wide range of smeary signals from less than 0.5 kb up to 21 kb, and *Chernes* a few additional bands in the low-molecular-weight range. Finally, a smear of signals with several bands only in the range below 3 kb was obtained in *Ixodes ricinus* (Parasitiformes). The shorter restriction fragments in the latter four species could result from short telomeres or from integrated non-telomeric DNA domains containing the respective restriction sites. Such domains were found in telomeres and subtelomeric regions of the silkworm *Bombyx mori* (e.g. Anzai *et al.* 2001). Some shorter restriction fragments might also originate from non-telomeric sites, the so-called

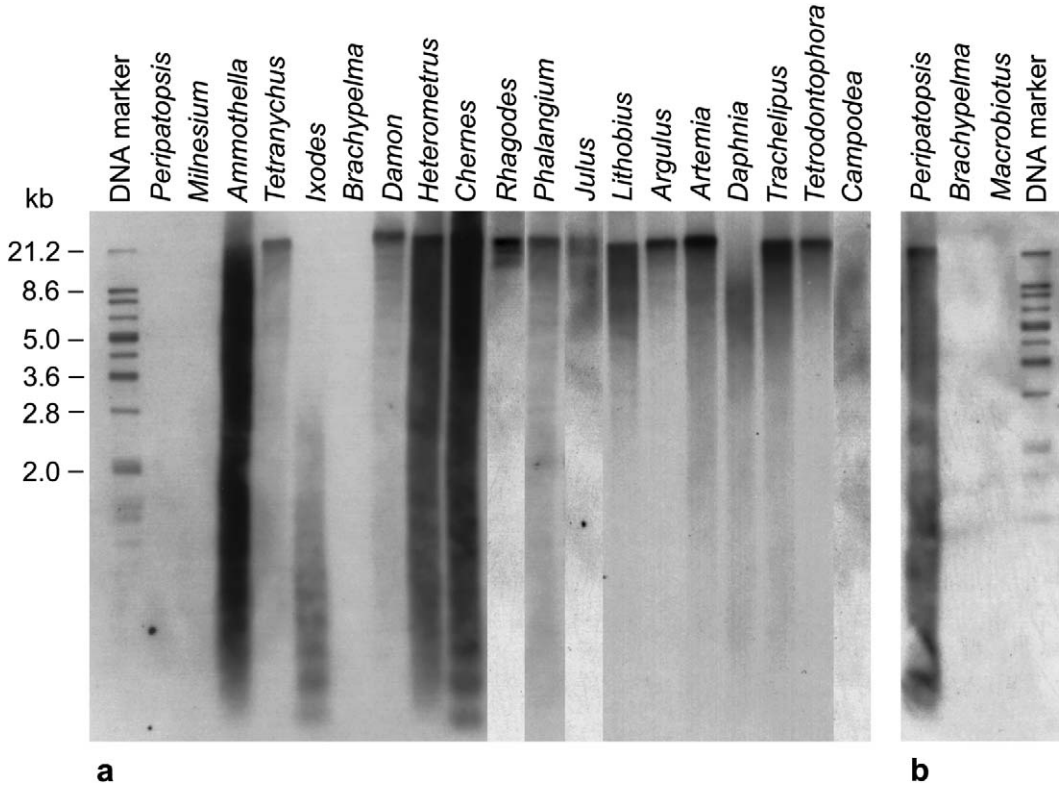
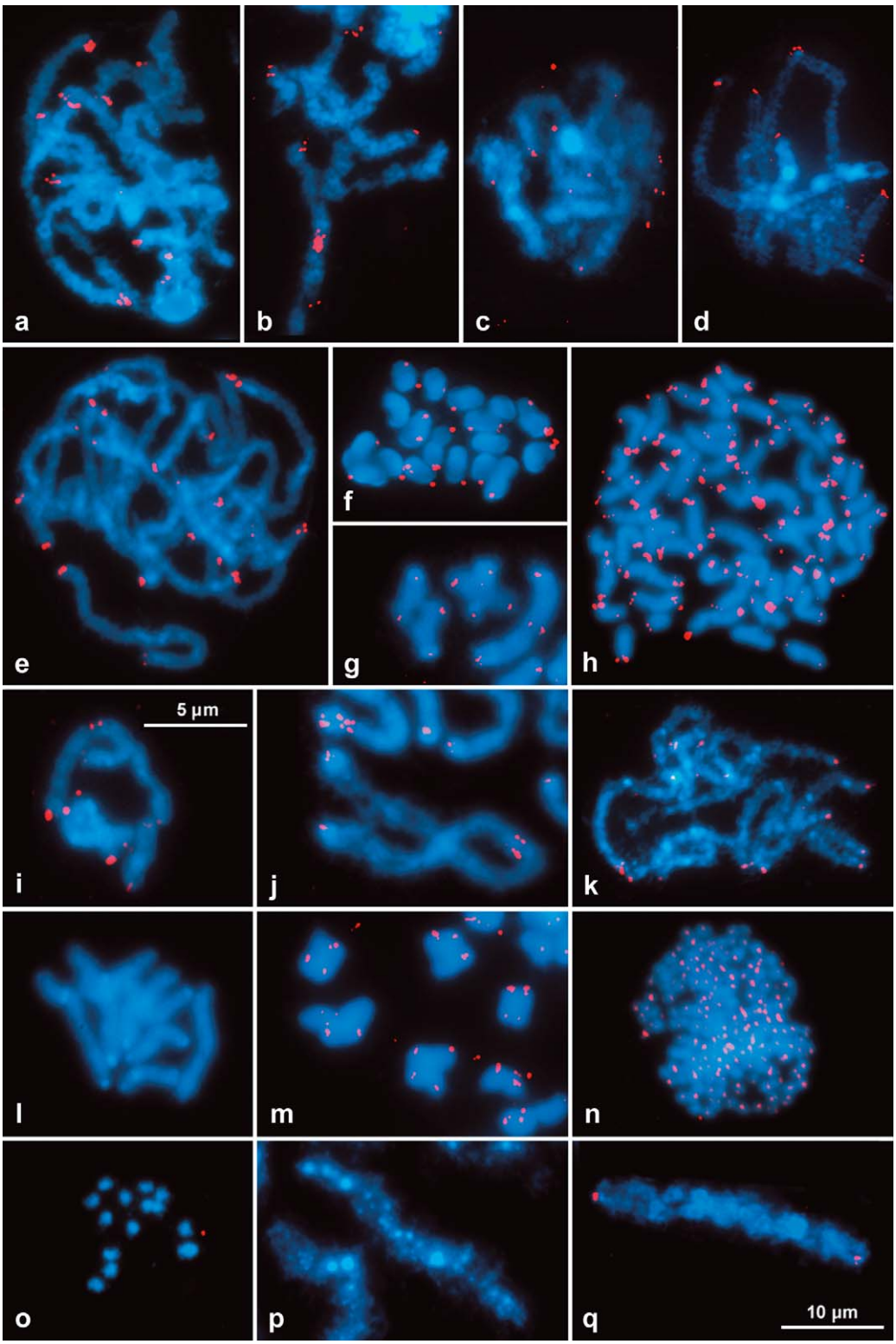


Figure 1. Southern hybridization of $(TTAGG)_n$ probe (**a**) or $(TTAGGG)_n$ probe (**b**) to Hinf I/Rsa I-digested genomic DNAs from 20 species of arthropods and related taxa. Full species names are given in Table 1.


a
b
c
d
e
f
g
h
5 µm
i
j
k
l
m
n
o
p
q
10 µm

interstitial telomeric-like sequences which were reported for $(TTAGG)_n$ repeats in the vapourer moth *Orgyia antiqua* (Rego & Marec 2003).

In the TTAGG-negative species, we performed Southern hybridization also with the $(TTAGGG)_n$ probe (Figure 1b). Two repeated experiments revealed the presence of the 'vertebrate-type' telomeric repeats in the genome of *Peripatopsis*, whereas no hybridization signals were observed in *Milnesium* (not shown), *Macrobiotus*, and *Brachypelma*.

Fluorescence in-situ *hybridization (FISH) with* $(TTAGG)_n$ *probe*

FISH with $(TTAGG)_n$ probe was made in a total of 22 species. In each species, at least two FISH preparations were examined. The experiments confirmed the results of Southern hybridization except in Tardigrada where they were inconclusive (see below). Southern hybridization and FISH data are summarized in Table 1. New data on chromosome numbers, determined in some species, along with published information on karyotypes are also included in Table 1.

Entognathous Hexapoda

FISH confirmed that the TTAGG sequence is a component of telomeres in both representatives. In *Campodea* sp. (Diplura), pachytene spermatocytes were observed with telomeric signals at the ends of bivalents (not shown). In the giant springtail *Tetrodontophora bielanensis* (Collembola), we obtained polytene chromosomes with a low level of polyploidy (Figure 2a).

Crustacea

$(TTAGG)_n$ probe hybridized to chromosome ends of five species examined. In the isopod *Trachelipus ratzeburgii* (Malacostraca; Figure 2b), the tongue worm *Raillietiella* sp. (Pentastomida; Figure 2c), and the fishlouse *Argulus foliaceus* (Branchiura; not shown), pachytene spermatocytes were examined. In the freshwater flea *Daphnia galeata* (Branchiopoda, Cladocera) and the brine shrimp *Artemia franciscana* (Branchiopoda, Anostraca), mitotic metaphase chromosomes with telomeric signals were seen (not shown).

Myriapoda

Representatives of both myriapod classes, Diplopoda and Chilopoda, showed clear hybridization signals. In the millipede *Julus scandinavius*, signals were observed in interphase nuclei and sperm (we did not find sufficiently early stages of spermatogenesis for chromosomes). In the centipede *Lithobius forficatus*, the telomere signals were checked and found at the ends of pachytene chromosomes (Figure 2d).

Chelicerata

Nine species were examined. Strong $(TTAGG)_n$ signals were observed at chromosome termini of eight species. These were the harvestman *Leiobunum limbatum* (Opiliones; Figure 2e), *Rhagodes* sp. (Solifugae; Figure 2f), *Chernes hahnii* (Pseudoscorpiones; Figure 2g), *Heterometrus spinifer* (Scorpiones; Figure 2h), the two-spotted spider mite *Tetranychus urticae* (Acariformes; Figure 2i), the tick *Ixodes*

Figure 2. FISH of $(TTAGG)_n$ (**a–n, p**) or $(CCCTAA)_3$ PNA (**o, q**) probes (red signals) to chromosome spreads (counterstained with DAPI). *Tetrodontophora bielanensis* (**a**): polytene chromosomes with clusters of strong hybridization signals at the ends of most chromosome arms. *Trachelipus ratzeburgii* male (**b**): pachytene bivalents with telomeric signals. *Raillietiella* sp. male (**c**): pachytene nucleus with telomeric signals on bivalent ends. *Lithobius forficatus* male (**d**): a part of pachytene nucleus with twin hybridization signals at the ends of bivalents. *Leiobunum limbatum* male (**e**): pachytene complement with twin hybridization signals at the ends of bivalents. *Rhagodes* sp. post-embryo (**f**): mitotic metaphase complement with telomeric signals at most but not all chromosome ends, perhaps as a result of a lower penetrance of the probe into compact chromatin. *Chernes hahnii* male (**g**): several diplotene bivalents with telomeric signals indicating the position of actual chromosome ends. *Heterometrus spinifer* larva (**h**): mitotic metaphase with two hybridization signals at most chromosome ends. *Tetranychus urticae* female larva (**i**): mitotic prometaphase ($2n = 6$) showing telomeric signals at each chromosome end. *Ixodes ricinus* male (**j**): a diplotene bivalent with strong telomeric signals. *Eukoenenia spelaea* male (**k**): pachytene complement with hybridizations signals at the end of bivalents. *Brachypelma albopilosa* male nymph (**l**): several mitotic prometaphase chromosomes without hybridization signals. *Damon medius* male (**m**): several metaphase I bivalents with telomeric signals. *Ammothella biunguiculata* male (**n**): mitotic metaphase showing tiny and numerous chromosomes with strong telomeric signals. *Milnesium tardigradum* (**o**): mitotic metaphase chromosomes without hybridization signals of the $(CCCTAA)_3$ PNA probe. *Peripatopsis stelliporata* male (**p, q**): (**p**) a pachytene bivalent without hybridization signals after FISH with the $(TTAGG)_n$ probe; (**q**) a pachytene bivalent with clear telomeric signals of the $(CCCTAA)_3$ PNA probe at both ends. The scale bar of 10 μm holds for all figures except (**i**).

ricinus (Parasitiformes; Figure 2j), *Eukoenenia spelaea* (Palpigradi; Figure 2k), and *Damon medius* (Amblypygi; Figure 2m). In contrast, no hybridization signals were found in chromosomes from male nymphs of the tarantula *Brachypelma albopilosa* (Figure 2l). Thus, FISH confirmed the absence of the $(TTAGG)_n$ motif in this representative of spiders (Araneae).

Pycnogonida

Preparations of testes from late nymphs and adult males in the sea spider, *Ammothella biunguiculata*, showed various stages of spermatogenesis from mitotic spermatogonia to metaphase I spermatocytes. Mitotic metaphases consisted of numerous and very tiny elements (we estimated a diploid chromosome number of 70–80) with numerous hybridization signals. Where single chromosomes lay separated from the more compact mass, the signals were clearly telomeric (Figure 2n).

Tardigrada

Most water bears possess low chromosome numbers of n = 5–6; however, these numbers vary even within species due to polyploidy (Rebecchi *et al*. 2002). The whole body preparations made from *Macrobiotus hufelandi* and *Milnesium tardigradum* showed mainly interphase nuclei or, in several specimens, elongated nuclei of sperm, but chromosome plates were only found exceptionally. The FISH procedure used resulted in heavy noise of fluorescent label on nuclei and in their vicinity (not shown). The unspecific signals, which were most probably caused by cross-reaction of antibodies with cytoplasmic and/or nuclear proteins, made evaluation of true hybridization signals impossible.

Onychophora

In accordance with the negative result of Southern hybridization, no hybridization signals were found on meiotic chromosomes of the velvet worm *Peripatopsis stelliporata* (Onychophora), prepared from testes of an adult male (Figure 2p).

FISH with $(TTAGGG)_n$ and $(CCCTAA)_3$ probes

No hybridization of the $(CCCTAA)_3$ PNA probe was observed in meiotic chromosomes of the tarantula *B. albopilosa* (Araneae; not shown). In *M. hufelandi* and *M. tardigradum* (Tardigrada), after standard FISH with biotin-labelled $(TTAGGG)_n$ probe, nuclei and chromosomes were covered with unspecific signals (not shown), similar to FISH with the 'insect' telomeric probe (see above). With the PNA probe, however, interphase nuclei and chromosomes remained free of any hybridization signals (Figure 2o). On the other hand, pachytene bivalents of *P. stelliporata* (Onychophora) displayed clear hybridization signals of the PNA probe with typical locations at telomeric regions (Figure 2q).

Discussion

Origin of the $(TTAGG)_n$ motif in arthropods

Previous studies suggested that the $(TTAGG)_n$ sequence was an ancestral motif of telomeres in insect chromosomes. It is present in most insect groups but has been lost and most probably replaced with alternative mechanisms of telomere maintenance in several insect orders (see Introduction). Sahara *et al*. (1999) showed that the $(TTAGG)_n$ sequence is a component of telomeres in the amphipod *Gammarus pulex* (Crustacea: Malacostraca). Their finding was in accordance with data obtained in two other malacostracans, namely positive Southern hybridization in the green tiger prawn *Penaeus semisulcatus* (Okazaki *et al*. 1993) and TTAGG-specific telomerase activity in the lobster *Homarus americanus* (Klapper *et al*. 1998). On the other hand, TTAGG repeats were not found in the spider *Tegenaria ferruginea* (Sahara *et al*. 1999). Taken together, the results appeared to indicate the origin of $(TTAGG)_n$ within the Arthropoda, possibly in a common ancestor of Crustacea plus Hexapoda, as the two taxa represent sister groups according to the recent phylogenetic concept of Pancrustacea (Giribet *et al*. 2001, Mallatt *et al*. 2004).

In this study we found that the $(TTAGG)_n$ telomere motif is present not only in Pancrustacea, but also in Myriapoda, Chelicerata, and Pycnogonida (Figure 3). In most species, Southern hybridization to Hinf I/Rsa I restriction fragments longer than 21 kb indicated considerable lengths of the telomeric DNA. The only TTAGG-negative species among the analysed non-insect arthropods

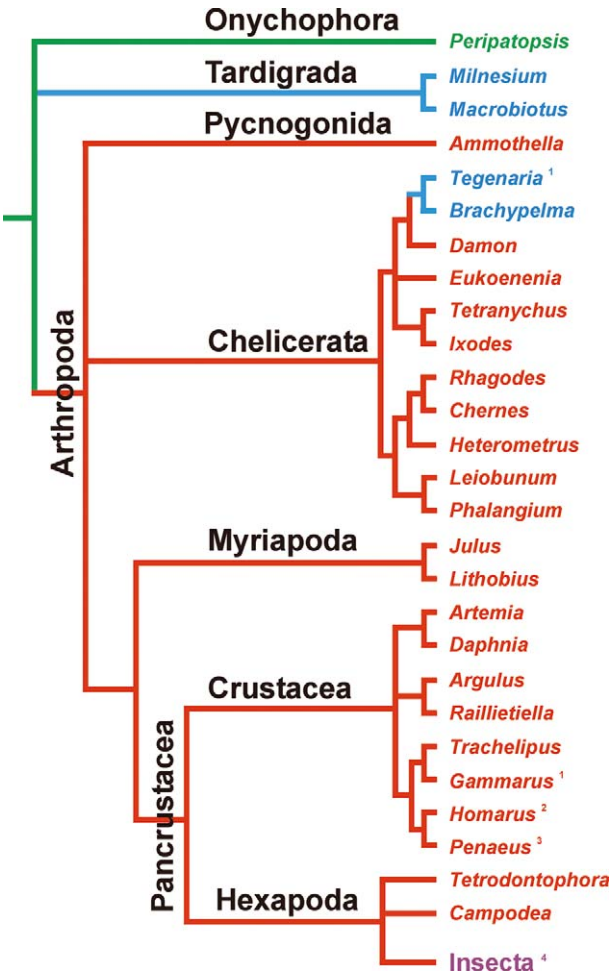


Figure 3. Phylogeny of arthropod and arthropod-related species examined for the presence of $(TTAGG)_n$ (red) or $(TTAGGG)_n$ (green) telomeric repeats or for the absence of both telomeric repeats (blue) by Southern hybridization and/or FISH (for details see text). In Insecta (violet), the $(TTAGG)_n$ is an ancestral motif of telomeric DNA but was repeatedly lost in several insect orders. The cladogram is based on Giribet *et al.* (2001) for Arthropoda and Giribet *et al.* (2002) for Chelicerata. [1]Sahara *et al.* (1999); [2]Klapper *et al.* (1998); [3]Okazaki *et al.* (1993); [4]Frydrychová *et al.* (2004).

was the tarantula *Brachypelma albopilosa*. Since *B. albopilosa* (this study) and *T. ferruginea* (Sahara *et al.* 1999) represent two of the three major clades of spiders (Araneae), the Mygalomorphae and Araneomorphae, respectively, it is possible that the absence of TTAGG telomeric repeats is a common trait for the Araneae. The amblypygid *Damon medius*, a member of the spider sister group (Giribet *et al.* 2002), has the $(TTAGG)_n$ telomere motif (this study); hence the motif has probably been lost between an ancestor of the amblypygids and spiders and that of the mygalomorph and araneomorph spiders. This could happen at the Paleozoic, since fossils of the spiders are known from the Triassic (Selden & Gall 1992, Selden *et al.* 1999). We also showed that the tarantula lacks the 'vertebrate-type' telomeric sequence, $(TTAGGG)_n$. Thus spiders may either have another short sequence repeat, or the telomerase-dependent system of telomere maintenance has been replaced with a yet unknown alternative mechanism.

The presence of TTAGG repeats in *Raillietiella* (Pentastomida) corroborates placement of the tongue worms within the Arthropoda, which is based on shared spermatology, ribosomal RNA (review: Zrzavý 2001), and mitochondrial genomes (Lavrov *et al.* 2004) but doubted by some paleontologists (Maas & Waloszek 2001).

The presence of TTAGG repeats in all main arthropod clades strongly suggests that the 'insect' telomeric sequence is, in fact, an ancestral motif of telomeres for arthropods. Since the $(TTAGG)_n$ motif was not found in the closest arthropod relatives (Onychophora and Tardigrada), its origin can be determined to the arthropod stem lineage at the Precambrian (probably Ediacaran) period at least 545 MYA (Fortey *et al.* 1997, Peterson *et al.* 2004).

Phylogeny of short telomeric repeats in animals and other eukaryotes

Two animal phyla seem to possess unique telomeric repeats, a TTAGGC hexamer in the Nematoda and the TTAGG pentamer in the Arthropoda. The $(TTAGGC)_n$ telomere motif was discovered in *Ascaris lumbricoides* (Muller *et al.* 1991), and later found in two other species, *A. suum* and *Parascaris univalens* (Niedermaier & Moritz 2000). The $(TTAGG)_n$ and $(TTAGGC)_n$ sequences represent evolutionary novelties of the Arthropoda and Nematoda, respectively, whereas the so-called 'vertebrate' motif $(TTAGGG)_n$ is evidently much older than the Vertebrata. The TTAGGG sequence was first identified in telomeres of human chromosomes (Moyzis *et al.* 1988) and then found to be conserved in all vertebrates (Meyne *et al.* 1989). At present, it is known to occur also in Cephalochordata (Castro & Holland

2002), Echinodermata (Plohl *et al.* 2002), Onychophora (this study), Platyhelminthes (Hirai & LoVerde 1996), Annelida (Vitturi *et al.* 2002, and references therein), and Mollusca (e.g. Wang & Guo 2001, Vitturi *et al.* 2004, and references therein). Thus, $(TTAGGG)_n$ appears to be a common telomere DNA motif in bilaterian animals. Based on its distribution in Bilateria, the 'vertebrate' telomeric motif has to be hypothesized as an ancestral motif of telomeres in the whole Bilateria (Figure 4). Hence, the other two telomeric repeats occurring in Bilateria, TTAGGC and TTAGG, evolved independently from the ancestral sequence TTAGGG.

So far, no data are available concerning telomeric sequences in non-Bilateria multicellular animals (like sponges and cnidarians). However, the occurrence of the $(TTAGGG)_n$ motif in other eukaryots suggests that ancestral metazoans could have it too. Although various telomeric repeats have been reported for the fungi (Zakian 1995, Underwood *et al.* 1996, Bhattacharyya & Blackburn 1997, Levis *et al.* 1997, Abdennadher & Mills 2000, Krupp *et al.* 2000), the $(TTAGGG)_n$ motif is the only telomeric sequence present in several representatives of both major fungal clades, the Ascomycota (*Neurospora*, *Podospora*, *Cladosporium*, *Aspergillus*, *Botrytis*, *Pneumocystis*) and Basidiomycota (*Ustilago*). As the fungi and animals are sister groups (Patterson 1999, Cavalier-Smith 2002, 2003), it seems even possible that the animal–fungal superclade as a whole was ancestrally characterized by the $(TTAGGG)_n$ motif. Moreover, the same motif was also found in two slime molds (*Physarum* and *Didymium*; Zakian 1995), representatives of the newly discovered eukaryote phylum, Amoebozoa, a probable sister group of the animal–fungal superclade (Cavalier-Smith 2003, Cavalier-Smith *et al.* 2004).

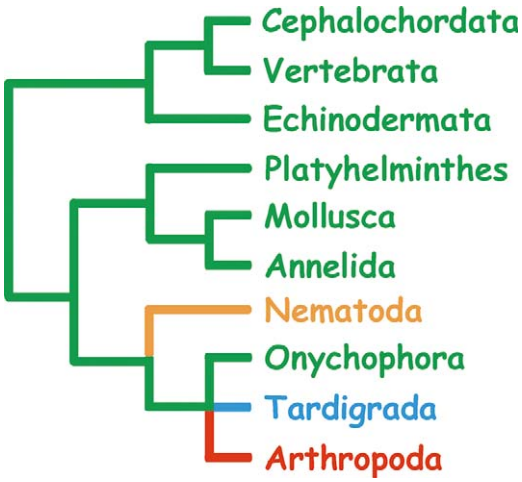


Figure 4. Phylogeny of short telomeric repeats in Bilateria. Green, $(TTAGGG)_n$; orange, $(TTAGGC)_n$; red, $(TTAGG)_n$; blue, unknown telomeric DNA motif. The cladogram is based on Giribet (2002).

The presence of TTAGGG telomeric repeats in the phylogenetically distant kinetoplastid protozoans (*Leishmania*, *Trypanosoma;* see Zakian 1995, Chiurillo & Ramirez 2002) may have a different origin. Before speculating about the $(TTAGGG)_n$ motif in the eukaryote ancestor, we should be aware that at least once, within the plant order Asparagales, the $(TTAGGG)_n$ motif evolved independently from the original plant telomeric sequence, $(TTTAGGG)_n$ (Sýkorová *et al.* 2003a).

Phylogenetic analyses of short telomeric repeats throughout the Eukaryota show that they are often conserved evolutionarily for hundreds of millions of years. This particularly concerns TTAGG, TTAGGG, and TTTAGGG repeats. On the other hand, a few minor monophyletic groups escaped that rule and displayed heterogeneity in telomeric DNA like Coleoptera (Frydrychová & Marec 2002) and Asparagales (Sýkorová *et al.* 2003a), and/or evolved telomerase-independent mechanisms of telomere maintenance like Diptera (Biessmann & Mason 2003).

Acknowledgements

We wish to thank M. Korchová for technical assistance and our colleagues, R. Frydrychová and A. Rego, for help with laboratory procedures. We also thank J. Erhart, B. Hlivka, J. Matěna, R. Pichlová, L. Stará, K. Voženílková, and R. Zemek (České Budějovice), F. Kovařík, S. Pekár, and F. Šťáhlavský (Prague), L'. Kováč and P. L'uptáčik (Košice, Slovakia), and P. Weygoldt (Freiburg, Germany) for collecting and/or supplying the specimens examined. Our thanks are also due to K. Tajovský (České Budějovice) for determination of isopods and myriapods, and L. Klimeš (Třeboň, Czech Republic) for determination of harvestmen. We are very grateful to W. Böckeler (Kiel, Germany) for help in obtaining

the pentastomids, K. Miyazaki (Kyoto, Japan) for donating the pycnogonids, H. Dastych (Hamburg, Germany) for suggestions and determination of water bears, T. Altiero (Modena, Italy) for advice on chromosome preparation from water bears, and M. H. Walker (Leicester, UK) and H. Ruhberg (Hamburg, Germany) for providing us with tissues of velvet worms. This work was supported by Grant No. 206/00/0750 of the Grant Agency of the Czech Republic (Prague) and from the Entomology Institute projects Z5007907 and K5052113. J.K. was supported by grant No. 113100003 from the Ministry of Education, Youth and Sports of the Czech Republic.

References

Abdennadher M, Mills D (2000) Telomere-associated RFLPs and electrophoretic karyotyping reveal lineage relationships among race-specific strains of *Ustilago hordei*. *Curr Genet* **38**: 141–147.

Anzai T, Takahashi H, Fujiwara H (2001) Sequence-specific recognition and cleavage of telomeric repeat $(TTAGG)_n$ by endonuclease of non-long terminal repeat retrotransposon TRAS1. *Mol Cell Biol* **21**: 100–108.

Bhattacharyya A, Blackburn EH (1997) *Aspergillus nidulans* maintains short telomeres throughout development. *Nucleic Acids Res* **25**: 1426–1431.

Beaton MJ, Hebert PDN (1994) Variation in chromosome numbers of *Daphnia* (Crustacea, Cladocera). *Hereditas* **120**: 275–279.

Biessmann H, Mason JM (2003) Telomerase-independent mechanisms of telomere maintenance. *Cell Mol Life Sci* **60**: 2325–2333.

Blackburn EH (1991) Structure and function of telomeres. *Nature* **350**: 569–573.

Castro LFC, Holland PWH (2002) Fluorescent *in situ* hybridisation to amphioxus chromosomes. *Zool Sci* **19**: 1349–1353.

Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int J Syst Evol Microbiol* **52**: 297–354.

Cavalier-Smith T (2003) Protist phylogeny and the high-level classification of Protozoa. *Eur J Protistol* **39**: 338–348.

Cavalier-Smith T, Chao EEY, Oates B (2004) Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont *Phalansterium*. *Eur J Protistol* **40**: 21–48.

Chiurillo MA, Ramirez JL (2002) Charaterization of *Leishmania major* Friedlin telomeric terminus. *Mem Inst Oswaldo Cruz* **97**: 343–346.

Colihueque N, Gajardo G (1996) Chromosomal analysis in *Artemia* populations from South America. *Cytobios* **88**: 141–148.

de Lange T (2004) T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* **5**: 323–329.

Fajkus J, Zentgraf U (2002) Structure and maintenance of chromosome ends in plants. In: Krupp G, Parwaresch R, eds. *Telomeres and Telomerases: Cancer and Biology*. Austin, Texas: Landes Bioscience, pp 314–331.

Fortey RA, Briggs DEG, Wills MA (1997) The Cambrian evolutionary 'explosion' recalibrated. *BioEssays* **19**: 429–434.

Frydrychová R, Marec F (2002) Repeated losses of TTAGG telomere repeats in evolution of beetles (Coleoptera). *Genetica* **115**: 179–187.

Frydrychová R, Grossmann P, Trubač P, Vítková M, Marec F (2004) Phylogenetic distribution of TTAGG telomeric repeats in insects. *Genome* **47**: 163–178.

Giribet G (2002) Current advances in the phylogenetic reconstruction of metazoan evolution: a new paradigm for the Cambrian explosion? *Mol Phylogen Evol* **23**: 345–357.

Giribet G, Edgecombe GD, Wheeler WC (2001) Arthropod phylogeny based on eight molecular loci and morphology. *Nature* **413**: 157–161.

Giribet G, Edgecombe GD, Wheeler WC, Babbitt C (2002) Phylogeny and systematic position of Opiliones: a combined analysis of chelicerate relationships using morphological and molecular data. *Cladistics* **18**: 5–70.

Greider CW (1996) Telomere length regulation. *Annu Rev Biochem* **65**: 337–365.

Helle W, Bolland HR (1967) Karyotypes and sex-determination in spider mites (Tetranychidae). *Genetica* **38**: 43–53.

Hirai H, LoVerde PT (1996) Identification of the telomeres on *Schistosoma mansoni* chromosomes by FISH. *J Parasitol* **82**: 511–512.

Kahn J (1964) Cytotaxonomy of ticks. *Q J Microsc Sci* **105**: 123–137.

Klapper W, Kühne K, Singh KK, Heidorn K, Parwaresch R, Krupp G (1998) Longevity of lobsters is linked to ubiquitous telomerase expression. *FEBS Letters* **439**: 143–146.

Krupp G, Klapper W, Parwaresch R (2000) Cell proliferation, carcinogenesis and diverse mechanisms of telomerase regulation. *Cell Mol Life Sci* **57**: 464–486.

Lavrov DV, Brown WM, Boore JL (2004) Phylogenetic position of the Pentastomida and (pan)crustacean relationships. *Proc R Soc Lond B Biol Sci* **271**: 537–544.

Levis C, Giraud T, Dutertre M, Fortini D, Brygoo Y (1997) Telomeric DNA of *Botrytis cinerea*: a useful tool for strain identification. *FEMS Microbiol Let* **157**: 267–272.

Maas A, Waloszek D (2001) Cambrian derivatives of the early arthropod stem lineage, pentastomids, tardigrades and lobopodians – an 'Orsten' perspective. *Zool Anz* **240**: 451–459.

Mallatt JM, Garey JR, Shultz JW (2004) Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Mol Phylogen Evol* **31**: 178–191.

Meyne J, Ratliff RL, Moyzis RK (1989) Conservation of the human telomere sequence $(TTAGGG)_n$ among vertebrates. *Proc Natl Acad Sci USA* **86**: 7049–7053.

Moyzis RK, Buckingham JM, Cram LS *et al*. (1988) A highly conserved repetitive DNA sequence, $(TTAGGG)_n$, present at the telomeres of human chromosomes. *Proc Natl Acad Sci USA* **85**: 6622–6626.

Muller F, Wicky C, Spicher A, Tobler H (1991) New telomere formation after developmentally regulated chromosomal

breakage during the process of chromatin diminution in *Ascaris lumbricoides*. *Cell* **67**: 815–822.

Niedermaier J, Moritz KB (2000) Organization and dynamics of satellite and telomere DNAs in *Ascaris*: Implications for formation and programmed breakdown of compound chromosomes. *Chromosoma* **109**: 439–452.

Okazaki S, Tsuchida K, Maekawa H, Ishikawa H, Fujiwara H (1993) Identification of a pentanucleotide telomeric sequence, $(TTAGG)_n$ in the silkworm *Bombyx mori* and in other insects. *Mol Cell Biol* **13**: 1424–1432.

Pardue ML, DeBaryshe PG (2003) Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu Rev Genet* **37**: 485–511.

Patterson DJ (1999) The diversity of eukaryotes. *Am Nat* **154** (Suppl): S96–S124.

Peterson KJ, Lyons JB, Nowak KS, Takacs CM, Wargo MJ, McPeek MA (2004) Estimating metazoan divergence times with a molecular clock. *Proc Natl Acad Sci USA* **101**: 6536–6541.

Pich U, Fuchs J, Schubert I (1996) How do Alliaceae stabilize their chromosome ends in the absence of TTTAGGG sequences? *Chromosome Res* **4**: 207–213.

Plohl M, Prats E, Martinez-Lage A, Gonzalez-Tizon A, Mendez J, Cornudella L (2002) Telomeric localization of the vertebrate-type hexamer repeat, $(TTAGGGG)_n$, in the wedgeshell clam *Donax trunculus* and other marine invertebrate genomes. *J Biol Chem* **277**: 19839–19846.

Rebecchi L, Altiero T, Bertolani R (2002) Banding techniques on tardigrade chromosomes: the karyotype of *Macrobiotus richtersi* (Eutardigrada, Macrobiotidae). *Chromosome Res* **10**: 437–443.

Rego A, Marec F (2003) Telomeric and interstitial telomeric sequences in holokinetic chromosomes of Lepidoptera: Telomeric DNA mediates association between post-pachytene bivalents in achiasmatic meiosis of females. *Chromosome Res* **11**: 681–694.

Sahara K, Marec F, Traut W (1999) TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res* **7**: 449–460.

Selden PA, Gall JC (1992) A Triassic mygalomorph spider from the northern Vosges, France. *Palaeontology* **35**: 211–235.

Selden PA, Anderson JM, Anderson HM, Fraser NC (1999) Fossil araneomorph spiders from the Triassic of South Africa and Virginia. *J Arachnol* **27**: 401–414.

Sherborn BJ, Walker MH (2004) A new species of *Peripatopsis* from South Africa, *P. stelliporata*, with observations on embryonic development and sperm degradation (Onychophora, Peripatopsidae). *J Zool* **264**: 295–305.

Šťáhlavský F (2000) *Pseudoscorpions (Pseudoscorpiones) of Prague (Ecology, Morphology, Karyology)*. MSc thesis (in Czech). Charles University, Prague.

Sýkorová E, Lim KY, Kunická Z *et al.* (2003a) Telomere variability in the monocotyledonous plant order Asparagales. *Proc R Soc Lond B Biol Sci* **270**: 1893–1904.

Sykorova E, Lim KY, Chase MW *et al.* (2003b) The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (Solanaceae): first evidence from eudicots. *Plant J* **34**: 283–291.

Underwood AP, Louis EJ, Borts RH, Stringer JR, Wakefield AE (1996) *Pneumocystis carinii* telomere repeats are composed of TTAGGG and the subtelomeric sequence contains a gene encoding the major surface glycoprotein. *Mol Microbiol* **19**: 273–281.

Vitturi R, Libertini A, Armetta F, Sparacino L, Colomba MS (2002) Chromosome analysis and FISH mapping of ribosomal DNA (rDNA), telomeric $(TTAGGG)_n$ and $(GATA)_n$ repeats in the leech *Haemopis sanguisuga* (L.) (Annelida: Hirudinea). *Genetica* **115**: 189–194.

Vitturi R, Sineo L, Volpe N, Lannino A, Colomba M (2004) Repetitive DNAs in the slug *Milax nigricans*: association of ribosomal (18S–28S and 5S rDNA) and $(TTAGGG)_n$ telomeric sequences) in the slug *M. nigricans* (Mollusca: Gastropoda: Pulmonata). *Micron* **35**: 255–260.

Wang YP, Guo XM (2001) Chromosomal mapping of the vertebrate telomeric sequence $(TTAGGG)_n$ in four bivalve molluscs by fluorescence *in situ* hybridization. *J Shellfish Res* **20**: 1187–1190.

Woznicki P, Wytwer J, Kulesza M (2003) Chromosome study of *Lithobius forficatus* (Lithobiomorpha, Chilopoda). *Folia Biol Krakow* **51**: 147–150.

Zakian VA (1995) Telomeres: Beginning to understand the end. *Science* **270**: 1601–1607.

Zrzavý J (2001) The interrelationships of metazoan parasites: a review of phylum- and higher-level hypotheses from recent morphological and molecular phylogenetic analyses. *Folia Parasitol* **48**: 81–103.