diversity

MDPI

Article

New Insights into the Telomere Structure in Hemiptera (Insecta) Inferred from Chromosome-Level and Scaffold-Level Genome Assemblies

Desislava Stoianova [1,2,*], Snejana Grozeva [1], Nadezhda Todorova [1], Miroslav Rangelov [3], Vladimir A. Lukhtanov [4] and Valentina G. Kuznetsova [4,*]

[1] Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 1000 Sofia, Bulgaria; sngrov@gmail.com (S.G.); nadeshdahr@gmail.com (N.T.)
[2] National Museum of Natural History, Bulgarian Academy of Sciences, 1000 Sofia, Bulgaria
[3] Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria; miroslav.rangelov@orgchm.bas.bg
[4] Department of Karyosystematics, Zoological Institute, Russian Academy of Sciences, 199034 St. Petersburg, Russia; vladimir.lukhtanov@zin.ru
* Correspondence: d.st.stoianova@gmail.com (D.S.); valentina.kuznetsova@zin.ru (V.G.K.)

Abstract

Telomeres are terminal regions of chromosomes that protect and stabilize chromosome structures. Telomeres are usually composed of specific DNA repeats (motifs) that are maintained by telomerase and a complex of specific proteins. Telomeric DNA sequences are generally highly conserved throughout the evolution of different groups of eukaryotes. The most common motif in insects is TTAGG, but it is not universal, including in the large order Hemiptera. In particular, several derived telomeric motifs were identified in this order by analyzing chromosome-level genome assemblies or by FISH screening the chromosomes of target species. Here, we analyzed chromosome-level genome assemblies of 16 species from three hemipteran suborders, including Sternorrhyncha (Coccoidea: *Planococcus citri*, *Acanthococcus lagerstroemiae*, and *Trionymus diminutus*; Aphidoidea: *Tuberolachnus salignus*, *Metopolophium dirhodum*, *Rhopalosiphum padi*, and *Schizaphis graminum*), Auhenorrhyncha (Cicadomorpha: *Allygus modestus*, *Arthaldeus pascuellus*, *Aphrophora alni*, *Cicadella viridis*, *Empoasca decipiens*, and *Ribautiana ulmi*), and Heteroptera (Gerromorpha: *Gerris lacustris*; Pentatomomorpha: *Aradus depressus* and *A. truncatus*). In addition, scaffold-level genome assemblies of three more species of Heteroptera (Gerromorpha: *Gerris buenoi*, *Microvelia longipes*, and *Hermatobates lingyangjiaoensis*) were examined. The presumably ancestral insect motif TTAGG was found at the ends of chromosomes of all species studied using chromosome-level genome assembly analysis, with four exceptions. In *Aphrophora alni*, we detected sequences of 4 bp repeats of TGAC, which were tentatively identified as a telomeric motif. In *Gerris lacustris*, from the basal true bug infraorder Gerromorpha, we found a 10 bp motif TTAGAGGTGG, previously unknown not only in Heteroptera or Hemiptera but also in Arthropoda in general. Blast screening of the scaffold-level assemblies showed that TTAGAGGTGG is also likely to be a telomeric motif in *G. buenoi* and *Microvelia. longipes*, while the results obtained for *H. lingyangjiaoensis* were inconclusive. In *A. depressus* and *A. truncatus* from the basal for Pentatomomorpha family Aradidae, we found a 10 bp motif TTAGGGATGG. While the available data allowed us to present two alternative hypotheses about the evolution of telomeric motifs in Heteroptera, further data are needed to verify them, especially for the yet unstudied basal infraorders Enicocephalomorpha, Dipsocoromorpha, and Leptopodomorpha.

check for updates

Academic Editors: Jader de Oliveira and Hélcio Reinaldo Gil-Santana

Received: 29 June 2025
Revised: 23 July 2025
Accepted: 29 July 2025
Published: 4 August 2025

Citation: Stoianova, D.; Grozeva, S.; Todorova, N.; Rangelov, M.; Lukhtanov, V.A.; Kuznetsova, V.G. New Insights into the Telomere Structure in Hemiptera (Insecta) Inferred from Chromosome-Level and Scaffold-Level Genome Assemblies. *Diversity* **2025**, *17*, 552. https://doi.org/10.3390/d17080552

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Keywords: telomeres; Hemiptera; true bugs; chromosome evolution; TTAGG; genome assemblies

1. Introduction

Telomeres, the physical ends of chromosomes in eukaryotes, are essential structures containing, most commonly, short tandem DNA repeats and associated proteins. Telomeres play a vital role in multiple cellular processes. They ensure the integrity and replication of chromosomes by protecting their ends from rupture and subsequent fusion with other chromosomes using a telomere-specific multiprotein structure called the telomere cap. In most eukaryotes, a ribonucleoprotein reverse transcriptase enzyme (telomerase) is involved in telomere length maintenance. This specialized reverse transcriptase (TERT) uses an internal RNA template molecule (TR) to add short, simple repeats to chromosome ends [1–6]. The main characteristics of telomeric DNA are its sequence and length. Because of their essential function, telomeres are thought to be under strong evolutionary control and resistant to molecular change. Indeed, telomeric sequences vary between large taxonomic groups and are considered group-specific [3]. Comparative analysis of telomeres in different eukaryotic organisms has shown that there are predominant telomeric repeats in certain groups, with $(TTTAGGG)_n$ in plants, $(TTAGGC)_n$ in nematodes, $(TTAGG)_n$ in arthropods, including insects, and $(TTAGGG)_n$ in vertebrates, although these latter are also observed in some molds, fungi, and protozoans [4,7–11]. Although the telomeric repeat monomers in these major taxa differ by only one or two nucleotides, even such minor variations in their sequence are considered evolutionarily significant [12].

Among insects, the $(TTAGG)_n$ sequence is found in most of the orders studied, and it is considered to be the ancestral telomeric motif for both them and all arthropods [9,13–16]. However, it has been shown that the ancestral motif was repeatedly lost during insect evolution, and although alternative motifs have been discovered in some groups, in most cases, telomeric motifs remain unknown [15,17–22]. By analyzing available chromosome-level genome assemblies, a number of alternative telomeric motifs have been identified in several insect orders [18,20,23]. The orders studied to date for telomere structure represent all four major clades of winged insects (Pterygota), such as Palaeoptera (Odonata [13]), Polyneoptera (Orthoptera [24–26], Dermaptera [20]), Paraneoptera (Hemiptera [15,18,19]), and Oligoneoptera (Neuroptera [27], Trichoptera, Coleoptera [17,28,29], Hymenoptera [16,18,22], Lepidoptera [13,30,31], and Diptera [32–35]). Based on the extensive material analyzed, Lukhtanov and Pazhenkova (2023) [18] hypothesized that in many insects there is a balance of telomerase-dependent and telomerase-independent mechanisms of telomere elongation, and a shift in this balance can lead to an increase in the variability of telomeric motifs or their complete disappearance. Furthermore, an interesting and rather unexpected conclusion was reached that the canonical telomeric sequence, which consists of only TTAGG repeats and is considered ancestral in insects, is not as widely distributed in insects as previously thought. These repeats, as such, were found in only 14 out of the 220 species studied, and in the remaining species the telomeric structure was altered to varying degrees. For example, three different non-canonical 5 bp repeats were found in Odonata, Dermaptera and Coleoptera, 10 bp repeats were found in Heteroptera (Hemiptera), and 6–11 bp repeats were found in Hymenoptera. In addition, Trichoptera, Coleoptera, Hemiptera, Hymenoptera, and Lepidoptera showed short telomeric repeats interspersed with long terminal repeats (non-LTR) of retrotransposons, such as TRAS, SART, or SART/TRAS, and Diptera showed long telomeric repeats

and telomere-targeting transposable elements instead of telomerase as a mechanism of telomere preservation [18].

The order Hemiptera (true bugs, moss bugs, cicadas, plantlice, whiteflies, scale insects, and aphids) is an ancient lineage whose fossils are known from the Early Permian [36]. It is the most species-rich hemimetabolous group (Hemimetabola, i.e., insects with incomplete metamorphosis, mostly sap-suckers), with more than 100,000 described species [37], making it the fifth largest order after Hymenoptera, Coleoptera, Lepidoptera, and Diptera [38]. Four main clades (suborders) are typically recognized in this order, including Sternorrhyncha (with superfamilies Psylloidea, Aleyrodoidea, Aphidoidea, and Coccoidea), Auchenorrhyncha (with infraorders Cicadomorpha and Fulgoromorpha, sometimes treated as separate suborders; see, e.g., [39]), Coleorrhyncha (with the only extant family Peloridiidae), and Heteroptera (with seven accepted infraorders) [40]. Currently, data on telomere structure are available for about 150 species of Hemiptera (Table S1) belonging to Auchenorrhyncha (39 species, 8 families), Sternorrhyncha (19 species, 5 families), Coleorrhyncha (1 species), and Heteroptera (90 species, 19 families) [13,15,18,19,23,41–77].

Data for the first three suborders were obtained exclusively through fluorescence in situ hybridization (FISH), which allows for direct labeling of telomeric sequences (if already known or inferred from available data on the taxonomic group to which the species belongs) at chromosome ends using oligonucleotide probes, or through blotting methods. In Sternorrhyncha and Auchenorrhyncha (formerly "Homoptera"), all species studied have canonical repeats of TTAGG at chromosome ends, although in some aphids these repeats were not detected by FISH [15]. In Coleorrhyncha, the only studied species, *Peloridium pomponorum* Shcherbakov, 2014, has a canonical telomeric motif [59]. True bugs (Heteroptera), unlike other groups of Hemiptera, show a high level of heterogeneity in telomere structure, although such heterogeneity has been found so far only in the advanced sister infraorders Pentatomomorpha and Cimicomorpha (=Terheteroptera), which are more fully studied in this respect. According to the data available to date, although still very scarce, the ancestral motif TTAGG is not dominant in true bugs, at least in Terheteroptera, where three non-canonical 10 bp motifs, TTAGGGATGG, TTAGGGGTGG, and TTAGGGTGGT, have been identified quite recently [18,19,23,78]. These are the first findings of derived telomeric motifs for the entire order Hemiptera.

The motifs listed above were identified by analyzing chromosome-level genome assemblies or by FISH screening the chromosomes of target species for candidate telomeric sequences derived from genomic assembly analysis. Thus, analyzing chromosome-level assemblies is by far the simplest straightforward method for the identification of true telomere sequence, type, and organization. Its wide application is limited only by the availability of high-quality genomic assemblies obtained by long-read sequencing and complex genomic analysis methods, including the Hi-C method, which combines chromosome conformation capture and high-throughput sequencing.

We obtained new data about the telomere sequences in 19 hemipteran species not examined before in this respect (except for *Rhopalosiphum padi* (Linnaeus, 1758), studied earlier through other methods [53]). The analyzed species represent eight families of three hemipteran suborders: Sternorrhyncha (seven species in seven genera, three families), Auchenorrhyncha (six species in six genera, two families), and Heteroptera (six species in four genera, four families). We used chromosome-level genome assemblies and, in some cases, scaffold-level assemblies.

2. Materials and Methods

Chromosome-level assemblies of Hemiptera genomes were obtained from GenBank, and their accession numbers are listed in Table 1. Most of these assemblies are based

on PacBio HiFi, whose combination of methodologies provides reliable and consistent sequence information at the chromosome level, including, to some extent, even repetitive arrays at the ends of chromosomes. Following the approach used in [18], we checked each chromosome-level assembly in the NCBI Genome Data Viewer using the "Graphics" panel and the integrated in NCBI function BLAST 2.17.0+. Genome-wide tandem repeats were identified with Tandem Repeats Finder v4.09 software [79] using the settings recommended on the TRF website:

trf <assembly.fasta> 2 7 7 80 10 50 100

These parameters correspond to Match = 2, Mismatch = 7, Delta = 7, PM = 80, PI = 10, and Minscore = 50, with MaxPeriod = 100 (search for repeat units $\leq$ 100 bp). The same configuration was applied to each assembly analyzed.

To determine whether the structure of a potential telomere motif is reproducible within a species, comparisons were made between different chromosomes and between their termini. We also followed five criteria that the identified sets of short motifs must satisfy to be listed here as the best candidates for telomeric sequences [18,20], namely, (1) they must be in a strictly terminal position, (2) their length should be at least 300–400 bp (specified as the minimum functional length required for t-loop formation in other eukaryotes (explained in [80], based on experiments with *Arabidopsis* plants; such information was not available for insects, so the minimum number of base pairs is given here as an auxiliary rather than a limiting criterion), (3) within a species, a candidate telomeric motif must be identical in all chromosomes where it is identified, (4) within each chromosome end analyzed, variant motifs must be absent or very rare, and (5) telomeric motifs at the left and right ends of the chromosome must have a reverse-complementary structure and orientation. In some eukaryotic organisms, including insects [81], insertions of specific transposable elements into telomeric tandem repeats, namely, SART and TRAS, have been described, resulting in telomeric sequences formed by both telomerase and transposable elements [18]. To search for SART and TRAS elements in the chromosomes of species with a canonical motif, we used the algorithm described in [18], where the sequences AAAAAAAAAACCTAACCTAA/TTAGGTTAGGTTTTTTTTT and CCTAACCTAACCTTTTTTTTTT/AAAAAAAAAAGGTTAGGTTAGG were used as queries for retrotransposons of the TRAS and SART families, respectively.

To identify candidate telomeric repeat motifs in each Hemiptera genome assembly, we also used the Tandem Identification Toolkit (tidk, v 0.2.65) [82], using the explore function. The search was restricted to the terminal 0.1% of every chromosome (--distance 0.001), because bona fide telomeric arrays are expected to reside in these extreme regions. We screened all k-mers from 4 to 15 bp in length (--minimum 4 --maximum 15) and reported a motif only if at least 30 uninterrupted tandem copies were detected in a single window (--threshold 30). The command used for each assembly was tidk explore --minimum 4 --maximum 15 --distance 0.01 --threshold 30 <assembly.fna>> <species>_telomere_kmers.tsv.

In addition, the scaffold-level genome assemblies of three Gerromorpha (Heteroptera) species, including *Gerris buenoi* Kirkaldy, 1911 (Gerridae), *Microvelia longipes* Uhler, 1894 (Veliidae), and *Hermatobates lingyangjiaoensis* Luo, Chen, and Wang, 2019 (Hermatobatidae), were examined. Although scaffold-level genome assemblies consist of shorter sequences (compared to chromosome-level assemblies) with unknown chromosomal positions, they are suitable for searching for sequences with telomere properties and identifying candidate telomeric repeats. Therefore, telomere-like sequences in these genomes were searched using the GenBank blastn function with the "Low complexity region" filter disabled and the $(TTAGG)_{30}$ and $(TTAGAGGTGG)_{30}$ sequences as queries

following the method used by [20]. Their accessions are, respectively, GCA_001010745.2, GCA_018340805.1, and GCA_026182355.1.

3. Results and Discussion

3.1. Variability of Telomeric Motifs in Chromosomes of Analyzed Hemiptera Species

In the chromosome-level genome assemblies of all of the studied Hemiptera species (Table 1), we found short sequence repeats at the ends of chromosomes. In 11 of the 14 "Homoptera" species, including four species of Aphidoidea (*Tuberolachnus salignus*, *Metopolophium dirhodum*, *Rhopalosiphum padi,* and *Schizaphis graminum*), three species of Coccoidea (*Planococcus citri*, *Trionymus* (=*Balanococcus*) *diminutus,* and *Acanthococcus* (=*Eriococcus*) *lagerstroemiae*), and four species of Auchenorrhyncha (*Allygus modestus*, *Arthaldeus pascuellus*, *Empoasca decipiens,* and *Ribautiana ulmi*), the canonical $(TTAGG)_n$ repeat was found at the ends of chromosomes.

A different pattern was observed in two species of Auchenorrhyncha, the European alder spittlebug *Aphrophora alni* (Aphrophoridae) and the green leafhopper *Cicadella viridis* (Cicadellidae). In the *A. alni* assembly, we found no TTAGG sequences at either end of the chromosomes, while the ends of four chromosomes were occupied by the 4 bp repeat $(TGAC)_n$, and its arrays were of considerable length, ranging from 1048 to 25,788 base pairs. The percentage of matches calculated for each array containing this repeat was 88–96%, indicating the relatively low degree of variability in this sequence characteristic of telomeric motifs. At the left and right ends of the chromosomes, where this motif was found, its sequence was $(GTCA)_n$ or $(TGAC)_n$, respectively, which corresponds to condition 5 (see above). Because it satisfies all five conditions (listed in "Materials and Methods") for being a telomeric repeat, we consider TGAC to be the best candidate for a telomeric motif in *A. alni*.

In the assembly of *C. viridis*, we found the repeat TTAGG only at only one end of chromosome 1, whereas the other end of this chromosome, one end of chromosome 2, one end of chromosome 5, one end of chromosome 6, both ends of chromosomes 7 and 8, and one end each of chromosomes 9 and 10 were occupied by a 10 bp repeat with a different nucleotide composition. Among these 10 bp repeats, the TTTGCGTGAG repeat was predominant, but, in addition to it, the following repeats were also encountered with high frequency: TTTGCGTTAG, TTTGCTTGAG, TTTGCGTGTG, TTTGCGAGAG, TTTGCGTTAG, and many others. The 10 bp structure of these repeats was strictly maintained. In the case of chromosomes 7 and 8, the repeats on opposite sides of each had a reverse-complementary structure and orientation. Thus, these repeats only partially satisfy condition 3 and do not satisfy condition 4 (see Materials and Methods) to be unequivocally recognized as telomeric motifs.

All three true bug species we studied using chromosome-level genome assembly analysis—the common pond skater *Gerris lacustris* (Gerridae, Gerromorpha) as well as the flat bugs, *Aradus depressus* (Fabricius, 1794) and *A. truncatus* Fieber, 1860 (Aradidae)—have non-canonical 10 bp telomeric motifs. According to the accepted criteria, *G. lacustris* has TTAGAGGTGG, while the two Aradidae species have TTAGGGATGG.

Note that TTAGAGGTGG was first discovered in Heteroptera and not previously known in insects in general. For three more species of the true bug infraorder Gerromorpha, namely, *G. buenoi* (Gerridae), *M. longipes* (Veliidae), and *H. lingyangjiaoensis* (Hermatobatidae), GenBank has scaffold-level genome assemblies. Blast of the *G. buenoi* and *M. longipes* genomes using the $(TTAGAGGTGG)_{30}$ sequence as a query revealed extended arrays of $(TTAGAGGTGG)_n$. Blast of these assemblies using the $(TTAGG)_{30}$ sequence did not reveal extended arrays of $(TTAGG)_n$. In contrast, blast of the *H. lingyangjiaoensis* assembly using

each of the two sequences as a query revealed that this assembly contained extended arrays of both $(TTAGAGGTGG)_{30}$ and $(TTAGG)_{30}$.

Table 1. Short telomeric motifs and presence/absence of non-LTR retrotransposons of the TRAS/SART families at chromosome ends (up to position 50,000, counting from the end of the corresponding chromosome) in Hemiptera species based on chromosome-level genome assemblies. The results for five species (marked with *) need to be verified.

Family	Species	Meioformula	Repeats at Chromosomes' Ends; Non-LTR Retrotransposons	GenBank
Auchenorrhyncha				
Aphrophoridae	*Aphrophora alni* (Fallén, 1805)	*n* = 14 + X	**(TGAC)$_n$**: Chr. 1 left end (94% matches), Chr. 4 left end, Chr. 11 left end, Chr. 14 right end; **no TRAS or SART**.	GCA_963513935.1
Cicadellidae	*Allygus modestus* Scott, 1876	*n* = 6 + X	**(TTAGG)$_n$**: Chr. 2 and 3 at both ends, Chr. 4 left end, Chr. 5 right end; **no TRAS or SART**.	GCA_963675035.1
Cicadellidae	*Arthaldeus pascuellus* (Fallén, 1826)	*n* = 8 + X	**(TTAGG)$_n$**: Chr. 1 right end, Chr. 2 left end Chr. X left end; Chr. 3 and Chr. 4 right end, Chr. 5 left end. Chr. 7 both ends; **no TRAS or SART**.	GCA_964204695.1
Cicadellidae	*Cicadella viridis* (Linnaeus, 1758) *	*n* = 10 + X	**(TTAGG)$_n$**: Chr, 1 left end; **TTTGCGTGAG, TTTGCGTTAG, TTTGCTTGAG, TTTGCGTGTG, TTTGCGAGAG, TTTGCGTTAG,** and many other 10 bp motifs: Chr. 1 right end, Chr. 2 left end, Chr. 5 left end, Chr. 6 right end; Chr. 7 and Chr. 8 both ends, Chr. 9 and Chr. 10 left end; **no TRAS or SART**.	GCA_964205205.1
Cicadellidae	*Empoasca decipiens* Paoli, 1930	*n* = 10 + X	**(TTAGG)$_n$**: Chr. 3 right end, Chr. 7 right end; **no TRAS or SART**.	GCA_964267455.1
Cicadellidae	*Ribautiana ulmi* (Linnaeus, 1758)	*n* = 8 + X	**(TTAGG)$_n$**: Chr. 1 left end, Chr. 2 left end, Chr. 7 right end; **no TRAS or SART**.	GCA_964468525.1
Sternorrhyncha				
Aphididae	*Metopolophium dirhodum* (Walker, 1849)	*n* = 9	**(TTAGG)$_n$**: Chr. 3 left end, Chr. 4 left end, Chr. 7 right end, 8 right end; **SART**: on the chromosomes where the **(TTAGG)$_n$** was present (after the 380 position counting from the ends).	GCA_019925205.1
Aphididae	*Rhopalosiphum padi* (Linnaeus, 1758) *	*n* = 4	**(TTAGG)$_n$**: Chr. 1 right end; **SART**: same chr. end (after the 3431 position).	GCA_020882245.1
Aphididae	*Schizaphis graminum* (Rondani, 1852) *	*n* = 6	**(TTAGG)$_n$**: Chr. 2 left end; **SART**: same chr. end (after the 633 position).	GCA_020882235.1

Table 1. *Cont.*

Family	Species	Meioformula	Repeats at Chromosomes' Ends; Non-LTR Retrotransposons	GenBank
Aphididae	*Tuberolachnus salignus* (Gmelin, 1790)	$n = 10$	**(TTAGG)$_n$**: Chr. 1 and Chr. 10 both ends, Chr. 4 right end, Chr. 7 right end; **SART**: Chr. 7 right end (after the 1357 position counting from the end).	GCA_956483605.1
Eriococcidae	*Acanthococcus lagerstroemiae* Borchsenius, 1960	$n = 9$	**(TTAGG)$_n$**: Chr. 1 both ends, Chr. 2 right end, Chr. 9 right end; **SART**: Chr. 1 both ends (after the 4800 position), Chr. 2 right end (after the 4800 position), Chr. 9 right end (after the 665 position).	GCA_031841125.1
Pseudococcidae	*Trionymus diminutus* (Leonardi, 1918)	$n = 5$	**(TTAGG)$_n$** Chr. 1, Chr. 3, Chr. 4 at both ends, Chr. 2 right end, Chr. 5 at both ends; **TRAS**: Chr. 2 left end (after the 3491 position), Chr. 3 left end (after the 7352 position).	GCA_959613365.1
Pseudococcidae	*Planococcus citri* (Risso, 1813)	$n = 5$	**(TTAGG)$_n$**: Chr. 2 right end; Chr. 5 right. end, Chr. 3 right end, Chr. 4 right end, Chr. 1 both ends; **TRAS** and **SART**: same chromosome ends after the 5000 position.	GCA_950023065.1
		Heteroptera		
Aradidae	*Aradus depressus* (Fabricius, 1794) *	$n = 12 + XY$	(TTAGGGATGG)$_n$: Chr. 1 left end. SART: same chromosome end (after the 7906 position).	GCA_963662175.1
Aradidae	*Aradus truncatus* Fieber, 1860	$n = 13 + XY$	(TTAGGGATGG)$_n$: Chr. Y left end, Chr. 12 left end. SART: same chromosome ends (after the 1443 position).	GCA_965153375.1
Gerridae	*Gerris lacustris* (Linnaeus, 1758),	$n = 10 + X$	**(TTAGAGGTGG)$_n$**: Chr. 2, 3, 4, 8, 10 both ends, Chr. 7 left end, Chr. X right end; **no TRAS or SART**.	GCA_951217055.1

The Tandem Identification Toolkit (tidk) function retrieved candidate telomeric motifs from each hemipteran genome assembly we analyzed (Supplement Table S2). In each case, the tool returned the sequence in its reverse complement, e.g., (TTAGG)$_n$ was represented by tidk as AACCT. Among the sequences selected by tidk, (TTAGG)$_n$ was the k-mer with the highest score in the results for six species, including *R. padi*, *P. citri*, *T. salignus*, *T. diminutus*, *A. modestus*, and *R. ulmi*, and ranked among the four top candidates for *M. dirhodum*, *S. graminum,* and *E. decipiens*. In the *A. pascuellus* results, (TTAGG)$_n$ was the 7th top candidate, in the *C. viridis* results, (TTAGG)$_n$ was the 10th candidate, and the (TTTGCGTGAG)$_n$ repeat was the 5th candidate.

The (TTAGG)$_n$ repeat was not present among the repeats selected by tidk for *G. lacustris*, *A. depressus*, and *A. alni*; instead, (TTAGAGGTGG)$_n$, (TTAGGGATGG)$_n$, and (TGAC)$_n$ were among the ten top telomeric repeat candidates, respectively.

3.2. Occurrence of SART and TRAS Retrotransposons in Hemiptera

The occurrence of non-LTR retrotransposons of the TRAS family near the chromosome ends has been reported for two aphid species, *Acyrthosiphon pisum* and *Myzus persicae* [45,54]. In five other aphid species, including four we reported here (Table 1) and *Drepanosiphum platanoidis* (Schrank, 1801) reported by [18], only SART elements were detected. We report for the first time the presence of non-LTR retrotransposons in mealybugs (Pseudococcidae); in *T. diminutus*, only TRAS insertions were detected, and in *P. citri* both TRAS and SART insertions were detected. In the present study, a telomeric repeat $(TTAGG)_n$ and SART non-LTR retrotransposons were detected in *A. lagerstroemiae*, providing the first evidence for telomeric sequences in Eriococcidae in general. There are no reports in the literature on the presence of TRAS or SART retrotransposon insertions in telomeric regions in Auchenorrhyncha. We also found no such insertions in any of the five studied species of the family Cicadellidae. In Heteroptera, SART family retrotransposon insertions have been reported in three families of Pentatomomorpha, including Coreidae, Pentatomidae, and Acanthosomatidae [18]; however, neither SART nor TRAS insertions were present in the chromosomes of *G. lacustris* (Gerridae).

Insertions of SART and TRAS retrotransposons into telomeric regions, resulting in the formation of multilayer telomeric DNA sequences, are found in many other insects belonging to the orders Coleoptera, Lepidoptera, Hymenoptera, and Trichoptera [18,81,83–85]. The compatibility of this type of telomeric sequence organization with a wide diversity of telomeric motifs suggests that multilayer telomeric sequences represent an ancestral state, at least for the clades Endopterygota [86] and Paraneoptera [18]. Thus, the absence of TRAS and SART retrotransposons is most likely the result of their secondary loss. According to [18], the dualism of the mechanisms of telomere length maintenance can explain the tendency of transition from the evolutionary stability of the TTAGG motif to the emergence of a greater diversity of motifs.

3.3. Diversity of Telomeric Repeats in Hemiptera

Although data on telomeric sequences in Hemiptera are sparse and fragmentary, their analysis suggests that the ancestral "insect" telomeric motif TTAGG is likely ancestral in this order as well [15]. This is evidenced by the fact that it occurs in each of the four major clades of Hemiptera, and derived motifs are still unknown in Auchenorrhyncha, Sternorrhyncha, and Coleorrhyncha [15,87]. In Sternorrhyncha, the only species in which the TTAGG motif was not found are the Russian wheat aphid *Diuraphis noxia* (Mordviko, 1913) and *Amphorophora tuberculata* Brown and Blackman, 1985 (Aphidoidea) [42,47,88]. It is worth adding that mapping of RNA-seq reads revealed the expression of six genes encoding telomere-associated protein in *D. noxia*, suggesting the existence of modified telomeric repeats [88]. Our results for ten "Homoptera" species, previously unstudied for telomeric motifs, are consistent with literature data indicating that aphids, scale insects, and leafhoppers are characterized by the canonical telomeric motif TTAGG [13,15,48,52,53,55,61,62]. Our results confirm the presence of the $(TTAGG)_n$ repeat in the telomeres of the bird cherry-oat aphid *R. padi*, as previously reported by [53]. In Auchenorrhyncha, the canonical motif was previously found in eight species of the spittlebug genus *Philaenus* Stål, 1864 [51]. Although *A. alni* we studied belongs to the same family, Aphrophoridae, it lacks the canonical motif. The detection of TGAC repeats at the chromosome ends of this species genome assembly, which are presumed to be true telomeres, is, to our knowledge, the first finding of such a nucleotide composition and such a number of base pairs in insects.

Surprising results were obtained for *C. viridis*. In the genome assembly of this species, the ends of most chromosomes were occupied by 10 bp sequences representing different variants of the predominant repeat $(TTTGCGTGAG)_n$, whereas TTAGG was found

at only one end of one chromosome. The 10 bp repeats do not satisfy two of the conditions, listed in Material and Methods section, therefore we cannot consider them telomeric motifs. On the one hand, we can assume that these repeats are part of the subtelomeric regions of chromosomes, and the true telomeres were not assembled. It is known that accurate assembly of telomeric regions can be very challenging due to their repetitive nature—repetitive sequences can cause misalignments or gaps in the assembly process [89–91], and as a result, telomeric regions may be missing or fragmented in some of the assembly datasets studied here. On the other hand, the very large length of the arrays of these telomere-like sequences (at least 17,000 base pairs) and the strict adherence to the 10 bp structure may suggest that these variable 10 bp repeats are part of true telomeres. Additionally, although both TGAC and TTTGCGTGAG lack the G-rich stretches typical for telomeric sequences, there are molecular mechanisms, making it possible to maintain chromosome integrity with such tandem repeats (in the studied species, respectively). It is the formation of G-quadruplexes in a intermolecular way, where no need of G-rich stretches at the ends are needed. Moreover, in repetitions of TTAGG motifs (found in many insects) appeared to be in equilibrium with non-G-quadruplex structures (likely hairpin-duplexes) [92]. Studies of additional specimens using different methods may provide insight into the organization of telomeric DNA in both *A. alni* and *C. viridis.*

As discussed in the Introduction, the true bug suborder Heteroptera is currently the only one among the Hemiptera in which alternative telomeric motifs in addition to the canonical TTAGG have been found. The three previously known alternative motifs, TTAGGGATGG, TTAGGGGTGG, and TTAGGGTGGT, are composed of 10 nucleotides, comprise the TTAGG set, and differ from each other by one or two bases [18,19,23,78]. Tracing the distribution of different telomeric motifs on the phylogenetic tree of Heteroptera (according to [93]) led to the suggestion that the $(TTAGGGATGG)_n$ repeat may be ancestral in the Pentatomomorpha + Cimicimorpha (=Terheteroptera) clade, although other hypotheses have also been proposed [23]. In the *G. lacustris* genome assembly, we found $(TTAGAGGTGG)_n$ sequences at the ends of chromosomes, i.e., another, previously unknown variant of the 10 bp motif. This species belongs to the family Gerridae (or water striders) from the basal infraorder Gerromorpha (semi-aquatic bugs). The genus *Gerris* Fabricius, 1794 comprises 42 species occurring mainly in the temperate zones of North America, Europe, and Asia [94]. Blast screening of scaffold-level assemblies of another species of this genus, *G. buenoi*, as well as *M. longipes* from the family Veliidae, using the $(TTAGAGGTGG)_{30}$ sequence as a query revealed that these assemblies contain extended arrays of TTAGAGGTGG, suggesting that this sequence is likely a telomeric motif in the chromosomes of these two species as well.

The four 10 bp motifs identified in Heteroptera show remarkable similarity, differing only by single nucleotide substitutions at positions 5 (GA), 7 (GAT), 8 (GT), and 10 (GT). The high degree of similarity suggests a common origin of these motifs, i.e., a single 10 bp motif arose within Heteroptera and subsequently diverged through modifications in different infraorders.

These results force a reconsideration of previous evolutionary scenarios for the 10 bp motif in true bugs. Initially, this motif was thought to have arisen exclusively in Pentatomomorpha, serving as a synapomorphy for this infraorder [18]. Subsequently, its discovery in Cimicomorpha led to the hypothesis that the motif arose in the common ancestor of these two sister infraorders [19,23]. However, our detection of TTAGAGGTGG in Gerridae and possibly also in Veliidae suggests a much earlier origin a 10 bp motif in Heteroptera evolution (Figure 1). Alternatively, each infraorder may have independently acquired the 10-bp repeat through extension of the ancestral pentameric sequence $(TTAGG)_n$. In this scenario, duplications in the telomerase RNA template would have

created a second pentanucleotide block, followed by the accumulation of point mutations at positions 5, 7, 8 and 10. This would represent convergent evolution of motif length rather than sequence. The phylogenetic distribution of tetomeric sequences in Heteroptera presents a complex pattern: in Cimicomorpha, the 5 bp state is found only in Reduviidae, the basal lineage of this infraorder, consistent with the hypothesis of independent evolution. However, in Pentatomomorpha, two species of the basal lineage Aradidae already possess the 10 bp motif, contrary to the expectation that basal lineages should retain the ancestral 5 bp sequence. This discrepancy suggests either that the 10 bp motif arose very early in the evolution of Pentatomomorpha, or that the evolutionary history of telomeric motifs in Heteroptera is more complex than either hypothesis alone can explain.

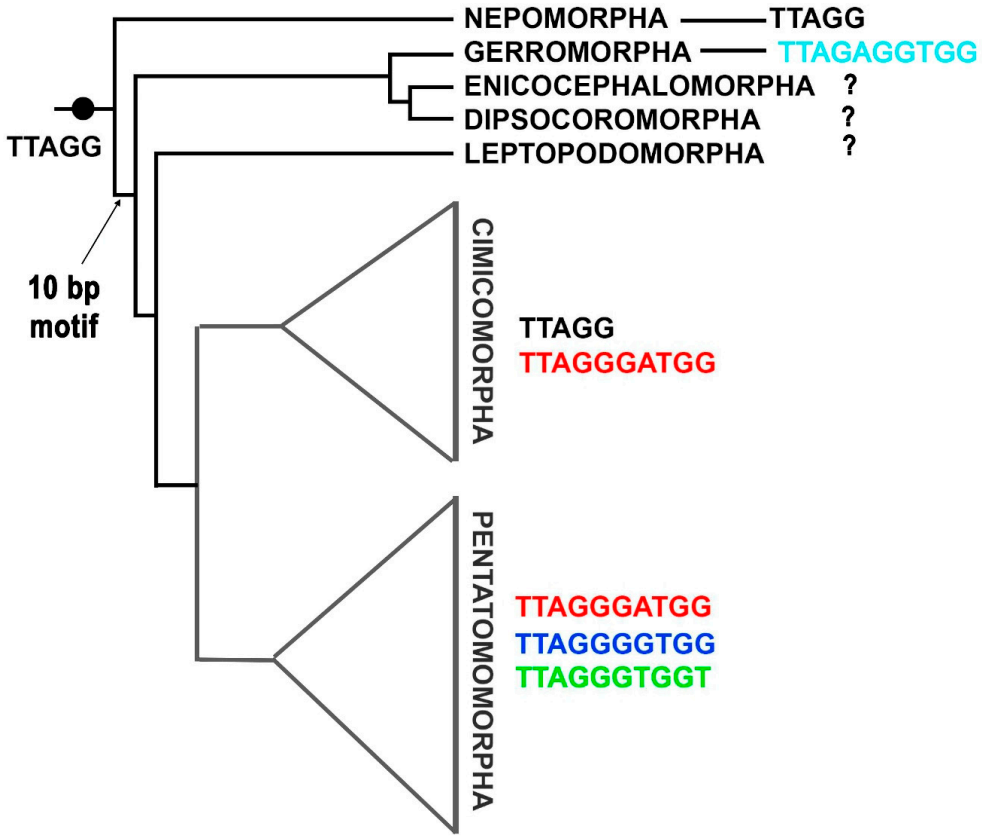

Figure 1. A hypothetical scheme describing one of the possible scenarios for the evolution of telomeric motifs in Heteroptera. The probable ancestral TTAGG motif is retained in the infraorder Nepomorpha. Within the infraorder Cimicomorpha, the canonical 5 bp insect motif TTAGG appears in one family (Reduviidae, see [23]), probably as a reversion to an ancestral state. The scheme is based on the phylogeny proposed by [93].

For a reliable reconstruction of the ancestral state of telomere organization in true bugs, data on telomeric motif of the potentially basal heteropteran infraorders—Enicocephalomorpha, Dipsocoromorpha and Leptopodomorpha, which remain unstudied are needed. Of no less importance are additional data on basal families of other infraorders. If the earliest-diverging taxa within each infraorder retain the $(TTAGG)_n$ sequence while the 10 bp repeat appears only in more derived families, this pattern would strongly support independent evolution of the 10 bp motif in each infraorder. Furthermore, we cannot exclude the possibility of finding 10 bp telomeric repeats in derived families of Nepomorpha, since current data are limited to the basal superfamily Nepoidea (families Belostomatidae and Nepidae [50,64,65]).

The marked variability in telomeric DNA sequence observed in Heteroptera (see also Table S1) is reminiscent of the high diversity of telomeric repeats found in Hymenoptera [16,18], where telomerase RNA has undergone a switch to plant/ciliate-like biogenesis [95]. This switch represents a fundamental difference from other animals and fungi, where RNA structure and its elements are different. This evolutionary transition has significant potential consequences. The presence of multiple stem loops located 3′ of the template and 5′ of the pseudoknot in telomerase RNA may explain the high diversity of probable telomere sequences (several prevalent telomere motifs) observed in Hymenoptera [16,18]. A similar mechanism may operate in Heteroptera, which might explain the documented diversity of telomeric sequences in true bugs. This hypothesis can be tested by studying telomerase RNA in Hemiptera using an approach similar to that used by Fajkus et al. [95].

4. Conclusions

The conducted study has expanded our knowledge of telomeres in Hemiptera, providing further evidence of variability of the molecular structure in this large insect order. By analyzing the chromosome-level or scaffold-level genome assemblies available for 17 hemipteran species, we (1) confirmed the predominance of the "insect" telomeric motif TTAGG in Auchenorrhyncha and Sternorrhyncha, (2) found a unique 4 bp telomeric motif in Auchenorrhyncha (in the family Aphrophoridae), which represents the first finding of the non-canonical motif in "Homoptera" as a whole, (3) found a 10 bp telomeric motif TTAGAGGTGG (for the first time for Heteroptera) in the basal true bug infraorder Gerromorpha (in the families Gerridae and Veliidae), and (4) found a 10 bp motif TTAGGGATGG in two species of the family Aradidae, which is basal for Pentatomomorpha. These new data allowed us to hypothesize the evolution of telomeric motifs in Heteroptera as a whole.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/d17080552/s1, Table S1: Telomeric motifs found in Hemiptera: published and new data; Table S2: Candidate telomeric motifs retrieved using the Tandem Identification Toolkit (tidk) function from each Hemipteran genome assembly.

Author Contributions: Conceptualization, D.S. and V.G.K.; methodology, D.S. and V.A.L.; software, D.S. and V.A.L.; validation, all authors.; formal analysis, all authors; investigation, all authors; resources, all authors; data curation, all authors; writing—original draft preparation, D.S., V.G.K. and S.G.; writing—review and editing, all authors; visualization, V.A.L.; supervision D.S. and V.G.K.; project administration, D.S. and V.G.K.; funding acquisition, D.S., V.G.K., V.A.L. and N.T. All authors have read and agreed to the published version of the manuscript.

Funding: The study of the genome assemblies in Auchenorrhyncha was performed within the framework of the State Research Project No. 125012901042-9 (V.G.K. and V.L). Analysis of the genome assemblies in Sternorrhyncha was supported by the Russian Science Foundation Grant No. 24-14-00047 (https://rscf.ru/project/24-14-00047/ (accessed on 27 June 2025)), awarded to the Zoological Institute of the Russian Academy of Sciences (St. Petersburg, Russian Federation). Analysis of the genome assemblies in Heteroptera was supported by the Bulgarian National Science Fund Grant No. KP06-PN71/13, BG-175467353-2023-130067, awarded to the National Museum of Natural History, Bulgarian Academy of Sciences. The research carried out by N.T. was supported by the Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences budget project task 3 "Ecological and genetic assessment of the environment state, management and strategies for overcoming risk".

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article and in the Supplementary Material. Further inquiries can be directed to the corresponding authors.

Acknowledgments: We thank Natalia Khabazova (Zoological Institute of the Russian Academy of Sciences, St. Petersburg) for her help in compiling the list of references. We thank Nikolay Simov and Fedor Konstantinov (National Museum of Natural History, BAS) for consultations on the correct use of taxonomic names of Heteroptera.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Muller, H.J. The remaking of chromosomes. *Collect. Net.* **1938**, *13*, 181–198.
2. Greider, C.W.; Blackburn, E.H. A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* **1989**, *337*, 331–337. [CrossRef] [PubMed]
3. Blackburn, E.H. Structure and function of telomeres. *Nature* **1991**, *350*, 569–573. [CrossRef]
4. Zakian, V.A. Telomeres: Beginning to understand the end. *Science* **1995**, *270*, 1601–1607. [CrossRef]
5. Bochman, M.L.; Paeschke, K.; Zakian, V.A. DNA secondary structures: Stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **2012**, *13*, 770–780. [CrossRef] [PubMed]
6. Greider, C.W. Telomere length regulation. *Annu. Rev. Biochem.* **1996**, *65*, 337–365. [CrossRef] [PubMed]
7. Fajkus, J.; Sýkorová, E.; Leitch, A.R. Telomeres in evolution and evolution of telomeres. *Chromosome Res.* **2005**, *13*, 469–479. [CrossRef]
8. Traut, W.; Szczepanowski, M.; Vítková, M.; Opitz, C.; Marec, F.; Zrzavý, J. The telomere repeat motif of basal Metazoa. *Chromosome Res.* **2007**, *15*, 371–382. [CrossRef]
9. Lukhtanov, V.A.; Kuznetsova, V.G. What genes and chromosomes say about the origin and evolution of insects and other arthropods. *Russ. J. Genet.* **2010**, *46*, 1115–1121. [CrossRef]
10. Anokhin, B.A.; Kuznetsova, V.G. FISH-based karyotyping of *Pelmatohydra oligactis*, *Hydra oxycnida*, and *H. magnipapillata* (Cnidaria, Hydrozoa). *Comp. Cytogenet.* **2018**, *12*, 539–548. [CrossRef]
11. Čapková Frydrychová, R. Telomerase as a possible key to bypass reproductive cost. *Mol. Ecol.* **2023**, *32*, 2134–2143. [CrossRef]
12. Peška, V.; García, S. Origin, diversity, and evolution of telomere sequences in plants. *Front. Plant Sci.* **2020**, *11*, 117. [CrossRef]
13. Frydrychová, R.; Grossmann, P.; Trubač, P.; Vítková, M.; Marec, F. Phylogenetic distribution of TTAGG telomeric repeats in insects. *Genome* **2004**, *47*, 163–178. [CrossRef]
14. Vítková, M.; Král, J.; Traut, W.; Zrzavý, J.; Marec, F. The evolutionary origin of insect telomeric repeats, (TTAGG)*n*. *Chromosome Res.* **2005**, *13*, 145–156. [CrossRef]
15. Kuznetsova, V.; Grozeva, S.; Gokhman, V. Telomere structure in insects: A review. *J. Zool. Syst. Evol. Res.* **2020**, *58*, 127–158. [CrossRef]
16. Zhou, Y.; Wang, Y.; Xiong, X.; Appel, A.G.; Zhang, C.; Wang, X. Profiles of telomeric repeats in Insecta reveal diverse forms of telomeric motifs in Hymenopterans. *Life Sci. Alliance* **2022**, *5*, e202101163. [CrossRef]
17. Prušáková, D.; Peška, V.; Pekár, S.; Bubeník, M.; Čížek, L.; Bezděk, A.; Čapková Frydrychová, R. Telomeric DNA sequences in beetle taxa vary with species richness. *Sci. Rep.* **2021**, *11*, 13319. [CrossRef]
18. Lukhtanov, V.A.; Pazhenkova, E.A. Diversity and evolution of telomeric motifs and telomere DNA organization in insects. *Biol. J. Linn. Soc.* **2023**, *140*, 536–555. [CrossRef]
19. Stoianova, D.; Grozeva, S.; Golub, N.V.; Anokhin, B.A.; Kuznetsova, V.G. The first FISH-confirmed non-canonical telomeric motif in Heteroptera: *Cimex lectularius* and *C. hemipterus* have a 10 bp motif (TTAGGGATGG)n. *Genes* **2024**, *15*, 1026. [CrossRef]
20. Lukhtanov, V.A. Telomere DNA in the insect order Dermaptera and the first evidence for the non-canonical telomeric motif TTCGG in Arthropoda. *Comp. Cytogenet.* **2025**, *19*, 13–18. [CrossRef]
21. Rico-Porras, J.M.; Mora, P.; Palomeque, T.; Cabral-de-Mello, D.C.; Lorite, P. Comparative cytogenetics of *Lachnaia* species (Coleoptera: Chrysomelidae) reveals a novel telomeric motif (TTTGG) in insects. *Org. Divers. Evol.* **2025**, *25*, 329–342. [CrossRef]
22. Menezes, R.S.T.; Bardella, V.B.; Cabral-de-Mello, D.C.; Lucena, D.A.A.; Almeida, E.A.B. Are the TTAGG and TTAGGG telomeric repeats phylogenetically conserved in aculeate Hymenoptera? *Sci. Nat.* **2017**, *104*, 85. [CrossRef]
23. Kuznetsova, V.; Golub, N.; Anokhin, B.; Stoianova, D.; Lukhtanov, V. Diversity of telomeric sequences in true bugs (Heteroptera): New data on the infraorders Pentatomomorpha and Cimicomorpha. *Cytogenet. Genome Res.* 2025, *in press*. [CrossRef]
24. López-Fernández, C.; Pradillo, E.; Zabal-Aguirre, M.; Fernández, J.L.; Garcia de la Vega, C.; Gosálvez, J. Telomeric and interstitial telomeric-like DNA sequence in Orthoptera genomes. *Genome* **2004**, *47*, 757–763. [CrossRef]
25. Jetybayev, I.E.; Bugrov, A.G.; Karamysheva, T.V.; Camacho, J.P.M.; Rubtsov, N.B. Chromosomal localization of ribosomal and telomeric DNA provides new insights on the evolution of Gomphocerinae grasshoppers. *Cytogenet. Genome Res.* **2012**, *138*, 36–45. [CrossRef]

26. Sasaki, T.; Fujiwara, H. Detection and distribution patterns of telomerase activity in insects. *Eur. J. Biochem.* **2000**, *267*, 3025–3031. [CrossRef]
27. Cabral-de-Mello, D.C.; Gasparotto, A.E.; Rico-Porras, J.M.; Ferretti, A.B.S.; Mora-Ruiz, P.; Alves-Gomes, R.T.; Lourejan, V.; Scudeler, E.L.; Lorite, P.; Bardella, V.B. First insights into the satellitomes and new evidence for the absence of canonical insect telomere in the Neuroptera order. *Genome* 2025, *ahead of print*. [CrossRef]
28. Frydrychová, R.; Marec, F. Repeated losses of TTAGG telomere repeats in evolution of beetles (Coleoptera). *Genetica* **2002**, *115*, 179–187. [CrossRef]
29. Mravinac, B.; Meštrović, N.; Čavrak, V.; Plohl, M. TCAGG, an alternative telomeric sequence in insects. *Chromosoma* **2011**, *120*, 367–376. [CrossRef]
30. Vershinina, A.O.; Anokhin, B.A.; Lukhtanov, V.A. Ribosomal DNA clusters and telomeric (TTAGG)n repeats in blue butterflies (Lepidoptera, Lycaenidae) with low and high chromosome numbers. *Comp. Cytogenet.* **2015**, *9*, 161–171. [CrossRef]
31. Rego, A.; Marec, F. Telomeric and interstitial telomeric sequences in holokinetic chromosomes of Lepidoptera: Telomeric DNA mediates association between postpachytene bivalents in achiasmatic meiosis of females. *Chromosome Res.* **2003**, *11*, 681–694. [CrossRef]
32. Rosén, M.; Edström, J. DNA structures common for chironomid telomeres terminating with complex repeats. *Insect Mol. Biol.* **2000**, *9*, 341–347. [CrossRef]
33. Danilevskaya, O.N.; Tan, C.T.; Wong, J.; Alibhai, M.; Pardue, M.L. Unusual features of the *Drosophila melanogaster* telomere transposable element HeT-A are conserved in *Drosophila yakuba* telomere elements. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3770–3775. [CrossRef] [PubMed]
34. Madalena, C.R.G.; Fernandes, T.; Villasante, A.; Gorab, E. Curiously composite structures of a retrotransposon and a complex repeat associated with chromosome ends of *Rhynchosciara americana* (Diptera: Sciaridae). *Chromosome Res.* **2010**, *18*, 587–598. [CrossRef]
35. Biessmann, H.; Zurovcova, M.; Yao, J.G.; Lozovskaya, E.; Walter, M.F. A telomeric satellite in *Drosophila virilis* and its sibling species. *Chromosoma* **2000**, *109*, 372–380. [CrossRef]
36. Shcherbakov, D.E.; Popov, Y.A. Order Hemiptera Linné, 1758. The bugs, cicadas, plantlice, scale insects, etc. In *History of Insects*; Rasnitsyn, A.P., Quicke, D.L.J., Eds.; Kluwer: Dordrecht, The Netherlands, 2002; pp. 143–157.
37. Kieran, T.J.; Gordon, E.R.; Forthman, M.; Hoey-Chamberlain, R.; Kimball, R.T.; Faircloth, B.C.; Weirauch, C.; Glenn, T.C. Insight from an ultraconserved element bait set designed for hemipteran phylogenetics integrated with genomic resources. *Mol. Phylogenet. Evol.* **2019**, *130*, 297–303. [CrossRef]
38. Grimaldi, D.; Engel, M.S. *Evolution of the Insects*; Cambridge University Press: Cambridge, UK, 2005; pp. 1–755.
39. Chen, J.; Szwedo, J.; Wang, B.; Zheng, Y.; Jiang, H.; Jiang, T.; Wang, X.; Zhang, H. A new bizarre cicadomorph family in mid-Cretaceous Burmese amber (Hemiptera, Clypeata). *Cretac. Res.* **2019**, *97*, 1–15. [CrossRef]
40. Wang, Y.H.; Wu, H.Y.; Rédei, D.; Xie, Q.; Chen, Y.; Chen, P.P.; Dong, Z.E.; Dang, K.; Damgaard, J.; Štys, P. When did the ancestor of true bugs become stinky? Disentangling the phylogenomics of Hemiptera-Heteroptera. *Cladistics* **2019**, *35*, 42–66. [CrossRef]
41. Okazaki, S.; Tsuchida, K.; Maekawa, H.; Ishikawa, H.; Fujiwara, H. Identification of a pentanucleotide telomeric sequence, (TTAGG)n, in the silkworm *Bombyx mori* and in other insects. *Mol. Cell. Biol.* **1993**, *13*, 1424–1432. [CrossRef]
42. Spence, J.M.; Blackman, R.L.; Testa, J.M.; Ready, P.D. A 169-base-pair tandem repeat DNA marker for subtelomeric heterochromatin and chromosomal rearrangements in aphids of the *Myzus persicae* group. *Chromosome Res.* **1998**, *6*, 167–175. [CrossRef]
43. Sahara, K.; Marec, F.; Traut, W. TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res.* **1999**, *7*, 449–460. [CrossRef]
44. Bizzaro, D.; Mandrioli, M.; Zanotti, M.; Giusti, M.; Manicardi, G.C. Chromosome analysis and molecular characterization of highly repeated DNAs in the aphid *Acyrthosiphon pisum* (Aphididae, Hemiptera). *Genetica* **2000**, *108*, 197–202. [CrossRef] [PubMed]
45. International Aphid Genomics Consortium. Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* **2010**, *8*, e1000313. [CrossRef]
46. Mandrioli, M.; Azzoni, P.; Lombardo, G.; Manicardi, G.C. Composition and epigenetic markers of heterochromatin in the aphid *Aphis nerii* (Hemiptera: Aphididae). *Cytogenet. Genome Res.* **2011**, *133*, 67–77. [CrossRef] [PubMed]
47. Novotná, J.; Havelka, J.; Starý, P.; Koutecký, P.; Vítková, M. Karyotype analysis of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Hemiptera: Aphididae) reveals a large X chromosome with rRNA and histone gene families. *Genetica* **2011**, *139*, 281–289. [CrossRef]
48. Mohan, K.N.; Rani, B.S.; Kulashreshta, P.S.; Kadandale, J.S. Characterization of TTAGG telomeric repeats, their interstitial occurrence and constitutively active telomerase in the mealybug *Planococcus lilacinus* (Homoptera; Coccoidea). *Chromosoma* **2011**, *120*, 165–175. [CrossRef]
49. Grozeva, S.; Kuznetsova, V.; Anokhin, B. Karyotypes, male meiosis and comparative FISH mapping of 18S ribosomal DNA and telomeric (TTAGG)n repeat in eight species of true bug (Hemiptera, Heteroptera). *Comp. Cytogenet.* **2011**, *5*, 355–374. [CrossRef]

50. Kuznetsova, V.G.; Grozeva, S.M.; Anokhin, B.A. The first finding of (TTAGG)n telomeric repeat in chromosomes of true bugs (Heteroptera, Belostomatidae). *Comp. Cytogenet.* **2012**, *6*, 341–346. [CrossRef]
51. Maryańska-Nadachowska, A.; Kuznetsova, V.G.; Karamysheva, T.V. Chromosomal location of rDNA clusters and TTAGG telomeric repeats in eight species of the spittlebug genus *Philaenus* (Hemiptera: Auchenorrhyncha: Aphrophoridae). *Eur. J. Entomol.* **2013**, *110*, 411–418. [CrossRef]
52. Monti, V.; Manicardi, G.C.; Mandrioli, M. Cytogenetic and molecular analysis of the holocentric chromosomes of the potato aphid *Macrosiphum euphorbiae* (Thomas, 1878). *Comp. Cytogenet.* **2011**, *5*, 163–172. [CrossRef]
53. Monti, V.; Giusti, M.; Bizzaro, D.; Manicardi, G.C.; Mandrioli, M. Presence of a functional (TTAGG)n telomere–telomerase system in aphids. *Chromosome Res.* **2011**, *19*, 625–633. [CrossRef]
54. Monti, V.; Serafini, K.; Manicardi, G.C.; Mandrioli, M. Characterization of non-LTR retrotransposable TRAS elements in the aphids *Acyrthosiphon pisum* and *Myzus persicae* (Aphididae, Hemiptera). *J. Hered.* **2013**, *104*, 547–553. [CrossRef]
55. Golub, N.V.; Kuznetsova, V.G.; Rakitov, R.A. First karyotype data on the family Myerslopiidae (Hemiptera, Auchenorrhyncha, Cicadomorpha). *Comp. Cytogenet.* **2014**, *8*, 293–300. [CrossRef]
56. Korandová, M.; Krůček, T.; Vrbová, K.; Frydrychová, R. Distribution of TTAGG-specific telomerase activity in insects. *Chromosome Res.* **2014**, *22*, 495–503. [CrossRef]
57. Mandrioli, M.; Zanasi, F.; Manicardi, G.C. Karyotype rearrangements and telomere analysis in *Myzus persicae* (Hemiptera, Aphididae) strains collected on *Lavandula* sp. plants. *Comp. Cytogenet.* **2014**, *8*, 259–274. [CrossRef]
58. Kuznetsova, V.G.; Maryańska-Nadachowska, A.; Anokhin, B.; Aguin Pombo, D. Evidence for TTAGG telomere repeats and rRNA gene clusters in leafhoppers of the genus *Alebra* (Hemiptera: Auchenorrhyncha: Cicadellidae). *Eur. J. Entomol.* **2015**, *112*, 207–214. [CrossRef]
59. Kuznetsova, V.G.; Grozeva, S.M.; Hartung, V.; Anokhin, B.A. First evidence for (TTAGG)n telomeric sequence and sex chromosome post-reduction in Coleorrhyncha (Insecta, Hemiptera). *Comp. Cytogenet.* **2015**, *9*, 523–532. [CrossRef]
60. Golub, N.V.; Golub, V.B.; Kuznetsova, V.G. Variability of 18 rDNA loci in four lace bug species (Hemiptera, Tingidae) with the same chromosome number. *Comp. Cytogenet.* **2015**, *9*, 513–522. [CrossRef]
61. Anjos, A.; Rocha, G.C.; Paladini, A.; Mariguela, T.C.; Cabral-de-Mello, D.C. Karyotypes and repetitive DNA evolution in six species of the genus *Mahanarva* (Auchenorrhyncha: Cercopidae). *Cytogenet. Genome Res.* **2016**, *149*, 321–327. [CrossRef]
62. Maryańska-Nadachowska, A.; Anokhin, B.A.; Gnezdilov, V.M.; Kuznetsova, V.G. Karyotype stability in the family Issidae (Hemiptera, Auchenorrhyncha) revealed by chromosome techniques and FISH with telomeric (TTAGG)n and 18S rDNA probes. *Comp. Cytogenet.* **2016**, *10*, 347–369. [CrossRef]
63. Pita, S.; Panzera, F.; Mora, P.; Vela, J.; Palomeque, T.; Lorite, P. The presence of the ancestral insect telomeric motif in kissing bugs (Triatominae) rules out the hypothesis of its loss in evolutionarily advanced Heteroptera (Cimicomorpha). *Comp. Cytogenet.* **2016**, *10*, 427–437. [CrossRef]
64. Chirino, M.G.; Dalíková, M.; Marec, F.; Bressa, M.J. Chromosomal distribution of interstitial telomeric sequences as signs of evolution through chromosome fusion in six species of the giant water bugs (*Belostoma*, Hemiptera). *Ecol. Evol.* **2017**, *7*, 5227–5235. [CrossRef]
65. Angus, R.B.; Jeangirard, C.; Stoianova, D.; Grozeva, S.; Kuznetsova, V.G. Chromosomal analysis of *Nepa cinerea* Linnaeus, 1758 and *Ranatra linearis* (Linnaeus, 1758) (Heteroptera, Nepidae). *Comp. Cytogenet.* **2017**, *11*, 641–657. [CrossRef]
66. Golub, N.V.; Golub, V.B.; Kuznetsova, V.G. Distribution of the major rDNA loci among four hemipteran species of the family Tingidae (Heteroptera, Cimicomorpha). *Folia Biol.* **2017**, *65*, 155–158. [CrossRef]
67. Maryańska-Nadachowska, A.; Kuznetsova, V.G.; Golub, N.V.; Anokhin, B.A. Detection of telomeric sequences and ribosomal RNA genes in holokinetic chromosomes of five jumping plant-lice species: First data on the superfamily Psylloidea (Hemiptera: Sternorrhyncha). *Eur. J. Entomol.* **2018**, *115*, 632–640. [CrossRef]
68. Luan, J.; Sun, X.; Fei, Z.; Douglas, A.E. Maternal inheritance of a single somatic animal cell displayed by the bacteriocyte in whitefly *Bemisia tabaci*. *Curr. Biol.* **2018**, *28*, 459–465. [CrossRef] [PubMed]
69. Golub, N.V.; Golub, V.B.; Kuznetsova, V.G. New data on karyotypes of lace bugs (Tingidae, Cimicomorpha, Hemiptera) with analysis of the 18S rDNA clusters distribution. *Comp. Cytogenet.* **2018**, *12*, 515–528. [CrossRef] [PubMed]
70. Grozeva, S.; Anokhin, B.A.; Simov, N.; Kuznetsova, V.G. New evidence for the presence of the telomere motif (TTAGG)n in the family Reduviidae and its absence in the families Nabidae and Miridae (Hemiptera, Cimicomorpha). *Comp. Cytogenet.* **2019**, *13*, 283. [CrossRef]
71. Karagyan, G.; Golub, N.; Sota, T. Cytogenetic characterization of periodical cicadas (Hemiptera: Cicadidae: *Magicicada*). *Eur. J. Entomol.* **2020**, *117*, 474–480. [CrossRef]
72. Ferguson, K.B.; Visser, S.; Dalíková, M.; Provazníková, I.; Urbaneja, A.; Pérez-Hedo, M.; Marec, F.; Werren, J.H.; Zwaan, B.J.; Pannebakker, B.A.; et al. Jekyll or Hyde? The genome (and more) of *Nesidiocoris tenuis*, a zoophytophagous predatory bug that is both a biological control agent and a pest. *Insect Mol. Biol.* **2021**, *30*, 188–209. [CrossRef]

73. Gapon, D.A.; Kuznetsova, V.G.; Maryańska-Nadachowska, A. A new species of the genus *Rhaphidosoma* Amyot et Serville, 1843 (Heteroptera, Reduviidae), with data on its chromosome complement. *Comp. Cytogenet.* **2021**, *15*, 467. [CrossRef]
74. Golub, N.V.; Golub, V.B.; Anokhin, B.A.; Kuznetsova, V.G. Comparative cytogenetics of lace bugs (Tingidae, Heteroptera): New data and a brief overview. *Insects* **2022**, *13*, 608. [CrossRef]
75. Golub, N.V.; Maryańska-Nadachowska, A.; Anokhin, B.A.; Kuznetsova, V.G. Expanding the chromosomal evolution understanding of lygaeioid true bugs (Lygaeoidea, Pentatomomorpha, Heteroptera) by classical and molecular cytogenetic analysis. *Genes* **2023**, *14*, 725. [CrossRef]
76. Jing, T.; Yang, J.; Pan, J.; Liu, X.; Yang, X.; Farhan, M.; Su, H.; Ma, X.; Zhang, S. A near-complete genome reveals the population evolution of the cotton-melon aphid *Aphis gossypii*. *Insect Biochem. Mol. Biol.* **2025**, *176*, 104215. [CrossRef]
77. Palmer Droguett, D.H.; Fletcher, M.; Alston, B.T.; Kocher, S.; Cabral-de-Mello, D.C.; Wright, A.E. Neo-sex chromosome evolution in treehoppers despite long-term X chromosome conservation. *Genome Biol. Evol.* **2024**, *16*, evae264. [CrossRef]
78. Lukhtanov, V.A. Diversity and evolution of telomere and subtelomere DNA sequences in insects. *bioRxiv* **2022**, preprint. [CrossRef]
79. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [CrossRef] [PubMed]
80. Watson, J.M.; Trieb, J.; Troestl, M.; Renfrew, K.; Mandáková, T.; Fulneček, J.; Shippen, D.E.; Říha, K. A hypomorphic allele of telomerase uncovers the minimal functional length of telomeres in *Arabidopsis*. *Genetics* **2021**, *219*, iyab126. [CrossRef]
81. Kubo, Y.; Okazaki, S.; Anzai, T.; Fujiwara, H. Structural and phylogenetic analysis of TRAS, telomeric repeat-specific non-LTR retrotransposon families in lepidopteran insects. *Mol. Biol. Evol.* **2001**, *18*, 848–857. [CrossRef]
82. Brown, M.R.; De La Rosa, P.G.; Balxter, M. tidk: A toolkit to rapidly identify telomeric repeats from genomic datasets. *Bioinformatics* **2025**, *42*, btaf049. [CrossRef] [PubMed]
83. Fujiwara, H.; Osanai, M.; Matsumoto, T.; Kojima, K.K. Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm *Bombyx mori*. *Chromosome Res.* **2005**, *13*, 455–467. [CrossRef]
84. Osanai, M.; Kojima, K.K.; Futahashi, R.; Yaguchi, S.; Fujiwara, H. Identification and characterization of the telomerase reverse transcriptase of *Bombyx mori* and *Tribolium castaneum*. *Gene* **2006**, *376*, 281–289. [CrossRef]
85. Kirkness, E.F.; Haas, B.J.; Sun, W.; Braig, H.R.; Perotti, M.A.; Clark, J.M.; Lee, S.H.; Robertson, H.M.; Kennedy, R.C.; Elhaik, E.; et al. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 12168–12173. [CrossRef]
86. Robertson, H.M.; Gordon, K.H. Canonical TTAGG-repeat telomeres and telomerase in the honey bee, *Apis mellifera*. *Genome Res.* **2006**, *16*, 1345–1351. [CrossRef]
87. Vicari, M.R.; Bruschi, D.P.; Cabral-de-Mello, D.C.; Nogaroto, V. Telomere organization and the involvement of interstitial telomeric sites in insect and vertebrate chromosome evolution. *Genet. Mol. Biol.* **2022**, *45*, e20220071. [CrossRef]
88. Nicholson, S.J.; Nickerson, M.L.; Dean, M.; Song, Y.; Hoyt, P.R.; Rhee, H.; Kim, C.; Puterka, G.J. The genome of *Diuraphis noxia*, a global aphid pest of small grains. *BMC Genom.* **2015**, *16*, 429. [CrossRef]
89. Tang, H. Genome assembly, rearrangement, and repeats. *Chem. Rev.* **2007**, *107*, 3391–3406. [CrossRef]
90. Rhie, A.; McCarthy, S.A.; Fedrigo, O.; Damas, J.; Formenti, G.; Koren, S.; Uliano-Silva, M.; Chow, W.; Fungtammasan, A.; Kim, J.; et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **2021**, *592*, 737–746. [CrossRef]
91. Kim, S.; Chowdhury, T.; Yu, H.J.; Kahng, J.Y.; Lee, C.E.; Kang, H.; Lee, J.H.; Lee, S.-T.; Won, J.-K.; Kim, K.H.; et al. The telomere maintenance mechanism spectrum and its dynamics in gliomas. *Genome Med.* **2022**, *14*, 88. [CrossRef] [PubMed]
92. Tran, P.L.T.; Mergny, J.-L.; Alberti, P. Stability of telomeric G-quadruplexes. *Nucleic Acids Res.* **2011**, *39*, 3282–3294. [CrossRef] [PubMed]
93. Weirauch, C.; Schuh, R.T.; Cassis, G.; Wheeler, W.C. Revisiting habitat and lifestyle transitions in Heteroptera (Insecta: Hemiptera): Insights from a combined morphological and molecular phylogeny. *Cladistics* **2019**, *35*, 67–105. [CrossRef] [PubMed]
94. Andersen, N.M. Classification, phylogeny, and zoogeography of the pond skater genus *Gerris* Fabricius (Hemiptera: Gerridae). *Can. J. Zool.* **1993**, *71*, 2473–2508. [CrossRef]
95. Fajkus, P.; Adámik, M.; Nelson, A.D.L.; Kilar, A.M.; Franek, M.; Bubeník, M.; Frydrychová, R.Č.; Votavová, A.; Sýkorová, E.; Fajkus, J.; et al. Telomerase RNA in Hymenoptera (Insecta) switched to plant/ciliate-like biogenesis. *Nucleic Acids Res.* **2023**, *51*, 420–433. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.