

RESEARCH

Open Access



Integrative satellitomics reveals distinct patterns of organization, transcription and evolution of satellite DNAs in *Tenebrio molitor*

Patrik Majcen¹, Antonio Sermek¹, Đurđica Ugarković¹, Brenda Oppert^{2,3}, Miroslav Plohl¹ and Eva Šatović-Vukšić^{1*}

Abstract

Satellite DNAs (satDNAs) are repetitive sequences that play important roles in chromosomal architecture, genome evolution, and regulation. Here, we present a comprehensive characterization of *Tenebrio molitor* satellitome, integrating cytogenetic mapping, in silico genome annotation, divergence profiling, screening of extrachromosomal circular DNA (eccDNA), transcription analysis across developmental stages, and phylogenetic and age analyses. SatDNAs exhibited diverse chromosomal organizations, ranging from widespread to chromosome-restricted distributions. Discrepancies between assembly-based and physical mapping highlight limitations of individual approaches and underscore the importance of their integration. Divergence landscape analyses revealed different homogenization efficiencies and turnover rates, reflecting distinct evolutionary trajectories among individual satDNAs. Phylogenetic reconstruction revealed distinct patterns which include clear species-specific clustering of monomers, mixed interspecific clustering, and dispersed topologies. Comparative analyses across insect orders enabled age estimation, identifying both ancient (≥ 380 MYA) and lineage-specific satDNAs, apparently restricted to *T. molitor*. We designed and applied an approach that enables the simultaneous detection of multiple satDNAs within the eccDNA fraction which confirmed the presence of six satDNAs in eccDNA. RNA-seq analyses revealed coordinated, stage-specific transcription of all satDNAs, with elevated expression in late male pupae and early male adults. Together, these results reveal a highly dynamic, heterogeneous, and functionally relevant satDNA landscape in *T. molitor* and demonstrate the importance of integrative approaches for understanding molecular mechanisms and trajectories of satDNA evolution.

Keywords Satellite DNA, Satellitome, Extrachromosomal circular DNA, Transcription, Phylogeny, Molecular evolution, Insect genomics, *Tenebrio molitor*

*Correspondence:

Eva Šatović-Vukšić
esatovic@irb.hr

¹Ruđer Bošković Institute, Bijenička 54, Zagreb 10000, Croatia

²Kansas State University, Manhattan 66502, KS, USA

³Entovations LLC, Manhattan 66503, KS, USA



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Introduction

Although sometimes referred to as the “dark matter of the genome”, repetitive sequences are recognized as key components of genome structure and as drivers of its evolution, since genome dynamics is strongly influenced by processes that reorganize repetitive DNAs, and modify their composition and copy number [1–5]. Satellite DNAs (satDNAs) are sequences repeated in tandem, frequently located in (peri)centromeric and subtelomeric heterochromatin, forming arrays of hundreds to thousands of monomers [3, 6–8]. They can also be found throughout euchromatin, frequently near genes or within introns, occurring as individual repeats or short arrays [9]. SatDNAs contribute to genome evolution through their involvement in speciation processes and play an essential role in maintaining the genome within the nucleus (reviewed in [10]). Increasing evidence also points to active and tightly regulated satDNA transcription, with satDNA transcripts participating in diverse cellular processes. For instance, satDNA transcripts are implicated in heterochromatin establishment, centromere function, and gene expression regulation across various biological contexts, stress responses, and environmental adaptations (reviewed in [11]). Moreover, aberrant satDNA transcription has been linked to carcinogenesis, and satDNA transcripts show potential as biomarkers for cancer detection [11]. SatDNA transcription is often developmentally regulated and tissue-specific, indicating functional relevance [12, 13]. Mechanisms that are proposed to contribute to the dispersal of satDNA sequences throughout the genome are unequal crossing over, transposition and extrachromosomal circular DNA (eccDNA) formation and reintegration [8, 14]. EccDNA molecules are circular DNA elements derived from chromosomal sequences and are recognized as widespread components of eukaryotic genomes. Tandemly repeated sequences, including satDNAs, are particularly prone to eccDNA formation [15], due to their organization and susceptibility to intra-chromosomal homologous recombination and excision of circular forms. It is suggested that eccDNAs may contribute to satDNA array length changes and satDNA homogenization, as well as genomic plasticity [16].

The yellow mealworm, *Tenebrio molitor*, is becoming a model organism of interest in a number of studies in biology, biochemistry, evolution, immunology and physiology [17–24]. Extracts from *T. molitor* have demonstrated antiproliferative activity against colorectal adenocarcinoma and hepatocellular carcinoma [25], as well as wound-healing and antimicrobial properties [26, 27]. Furthermore, this species produces antifreeze proteins that can be used for vegetable preservation [28]. Due to its high protein content and nutritional value, *T. molitor* is increasingly used as an animal dietary supplement and

has been introduced to the human food market [29–31]. The importance of studying *T. molitor* is highlighted by ongoing efforts to improve the genome assembly and annotation [32–34]. While repetitive regions pose significant technical challenges for DNA sequencing and assembly [35], a substantial portion of the *T. molitor* genome consists of a single satDNA. This satDNA has a monomer length of 142 bp and is localized in the pericentromeric heterochromatin of all chromosomes, as confirmed by in situ restriction digestion of metaphase chromosomes [36–38]. The most recent, improved genome assembly of *T. molitor* was accompanied by in silico satellitome analysis [34], in which the abundance of the major 142-bp satDNA was determined to be 26.5% (TmSat01). Ten additional low-abundant satDNAs were identified (TmSat02 - TmSat11), together constituting 1.5% of the genome assembly [34]. However, the exact chromosomal localization and distribution patterns of *T. molitor* satDNAs remained unknown, as well as their transcription and divergence profiles. As genomic databases continue to expand with an exponentially increasing number of sequenced genomes, newly available data allow for more comprehensive and accurate insights into satDNA distribution across taxa, including the profiling of satDNA conservation and divergence, and a more detailed understanding of the relationships between satDNA distribution patterns and species phylogeny.

In this study, we determined the localization of all satDNAs (TmSat01–11) in the genome of *T. molitor* using complementary in situ and in silico approaches. We analyzed their transcription across different developmental stages in males and females using RNA-seq data. Furthermore, we designed and applied an approach that enables the simultaneous detection of multiple satDNAs within the eccDNA fraction, confirming the presence of six satDNAs in eccDNA. We additionally examined sequence divergence for each *T. molitor* satDNA and assessed their presence across other insect species. These analyses allowed us to reconstruct satDNA phylogenies, explore relationships between satDNA distribution and species phylogeny, and estimate the minimal age of each satDNA family.

This study represents the first comprehensive analysis of a species’ satellitome, integrating simultaneously satDNA genomic organization, inter- and intra-species distribution, evolutionary dynamics, eccDNA association and transcriptional activity. By combining cytogenetic, genomic, phylogenetic and transcriptomic data, we provide a wider perspective of satDNA dynamics across structural, functional, taxonomical and temporal levels. Our results underline the importance of such an integrative framework for understanding the complexity of satDNA biology and support its broader

application in future studies of repetitive DNA in other species.

Materials and methods

SatDNAs annotation, divergence profiling, and genomic context inspection

Consensus sequences of the 11 *T. molitor* satDNAs (NCBI accession numbers PX395431–PX395441) were annotated on two recent publicly available chromosome-level genome assemblies of this species, GCA_037075425.1 and GCA_963966145.1. Monomer consensus sequences were used to annotate each satDNA on chromosomes and scaffolds in Geneious Prime software (Biomatters Ltd., Auckland, New Zealand), allowing up to 45% divergence from the consensus to detect monomer variants. To assess sequence divergence, all monomers of each satDNA were extracted from the genome and compared to their respective consensus sequences within Geneious Prime. Divergence landscapes were generated using Kimura 2-parameter distances and grouped into 1% divergence bins. Peaks were defined as local maxima in the abundance distribution, where a given bin or a series of neighboring bins exhibited 50% higher values than preceding and following bins. For the analysis of the surrounding regions of TmSat06 and TmSat07, the corresponding arrays, along with 1,000 bp of flanking sequences upstream and downstream of each array, were extracted and analyzed in Geneious Prime. All nucleotide sequence alignments were performed in the same software.

The TideCluster software tool, implemented within the RepeatExplorer Galaxy platform (<https://repeatexplorer-elixir.cerit-sc.cz/galaxy/>), was used for *de novo* detection of tandem repeats in the reference genome assembly GCA_963966145.1. The TideCluster software integrates TideHunter [39] and TAREAN analyses [40], and was run using default parameters.

Preparation of chromosome spreads from male pupae

The sex of the pupae was determined following Bhatnagya et al. [41]. Chromosome preparations were obtained from male pupae. Testes were dissected and immersed in distilled water for 45 min to induce osmotic shock. The tissue was then fixed in Carnoy's solution (absolute ethanol:glacial acetic acid, 3:1) and subsequently macerated in 50% glacial acetic acid in an Eppendorf tube. Droplets of the suspension were spread onto glass slides and placed on a hot plate at 42 °C. After drying, the chromosomal preparations were dehydrated in series of ethanol solutions (70%, 90%, 100%; 30 s each). The slides were stored at –20 °C until utilized.

Probe labelling

DNA probes for fluorescent in situ hybridization (FISH) were labelled by PCR. Each of the 20 µL reaction contained ~20 ng of DNA, 2.5 U GoTaq Flexi G2 DNA polymerase (Promega), GoTaq Buffer, 1.5 mM MgCl₂, primers (0.1 µM each), 100 µM dATP/dCTP/dGTP (NEB), 50 µM dTTP (NEB), 50 µM biotin-16-dUTP (Jena Bioscience). After verification by agarose gel electrophoresis, DNA probes were purified using either the QIAquick PCR Purification Kit or the QIAquick Gel Extraction Kit (Qiagen). Probe concentrations were measured with a DeNovix fluorometer. Nucleotide sequences of the primers and PCR amplification conditions are provided in Supplementary Table S1.

Fluorescent in situ hybridization (FISH)

FISH experiments were performed following the protocol outlined in Pérez-García et al. [42], with a modification in the pepsin digestion step, which was carried out for 5 min at 37 °C. Each FISH assay utilized 50 ng of satDNA probe. Prior to use, probes were denatured at 80 °C for 8 min and immediately cooled on ice for 2 min. Chromosomal DNA was denatured for 1 minute and 45 seconds in 50% formamide (Sigma-Aldrich) in 2× SSC. Signal detection was carried out using fluorescein-conjugated streptavidin (Vector Laboratories) at a 1:200 dilution, biotinylated anti-streptavidin (Vector Laboratories) at 1:100, followed by a second layer of fluorescein-conjugated streptavidin at 1:200. Finally, chromosomes were counterstained with 100 ng/mL DAPI (Sigma-Aldrich) and mounted in Mowiol 4–88 mounting medium (Sigma-Aldrich). Signals were visualized and captured using a Leica TCS SP8 X laser-scanning confocal microscope.

Analyses of transcriptomes

Publicly available RNA-seq data from *T. molitor* was used for the analyses (NCBI BioProject accession PRJNA1012330). The dataset contains male and female samples spanning multiple developmental stages: eggs, early, mid, and late larvae, early and late pupae, and early and late adults. The number of accessible biological replicates varied between stages. Raw sequencing reads were evaluated for base quality and adapter content using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Adapter trimming and removal of low-quality bases were performed with TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Filtered paired-end reads were mapped to the *T. molitor* reference genome (RefSeq accession GCF_963966145.1, NCBI) using HISAT2 [43] with default parameters. Gene-level read counts were obtained with featureCounts [44]. To identify potentially problematic libraries, read count data were variance-stabilized using the rlog transformation (DESeq2 [45]), and

subjected to principal component analysis (PCA) in R (R Core Team, 2024) (Supplementary Figure S1). Sample outliers were detected using a robust covariance estimator (EllipticEnvelope [46]); based on the first principal components. Libraries with extremely low sequencing depth or detected as outliers were excluded from subsequent analyses.

We constructed a custom database comprising 11 *T. molitor* satDNAs. Reads were mapped to satellite dimer sequences from this database to improve alignment specificity. Mapping was performed using Bowtie2 [47] with the -a option to report all alignments, allowing for a single mismatch. Mapped reads were counted using SAMtools [48]. Counts for each satDNA were normalized using the fragments per kilobase per million mapped reads (FPKM) method [49]. Due to significant copy number variation among satellites FPKM values are valid for longitudinal analysis (the same satellite DNA across different developmental stages) and not for comparison between different satellite DNAs. Data processing and visualization were carried out in R (<https://www.R-project.org/>) using the ggplot2 package [50]. Stage-specific colors and plotting aesthetics were customized to facilitate comparison across developmental stages.

Extrachromosomal circular DNA (eccDNA) experiments

Genomic DNA was isolated from larvae using a standard phenol/chloroform protocol, followed by eccDNA isolation. To remove linear DNA fragments, 400 ng of purified DNA was digested with Exonuclease V (NEB). The initial digestion was performed in a 50 μ L reaction containing 20 U of Exonuclease V, 5 μ L of 10 \times NEBuffer 4, 5 μ L of 10 mM ATP, and nuclease-free water to a final volume of 50 μ L. The reaction was incubated at 37 $^{\circ}$ C for 15 hr. Subsequently, 5 μ L ATP, 1 μ L 10 \times NEBuffer 4, and 20 U Exonuclease V were added, and the reaction continued for additional 7 h at 37 $^{\circ}$ C. Further additions of 5 μ L ATP, 0.1 μ L 10 \times NEBuffer 4, and 20 U Exonuclease V were made, and the digestion proceeded for additional 18 h at 37 $^{\circ}$ C. Finally, 5 μ L ATP, 0.1 μ L 10 \times NEBuffer 4, and 30 U Exonuclease V were added, and digestion proceeded for 6 h at 37 $^{\circ}$ C. Afterwards, the reaction was heat-inactivated at 70 $^{\circ}$ C for 30 min.

To confirm the elimination of chromosomal linear DNA and the preservation of circular DNA, PCR was performed on Exonuclease V-treated sample. As a negative control for chromosomal DNA removal, PCR with primers targeting the histone H3 gene (TM_H3F 5'-CGC TTACCGTTCGAAG-3' and TM_H3R 5'-TATTTGCAA TTCGTCGTAG-3') was performed. As a positive control for the presence of circular DNA, PCR with primers for the mitochondrial 16S gene (Tm16S_F 5'-GTATTTTGA CTGTGCAAAG-3' and Tm16S_R 5'-TTTTAATCCAA CATCGAGG-3') was performed. Primer sequences were

designed based on *T. molitor* H3 gene (XM_069038316.1) and mitochondrial 16S sequences (EU048292, AJ633673.1, AJ438153.1), available in the NCBI database. PCR reactions (10 μ L total volume) contained 5 μ L 2 \times PCR Master Mix (Promega), 0.4 μ L of each primer (10 pmol/ μ L), and 1 μ L of the Exonuclease V-treated DNA. Amplification was carried out under the following conditions: initial denaturation at 95 $^{\circ}$ C for 5 min, followed by 25 cycles of 95 $^{\circ}$ C for 30 s, 50 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 30 s, with a final extension at 72 $^{\circ}$ C for 7 min.

Phylogeny and age reconstruction

In order to perform phylogeny reconstruction, for each satDNA of *T. molitor*, 100 monomers were randomly subsampled from the genomic assembly GCF_963966145.1 in Geneious Prime. For other insect species, homologous satDNA sequences were retrieved by querying GenBank using the consensus sequences of TmSat01–TmSat11. Searches were performed with the blastn algorithm against both the Whole Genome Shotgun (WGS) and Nucleotide (nt) databases (GenBank+EMBL+DDBJ+PDB+RefSeq). All contigs with hits were downloaded, locally annotated with TmSat01–11 consensus sequences in Geneious Prime. Monomers were extracted and manually inspected. Where more than 100 were available, they were randomly subsampled to 100 sequences, following the same procedure applied to the *T. molitor* ones.

Multiple sequence alignments of satDNA monomers were performed using MAFFT alignment algorithm implemented in Geneious Prime. Maximum likelihood phylogenies were inferred using PhyML under the TN93 substitution model. Consensus trees were reconstructed using the Consensus Tree builder algorithm, and nodal support values were obtained through 100 bootstrap replicates. Divergence times of insect taxa were obtained via Timetree of life (TTOL) database [51]. Divergence time estimates for all taxa in the TTOL are pre-computed and stored in the TTOL database. The evolutionary timeline shown was constructed by the TTOL algorithm, by locating queried taxa within its precomputed database and traversing the tree toward the root, collecting divergence times at each internal node. The approximate evolutionary age of each satDNAs was inferred manually, based on its distribution across insect taxa, considering only species in which the query coverage exceeded 50% of the monomer sequence.

Results

In silico analysis of chromosomal distribution of *T. molitor* satDNAs

The chromosomal distribution of the 11 *T. molitor* satDNAs (TmSat01–TmSat11) was first examined in silico. For that purpose, we have annotated consensus

sequences of each satDNA on the two recent chromosome-level assemblies of this species. The assembly GCA_037075425.1 [34] comprises linkage groups LG1–LG10 plus LGX and 1,006 unplaced scaffolds, while the assembly GCF_963966145.1 contains nine autosomes, the X and Y chromosomes, and 225 additional scaffolds. For both assemblies, the presence or absence of each of the 11 satDNAs was recorded for each chromosome or linkage group, along with the number of monomers on unplaced scaffolds (Table 1). A substantial difference in the number of annotated satDNA monomers was

observed between the two assemblies. For example, TmSat01 is represented by only 1,689 monomers in GCF_963966145.1, compared to 12,407 monomers in GCA_037075425.1. Conversely, the low-abundance satDNAs TmSat02–TmSat11 are considerably more enriched in the GCF_963966145.1 assembly (Table 1). Therefore, the GCF_963966145.1 assembly, currently designated as the reference genome assembly, was used for subsequent analyses.

The eleven *T. molitor* satDNAs display distinct patterns of placement, distribution, and organization of

Table 1 Abundance and distribution of monomers belonging to TmSat01–TmSat11 across two *T. molitor* genome assemblies

assembly accession: GCF_963966145.1	number of monomers on chromosomes	Chr	TmSat 01	TmSat 02	TmSat 03	TmSat 04	TmSat 05	TmSat 06	TmSat 07	TmSat 08	TmSat 09	TmSat 10	TmSat 11	
		1	123	904	849	469	135						316	
		2	30		371	187	80				3			
		3	808	1084	611	412	130				5			
		4	31	462	277	1105	459				3			
		5	61	323	898	211	21				93			
		6	9		600	487	84					1042		1000
		7	44	258	643	327	220				26			
		8	186	402	327	142	142							
		9	29	430	860	14	62				26			
		X		549	672	278					87			
	Y	114					163	101	103					
	monomers on unplaced scaffolds		254	733		58	27							
overall number of monomers		1689	5145	6108	3690	1523	101	103	243	1042	316	1000		
assembly accession: GCA_037075425.1	number of monomers on linkage groups	LG	TmSat 01	TmSat 02	TmSat 03	TmSat 04	TmSat 05	TmSat 06	TmSat 07	TmSat 08	TmSat 09	TmSat 10	TmSat 11	
		2	8		314	25	30				68			
		3a	48	301	145	8							29	
		3b	10		161		103							
		4	8		154		47				18	218		
		5	11		601	15	8				90			
		6a	9		63	13	23							
		6b	6	177	186	1	39				2			
		7	14	2	36	22	45				9			
		8	29		312	3					35			
		9a	10		192	20	49							45
		9b	9		87	35	161				6			
	10a	13		104	144			10	4					
	10b	8	180	27										
X	8		157						43					
monomers on unplaced scaffolds		12216	145	47	55	402	164	85	6				7	
overall number of monomers		12407	805	2586	341	907	174	89	277	218	29	52		

monomer arrays on the chromosomes (Fig. 1). TmSat01, TmSat03, and TmSat05 occur as numerous interspersed arrays across the majority of chromosomes (Fig. 1A). TmSat02 and TmSat04 form several large terminal blocks on most chromosomes, often arranged in a way that the arrays of one satDNA directly follow those of the other (Fig. 1B). TmSat08 is organized into several blocks of arrays distributed across seven chromosomes (Fig. 1C). TmSat06 and TmSat07 are confined to the Y chromosome, where their arrays intertwine and are separated by blocks of unrelated sequences (Fig. 1D). Finally, TmSat09, TmSat10, and TmSat11 each form a single large block of consecutive monomers, TmSat09 and TmSat11 on chromosome 6, and TmSat10 on chromosome 1 (Fig. 1E).

The TmSat06 and TmSat07 were each detected in 12 short arrays (Fig. 1D), ranging from 2 to 9 monomers for TmSat06 and 6–11 for TmSat07. Given their organization in a small number of short arrays, we further explored their genomic context by analyzing the sequences flanking the arrays of both satDNAs. For TmSat06, segments preceding all arrays exhibited very high nucleotide sequence similarity (>99%), particularly within the 1,013 bp region directly prior to the arrays. Upstream from this segment, similarity persisted but was interrupted by blocks of insertions or deletions shared among subsets of sequences (Supplementary Figure S2A). Similarly, the regions following TmSat06 arrays showed conserved sequence similarity (>95%) within a 3,370 bp stretch directly downstream of the arrays (Supplementary Figure S2B). The exceptions are 3 extractions: one containing a 430 bp insertion, one showing similarity only over the first 1,000 bp, and one omitted due to a sequencing gap. In the case of TmSat07, the 2,890 bp region preceding the arrays retains sequence similarity across most sequences, interrupted by several indels up to 100 bp in size. Two sequences deviated from this pattern, one due to the insertion of a TmSat05 array (Supplementary Figure S3). Downstream of TmSat07 arrays, a 4,755 bp segment remained conserved, with indel blocks of up to 300 bp. The flanking sequences of TmSat06 and TmSat07 show no mutual similarity and lack detectable homology to any sequences in the NCBI or RepBase databases.

In addition to the TAREAN-obtained consensus-based annotation of the 11 satDNAs, *de novo* detection from the assembly was performed using TideHunter, allowing a comparison between assembly-free and assembly-based approaches for satDNA identification. Using TideHunter, seven satDNAs (TmSat01, TmSat02, TmSat03, TmSat04, TmSat05, TmSat08, and TmSat11) were identified within the main tandem repeat clusters (TRCs) (Supplementary Figure S4A). Some sequences belonging to TmSat08, as well as TmSat09 and TmSat10, were detected in additional clusters (Supplementary Figure

S4B), whereas TmSat06 and TmSat07 were not detected by this approach.

In situ localization of *T. molitor* satDNAs

Fluorescence in situ hybridization (FISH) mapping of satDNAs revealed diverse chromosomal distribution patterns (Fig. 2). TmSat01 encompassed large pericentromeric and interstitial heterochromatic blocks on all chromosomes (Fig. 2A). Probes for TmSat02, 03, 04 and 05 satDNAs produced multiple interspersed signals (Fig. 2B, C, D). TmSat07, 08 also exhibited interspersed pattern of distribution, but with a fewer signals (Fig. 2E, G, H). TmSat06 signals were clustered terminally in two chromosomal pairs in addition to the several weaker and interspersed signals (Fig. 2F). Finally, TmSat09, TmSat10, and TmSat11 each hybridized to a single chromosomal pair, TmSat10 and TmSat11 located terminally (Fig. 2J, K), and TmSat09 positioned centrally on the chromosome (Fig. 2I).

SatDNA divergence

The evolution of satDNAs is shaped by homogenization processes, which reduce sequence divergence, and the accumulation of mutations, which increase divergence. Consequently, satDNA divergence landscapes provide valuable insights into the structure (and evolution) of sequence variants within a given satDNA family. Peaks at low divergence values are generally interpreted as signatures of recent amplification or active homogenization, whereas peaks at higher divergence values likely represent older or non-homogenized variants that have accumulated mutations over time. We therefore analyzed the divergence profiles of the 11 *T. molitor* satDNAs.

Figure 3 shows repeat landscape plots illustrating the divergence of individual monomers extracted from chromosome-level genome assemblies from their consensus sequence. TmSat02 and TmSat11 each exhibit a single dominant peak at 0–10% divergence (Fig. 3B, K). For TmSat09 this peak is shifted towards higher divergence values, 10–20% (Fig. 3I). Divergence plots of TmSat03, TmSat04, and TmSat07 display two distinct peaks: a major peak at very low divergence (up to 10%) and a smaller peak at higher divergence values (Fig. 3C, D, G). Several other satDNAs (TmSat01, TmSat05, TmSat06, TmSat07, TmSat08, TmSat10) exhibit a broad range of sequence divergence, with a dominant peak at low divergence accompanied by multiple smaller peaks at higher divergence levels (Fig. 3A, E, F, H, J).

SatDNAs in extrachromosomal circular DNA (eccDNA)

One of the mechanisms proposed to be included in the spread of satDNA sequences is via extrachromosomal circular DNA (eccDNA). To investigate this, we isolated and screened eccDNA from *T. molitor* for the presence

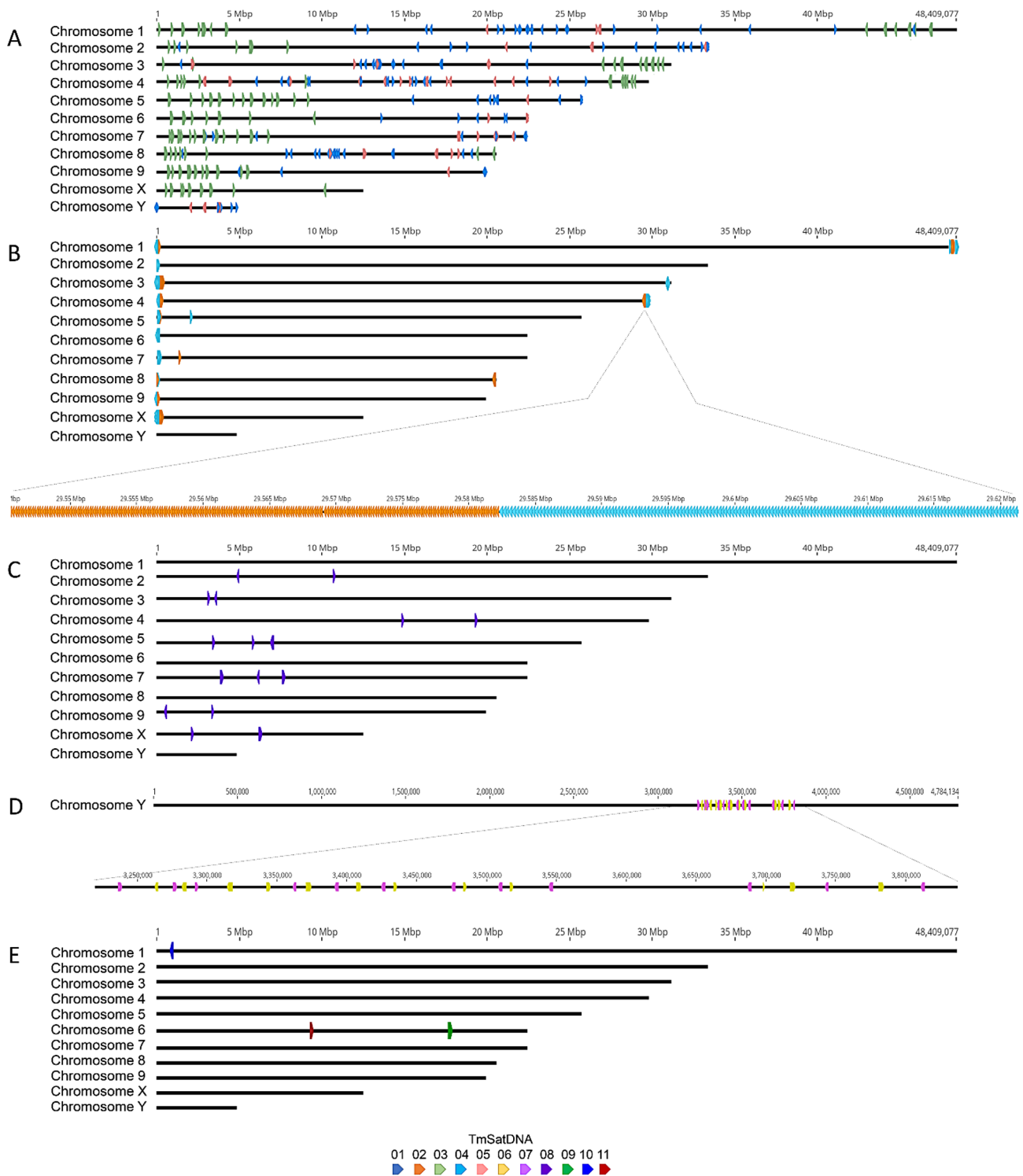


Fig. 1 In silico chromosomal localization of TmSat01 - TmSat11 on the *T. molitor* genome assembly GCF_963966145.1. Monomers of each satDNA are represented by a distinct color, denoted in the legend. Panels correspond to: **A)** TmSat01, TmSat03, and TmSat05. **B)** TmSat02 and TmSat04. **C)** TmSat08. **D)** TmSat06 and TmSat07. **E)** TmSat09, TmSat10, and TmSat11

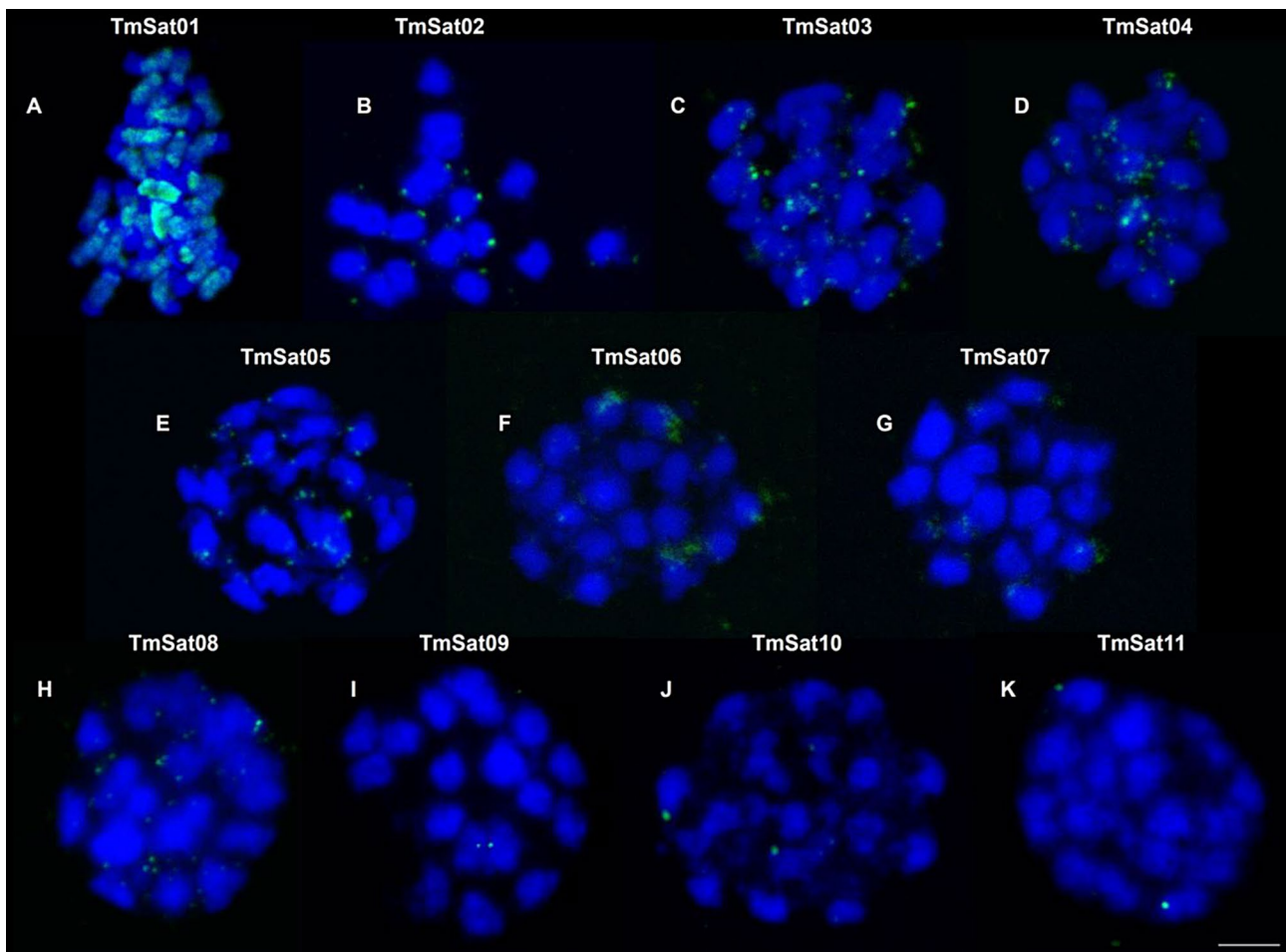


Fig. 2 Fluorescence in situ hybridization localization of TmSat01–TmSat11 on *Tenebrio molitor* chromosomes. Localization of satDNA repeats belonging to (A) TmSat01, (B) TmSat02, (C) TmSat03, (D) TmSat04, (E) TmSat05, (F) TmSat06, (G) TmSat07, (H) TmSat08, (I) TmSat09, (J) TmSat10, and (K) TmSat11. SatDNA signals are shown in green, and chromosomes are counterstained with DAPI (blue). Scale bar represents 3 μm

of the 11 satDNAs. Following total DNA isolation, linear genomic DNA was removed using Exonuclease V (Fig. 4). The absence of linear genomic DNA and the preservation of circular DNA were confirmed using appropriate controls: amplification of H3 (negative control, confirming the absence of linear DNA) and 16S (positive control for circular DNA) is shown in Fig. 4A. The exonuclease digestion result is shown in Fig. 4B. Figure 4C depicts the amplification results from eccDNA fraction. There, 16S amplification was observed, whereas H3 amplification was absent, confirming successful isolation of circular DNA and removal of linear DNA (Fig. 4C). Using this approach, we detected six satDNAs in the eccDNA fraction: TmSat01, TmSat02, TmSat03, TmSat07, TmSat10, and TmSat11, (Fig. 4C).

SatDNA transcription

The transcription of all 11 satDNAs was analyzed using RNA-seq libraries from *T. molitor* eggs, early, mid, and late larvae, early and late pupae, and early and late adults

of both sexes. Our analyses revealed that all 11 satDNAs are transcribed. For the majority, transcription peaked in late male pupae and was present at reduced levels in early male adults (Fig. 5). TmSat06 showed increased transcription in eggs and late larvae (Fig. 5F), while TmSat10 was most active during both early and late larval stages, as well as in late male pupae and late male adults (Fig. 5J). In contrast, TmSat11 was most strongly transcribed during mid-larval stages and late male pupae, as well as in both late female and male adults (Fig. 5K).

SatDNA taxonomic distribution, phylogeny and age

GenBank databases were scanned using the consensus sequences of 11 *T. molitor* satDNAs to gain insight into their taxonomic distribution. In the Supplementary Table 2, insect orders with available genomic data are presented, while for Coleoptera, the corresponding families are also shown. The BLAST search results revealed hits in 47 species across six insect orders: Coleoptera, Lepidoptera, Hymenoptera, Diptera, Hemiptera, and

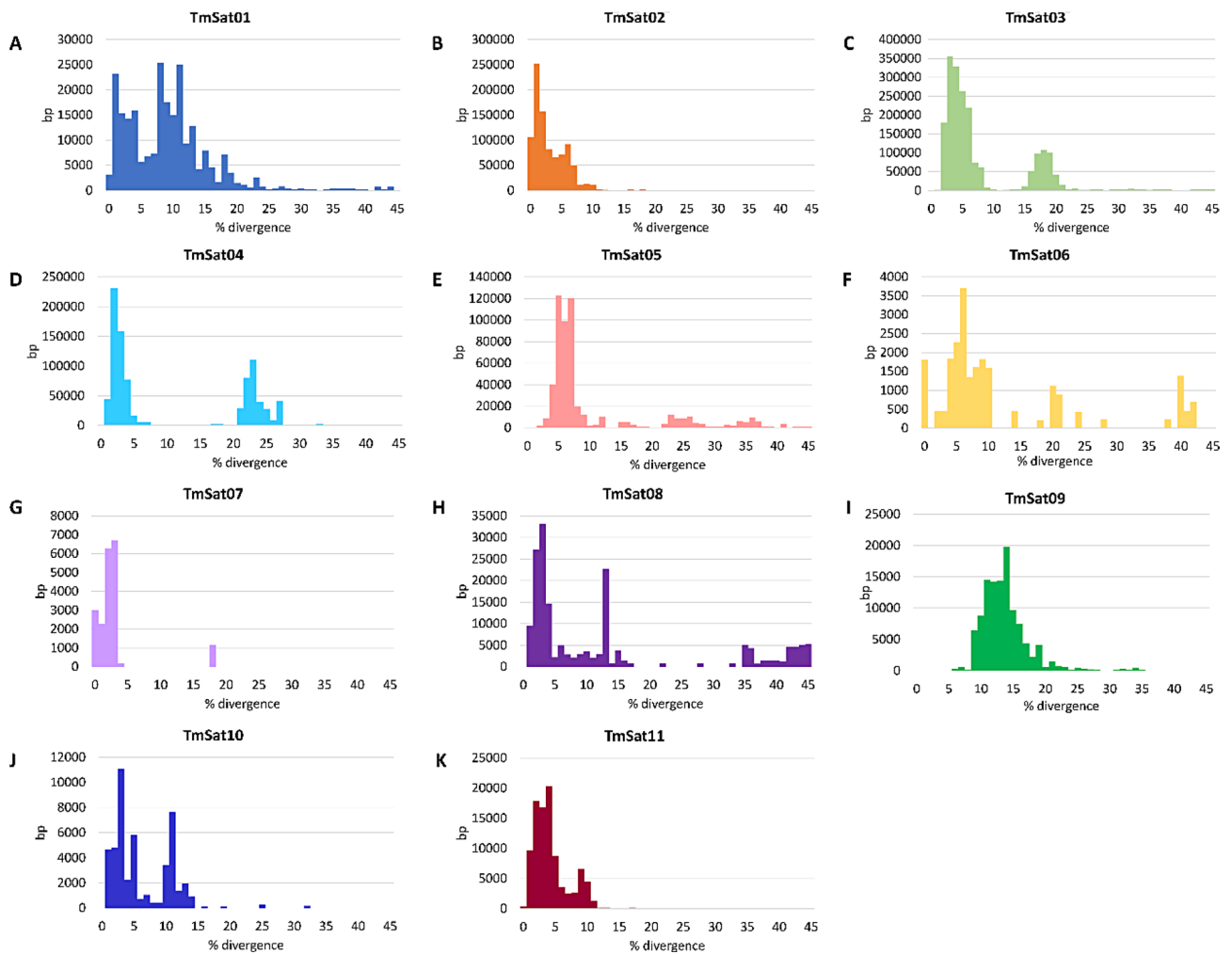


Fig. 3 The distribution of sequence divergence for each satDNA calculated relative to its consensus sequence using the Kimura 2-parameter model. The x-axis represents percentage divergence, while the y-axis indicates the satDNA abundance within each divergence class. Panels correspond to: **A)** TmSat01, **B)** TmSat02, **C)** TmSat03, **D)** TmSat04, **E)** TmSat05, **F)** TmSat06, **G)** TmSat07, **H)** TmSat08, **I)** TmSat09, **J)** TmSat10, **K)** TmSat11. SatDNA-specific colors are preserved in Figs. 1, 3 and 7

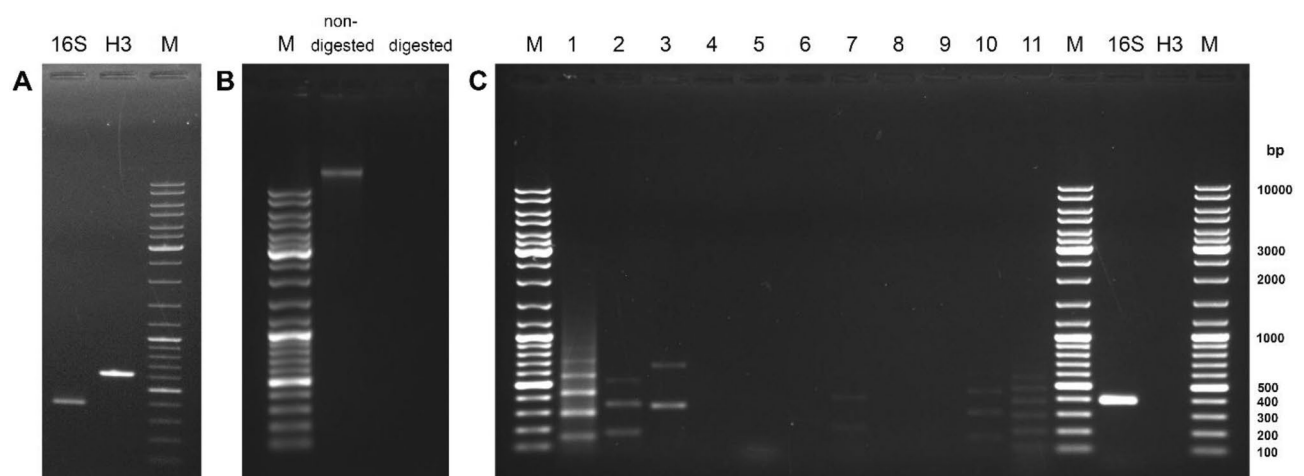


Fig. 4 EccDNA analysis. **A)** agarose gel separation of PCR products for 16S and H3 amplification from complete genomic DNA. **B)** equal amounts of non-digested and Exonuclease V-digested genomic DNA loaded on the gel. **C)** PCR amplification of satDNAs from the eccDNA fraction. Numbers 1–11 correspond to TmSat01 – TmSat11. The 16S lane displays amplification, while in H3 lane amplification was absent. M indicates the molecular size marker

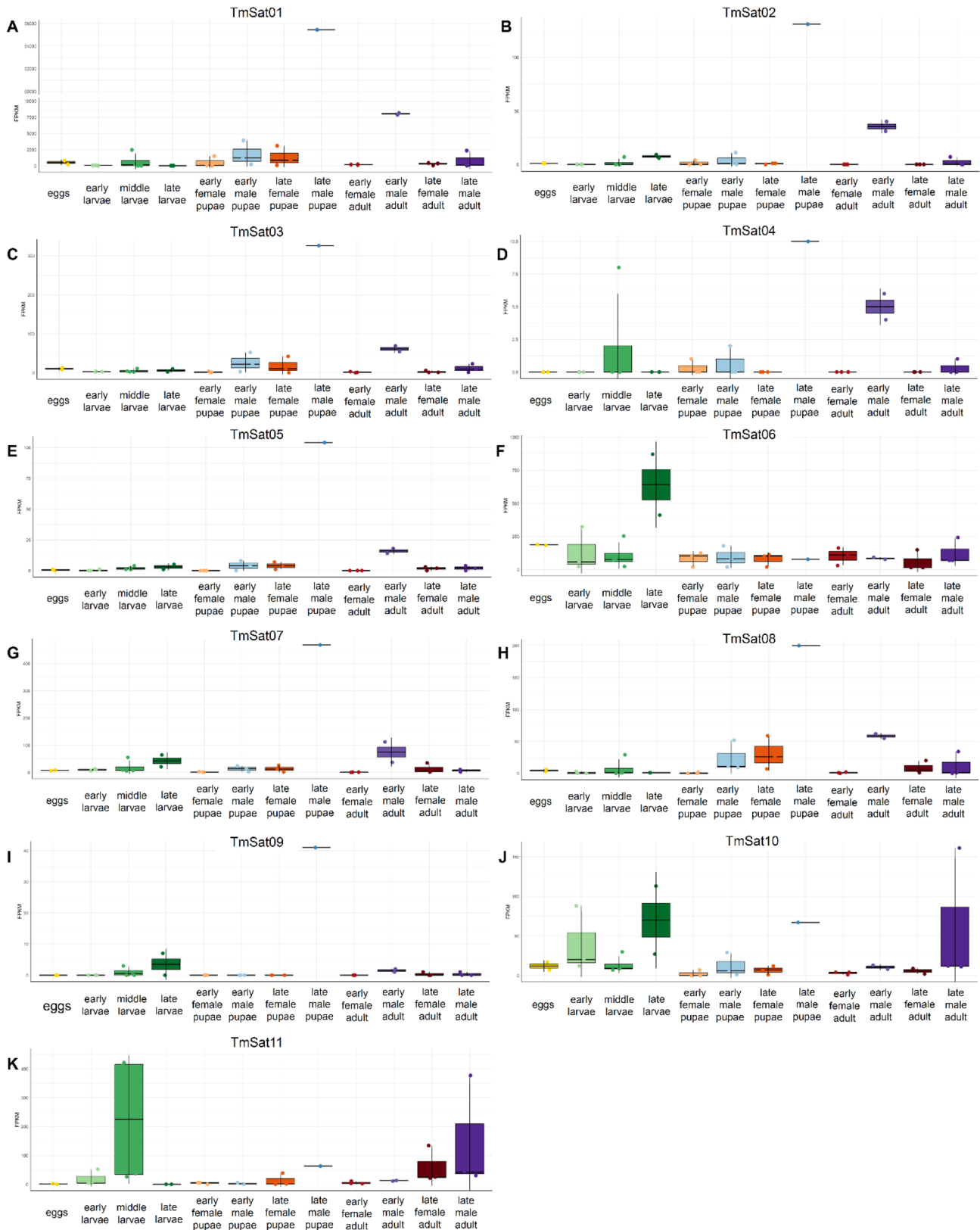


Fig. 5 Transcription dynamics of 11 satDNAs of *T. molitor* across different developmental stages in females and males. Each sample is based on biological replicates shown by dots and error bars indicate the standard deviations. Stage-specific color facilitate comparison across developmental stages. Panels correspond to: **A)** TmSat01, **B)** TmSat02, **C)** TmSat03, **D)** TmSat04, **E)** TmSat05, **F)** TmSat06, **G)** TmSat07, **H)** TmSat08, **I)** TmSat09, **J)** TmSat10, **K)** TmSat11

Blattodea (Supplementary Table S3). Monomers for phylogeny reconstruction were selected and subsampled as described above. The largest number of species and sets of monomers were found with TmSat01 and TmSat10, comprising 320 monomers from 13 species (TmSat01) and 885 monomers from 12 species (TmSat10). TmSat01 sequences did not form distinct clusters and showed no species-specific grouping (Fig. 6A). In contrast, TmSat10 exhibited pronounced clustering, with monomers of each species forming species-specific groups and/or subgroups (Fig. 6B). An exception was a single monomer from *Cataglyphis albicans*, which clustered with those from *T. molitor*. *Tribolium castaneum* and *Tribolium freemani* monomers cluster together, followed by further separation of *T. castaneum* monomers into species-specific group (Fig. 6B). Phylogenetic trees for TmSat02, 03, 04, 05, 08, and 09 were dominated by *T. molitor* monomers, with only limited contributions from *C. albicans*, *C. bombycina*, or *Prionychus ater* (Fig. 6C–F, I, J). In the case of TmSat03 and 09, *T. molitor* monomers form several subgroups. However, monomers from *C. albicans* did not cluster together; instead, they were dispersed among *T. molitor* sequences (Fig. 6D, J). *T. molitor* monomers from TmSat02, 04, 05, and 09 did not form distinct phylogenetic clusters but were observed in numerous clusters and subclusters (Fig. 6C, E, F, J). The phylogenetic tree of TmSat06 included monomers from *Achroia grisella*, *Periplaneta americana*, and *C. albicans* in addition to *T. molitor*. These monomers did not form species-specific groups and were instead scattered throughout the tree, except for those of *P. americana*, which clustered together (Fig. 6G). TmSat07 incorporated an additional species, *Dinoponera quadriceps*, yet similarly lacked species-specific grouping (Fig. 6H). TmSat11 was detected exclusively in *T. molitor* (Supplementary Table S3; Fig. 6K). Approximately half of the monomers selected for phylogenetic analysis present no evident grouping, while the rest form a number of small clusters (Fig. 6K).

We estimated the minimal age of the 11 satDNAs after juxtaposing their distribution against the timetree of insect orders (Fig. 7A). SatDNAs TmSat01, 06, 07 were detected in species belonging to Coleoptera, Lepidoptera, Hymenoptera, and Blattodea. The separation time of these orders dates to ~380 MYA, which would denote the minimal age of these satDNAs. The presence of TmSat04 within Coleoptera, Hymenoptera, Diptera, Hemiptera species suggest a minimal age of ~360 MYA. The divergence of Hymenoptera from Coleoptera, Lepidoptera and Diptera ~340 MYA marks this timepoint as the minimal age of TmSat02, 03, 05, 08, 09, and 10. Finally, TmSat11, which was detected only in *T. molitor* (Coleoptera), represents the youngest satDNA in our dataset, originating sometime after the emergence of the Tenebrionidae family, which is dated to ~165 MYA (Fig. 7B).

Discussion

Sequence analysis, chromosomal mapping, divergence profiles, phylogenetic analyses and age estimation enabled us to delineate the evolutionary origin and diversification of TmSatDNAs while transcription analysis and eccDNA provided insights into functional and distribution potential of these sequences.

Diverse distribution patterns of satDNA sequences

A massive difference in satDNA monomer counts between the two assembly versions has been observed, particularly for TmSat01. The difficulties arise in reconstructing exact sequential order and length of segments composed of highly similar repeat units present in high copy-numbers. Consequently, repetitive sequences are frequently omitted from datasets available in public databases, or even significant errors in genome annotation are introduced [35, 52]. Highly repetitive sequences often suffer from repeat collapse during genome assembly, where multiple copies are erroneously merged into one or certain regions are left as gaps in the final assembly [53]. This indicates that even chromosome-level assemblies may be incomplete and/or misassembled, a problem that is particularly pronounced in species with large genomic blocks composed of highly abundant tandemly repeated satDNA sequences, such as *T. molitor* TmSat01. For example, genome assembly of *T. molitor* is only ~400 Mb while cytological analysis estimated the actual genome size to ~780 Mb [34]. Therefore, relying solely on a (single) assembly can lead to a severe underestimation in the amount of major satDNA, or its positioning, masking their influence on the genome architecture and/or potential functions. The satDNA distribution patterns inspected in situ and in silico are largely consistent for most satellites, with the most notable discrepancies observed for TmSat01, TmSat06, and TmSat07. Although TmSat01 is known to represent the major satDNA of *T. molitor*, forming large pericentromeric blocks [54, 55], Fig. 2A), its in silico representation is significantly reduced, appearing only as numerous short arrays frequently interrupted by assembly gaps. On the other hand, TmSat06 and TmSat07 are detected exclusively on the Y chromosome in the genome assembly, whereas FISH reveals their presence on additional chromosomes (Fig. 2). It is important to highlight that the satDNA research benefits from complementing and juxtaposing in situ and in silico localizations, as their capacity in displaying the localization of long arrays and short satDNA segments are fundamentally opposite [4]. While very short arrays and interspersed monomers are likely to fall below the detection threshold of FISH, they will mostly be successfully included in the genome assembly. Oppositely, long arrays are mostly excluded or reduced in the assembly, while in the FISH analysis they display strong clustered

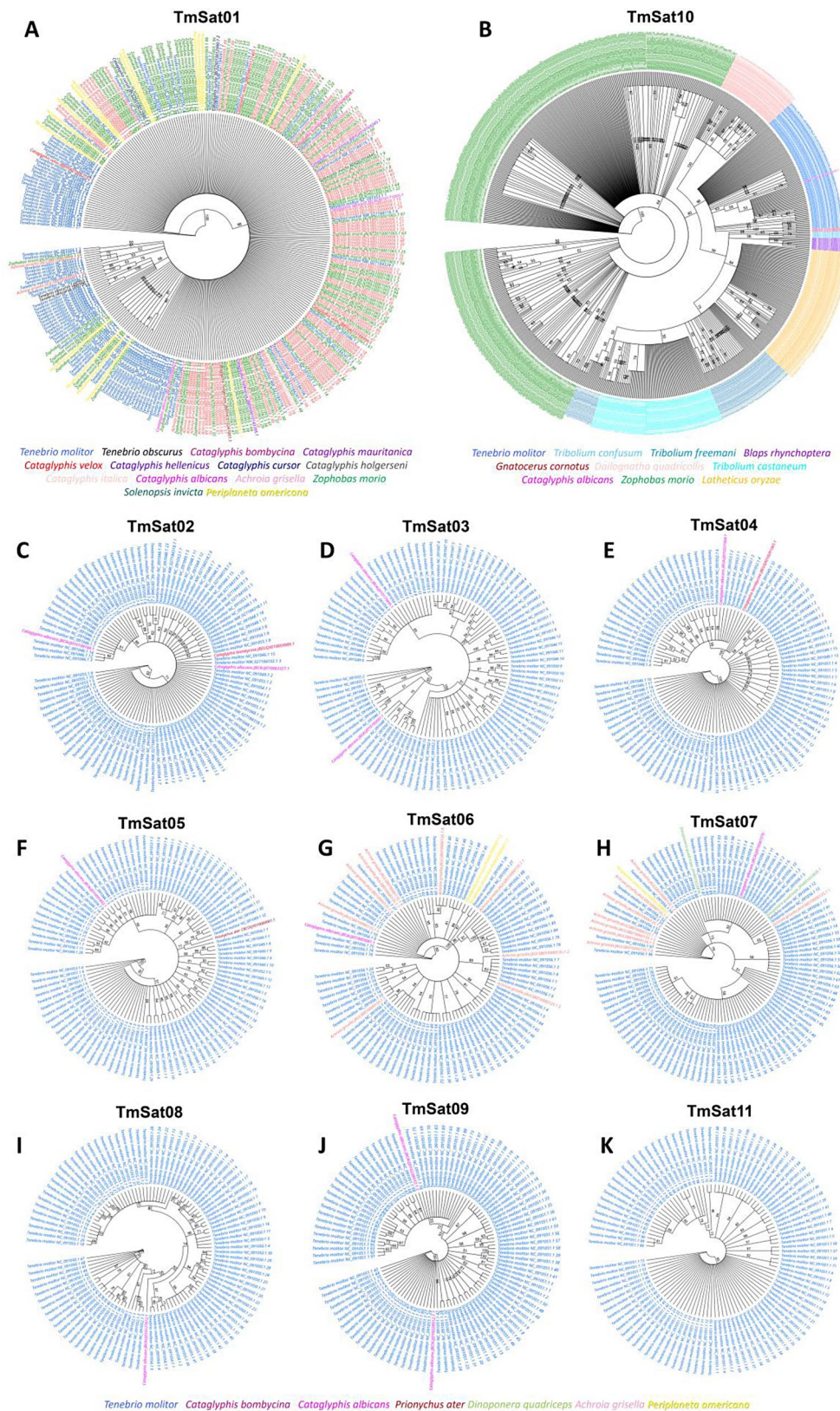


Fig. 6 Phylogenetic reconstruction for 11 satDNAs, based on monomer sequences from insect species. Monomer sequences are color-coded according to their species of origin, denoted in the legend. Panels correspond to: **A)** TmSat01, **B)** TmSat10, **C)** TmSat02, **D)** TmSat03, **E)** TmSat04, **F)** TmSat05, **G)** TmSat06, **H)** TmSat07, **I)** TmSat08, **J)** TmSat09, **K)** TmSat11

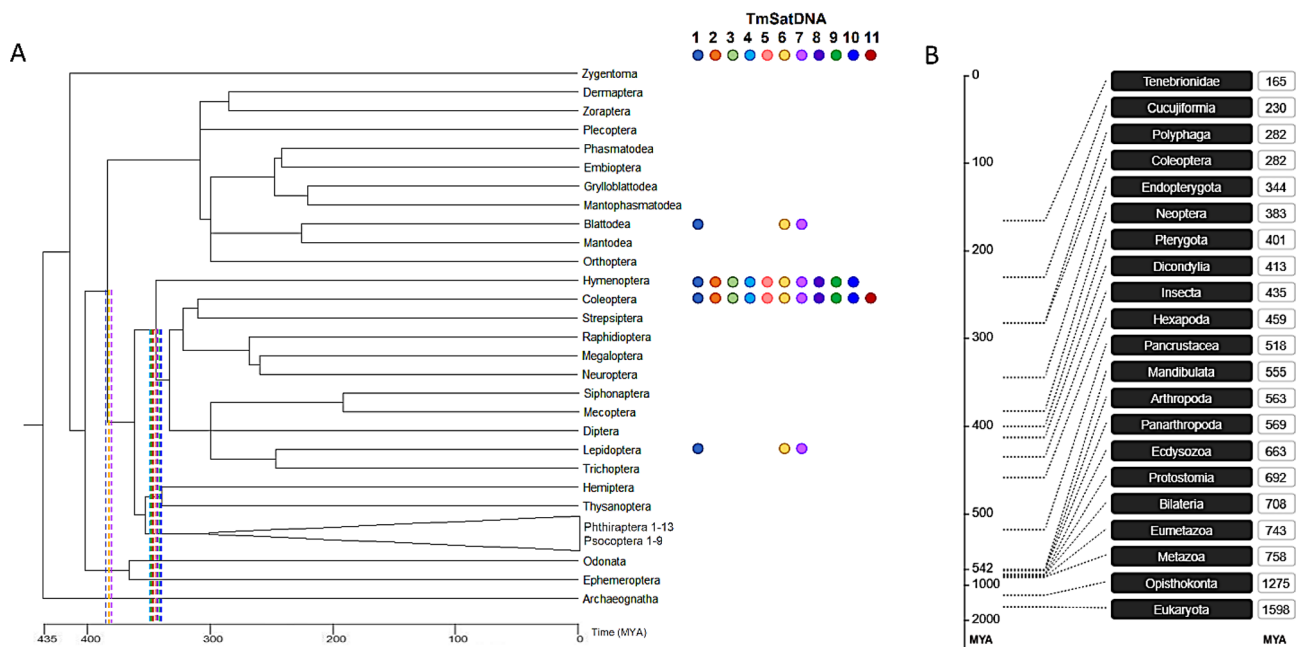


Fig. 7 Taxonomic distribution and age estimation of 11 satDNAs (TmSat01–11). **A)** time-calibrated phylogeny of insect orders with annotated occurrence of *T. molitor* satDNAs. Each satDNA is indicated by a distinct color. Circles denote its presence within a given insect order, while dashed lines represent its estimated minimal evolutionary age. **B)** an evolutionary timeline of family Tenebrionidae

signals [4]. While TmSat01 forms large heterochromatic blocks and is present on all chromosomes of *T. molitor*, TmSat09, 10, and 11 are restricted to only a single chromosome pair (Fig. 2, Table 1). Similar examples of chromosome-specific satDNAs have been reported in other taxa, such as the DBC-150 satellite family in *Drosophila buzzatii* [56], the MCSAT family in *Muscari* species [57], and OocSat42 in the flea beetle *Omophoita octoguttata* [58], and others. Such confinement of a satDNA to a single chromosome may result from several factors, including the relatively recent origin of the satDNA, or structural or functional constraints. In addition, reduced recombination, limited amplification, or low rates of transposition-mediated dispersal may further contribute to a limited spread of a satDNA across the genome. *T. molitor* telomeres are known to be composed of TCAGG telomeric repeats [59], which are omitted from the nucleotide sequence of TmSat02 and TmSat04. Nevertheless, based on their predominantly terminal localization on chromosomes *in situ* and *in silico*, TmSat02 and TmSat04 contribute to the (sub)telomeric regions structurally, and potentially even functionally.

In comparison to the consensus-based annotation, *de novo* detection using TideHunter recovered nine of the eleven known satDNAs (Supplementary Figure S4). This observation reinforces the well-established advantages of assembly-free approaches, particularly when complemented by subsequent annotation in genome assemblies. Conversely, the utility of the assembly-based approach is illustrated by the detection of several repeats

in additional TRCs that are predominantly present in the genome as a single array (Supplementary Figure S4B).

Distinct turnover dynamics and homogenization efficiencies for each *T. molitor* satDNA

The evolution of satDNAs is shaped by homogenization processes, such as unequal crossing-over, gene conversion, and eccDNA-mediated turnover, which act to maintain sequence similarity among monomers within a family [6, 16, 60]. In *T. molitor*, low divergence among certain satDNA families and different patterns in divergence landscapes presented in Figure 3 indicate the action of these processes in shaping their genomic organization and evolutionary dynamics. The divergence profiles of the 11 satDNAs differ from one another, (Fig. 3), indicating the lack of strong homogenization processes that would affect all satDNAs equally. Sequences with low divergence values exhibit minimal deviation from the consensus sequence, reflecting either well preserved copies, or recently amplified/homogenized copies (TmSat02 and TmSat11). High sequence preservation could also be a sign of functional significance of a particular sequence. In contrast, the sequences positioned at higher divergence values represent more degenerate copies with accumulated mutations (TmSat09). Multiple peaks (observed for TmSat01, TmSat05, TmSat06, TmSat07, TmSat08, TmSat10) could be a result of a more intensive propagation of several sequence variants forming subgroups that are homogenous within. Although homogenization is generally expected to be more

effective locally, within a given heterochromatic block, array, or chromosome, this is not necessarily the rule. For example, TmSat02 is distributed across a large number of chromosomes (Table 1, Fig. 2) yet retains high sequence identity (Fig. 3). TmSat09, TmSat10, and TmSat11 are each confined to a single chromosome pair (Table 1, Fig. 2), while exhibiting different divergence patterns (Fig. 3I, J, K). Together, these findings suggest distinct turnover dynamics and homogenization efficiencies for each satDNA in *T. molitor*.

The presence of satDNA in the eccDNA

EccDNA plays a significant role in the dynamics, propagation, and homogenization of satDNA. EccDNAs can originate from chromosomal DNA through several mechanisms, including recombination, DNA repair, and aberrant replication. In the case of satDNA, eccDNAs are frequently produced by intra-strand recombination within tandem arrays [16, 61]. Once formed, eccDNAs may serve as templates for rolling-circle replication, resulting in the amplification of satDNA repeats. This process can facilitate the homogenization of satDNA by enabling the redistribution of identical or nearly identical monomers across the genome.

Traditional detection of satDNAs in the eccDNA fraction relied on using two-dimensional agarose gel electrophoresis followed by Southern hybridization with probes for a specific satDNA (e.g. [15, 61, 62]). Through different and novel approach, based on eccDNA isolation and PCR amplification, we have confirmed the existence of six satDNAs in the eccDNA fraction, TmSat01, 02, 03, 07, 10 and 11 (Fig. 4). While the functional role of eccDNA in satDNA in turnover requires further structural and quantitative verification, we hypothesize some potential scenarios. As each of these satDNAs presents distinct abundance, distribution and chromosomal occupancy, their presence in the eccDNA could be related to different occurrences. For some, eccDNA could be a mean of propagation and distribution across the genome/chromosomes. For the others, it could be merely an instrument for array shortening via excision, or of the array prolonging via circle formation, amplification, and reintegration. TmSat10 and 11 are present in the eccDNA fraction, but still remain limited to one chromosome each, chromosome 1 (TmSat10) and 6 (TmSat11) (Table 1, Fig. 2). This pattern could also indicate that satDNA propagation via eccDNA is not a random process. Successful reintegration of eccDNA-derived satDNA potentially requires a certain “seed” of the corresponding satDNA to be present. Transposition-mediated spread of satDNAs could serve as satDNA seed distribution, enabling subsequent eccDNA-driven expansion. Alternatively, certain satDNAs may be constrained by functional or structural requirements that restrict their presence

to specific chromosomes or genomic regions. In these cases, eccDNA may serve primarily to modulate array length rather than to promote genome-wide dispersal.

Stage- and sex-dependent transcription of *T. molitor* satDNAs

There is increasing evidence that satDNAs undergo active and tightly regulated transcription, and that their transcripts play important roles in a variety of cellular processes, including heterochromatin formation, centromere function, and the regulation of gene expression across various biological contexts (reviewed in [11]). Although transcription of satDNAs in insects has been investigated within various orders [12, 59, 63–69], etc.), comprehensive studies covering all satDNA of the species’ satellitome, spanning at the same time different tissues and/or developmental stages, are still scarce.

Increased transcription for majority of satDNAs in late male pupae and early male adults of *T. molitor* indicates the role of satDNAs in development, with special focus on the male lineage (Fig. 5). For several satDNAs, transcription is also observed at the larval stage, which may likewise reflect contributions from the male lineage, as the sex of those developmental stages could not be distinguished. The majority of satDNAs present similar expression profiles, indicating their simultaneous activation, with a clear association to specific developmental timepoints. De Lima et al. [12] analyzed transcription of satDNAs pBuM (in *Drosophila mojavensis*) and CDSTR198 (in *D. buzzatii*) across embryos, larvae, pupae, and adult males and females, observing the highest levels in pupae and adult males. In *Rhodnius prolixus*, Montiel et al. [13] found that 33/39 satDNAs are transcribed and can be categorized into four expression patterns (ubiquitous, antenna-enriched, gonad-enriched, and gonad-specific), with some satDNAs showing pronounced sex/tissue biases. Our results, as well as other studies indicate that satDNA transcription in insects seems to be widespread and regulated, but varies depending on the satDNA family, tissue, developmental stage, or biological context.

Contrasting phylogenetic distributions of *T. molitor* satDNAs

The clustering patterns of satDNA sequences in the phylogenetic analyses reflect their evolutionary trajectories within and between species. TmSat01 and TmSat10 were detected in a larger number of species compared to other satDNAs, yet they exhibit contrasting phylogenetic clustering patterns. While TmSat01 sequences do not form species-specific groups, TmSat10 shows more pronounced species-specific clustering (Fig. 6). The monomer clustering observed in the TmSat10 phylogeny suggests that these sequences have evolved independently within specific lineages, accumulating mutations

over time and diverging from the ancestral sequence. Similarly, the BIV160 satDNA family, present in various bivalve species, shows a shared origin that subsequently diverged to produce species-specific variants [70]. In contrast, TmSat01 and several other satDNAs lack clear species-specific clustering, which could indicate preservation of the ancestral sequence. Comparable patterns of long-term sequence conservation have been observed in alligators and caimans, where a small number of satDNA families exhibit little variation both within and between species [71]. Alternatively, random distribution of mutations preventing clear separation, or even horizontal transfer events could lead to a mixed satDNA monomer distribution in the phylogenetic tree.

Age of satDNA sequences

For some of the *T. molitor* satDNAs, age estimation was previously performed in Oppert et al. [34]. However, the continuous expansion of NCBI GenBank with newly sequenced insect genomes allowed us to obtain an updated and more comprehensive overview of satDNA species-distribution. TmSat01, 06, and 07 represent the oldest satDNAs in the *T. molitor* genome, with a minimal age of ~380 MYA. TmSat01 shows broad distribution both chromosomally (forming large pericentromeric blocks on all chromosomes of *T. molitor*) and across species, being present in numerous insect genomes (Supplementary Table S3). In contrast, TmSat06 and TmSat07 are restricted to a limited number of chromosomes and are found in only a few currently available insect genomes. TmSat11 is detected exclusively in *T. molitor*, suggesting that it emerged after the speciation of this species. This estimates its age to less than 165 MYA, corresponding to the proposed age of the Tenebrionoidea [72]. Simultaneously, this enables the potential usage of TmSat11 sequence as an alternative marker for species confirmation, particularly relevant in the context of the integration of *T. molitor* into the animal feed and human food market.

The library hypothesis posits that closely related species share a common set of satDNAs inherited from a common ancestor [73]. However, novel satDNAs can arise independently in each species, while existing ones may undergo degeneration, resulting in partially overlapping yet divergent satDNA content across species. This partially explains the variable presence or absence of *T. molitor* satDNAs in different insect species/orders (Fig. 7). At the same time, despite the growing number of available insect genomes (Supplementary Table S2), *T. molitor* satDNAs were detected in only a limited number of species (Supplementary Table S3). The absence of detection does not imply true absence, but rather reflects current database limitations and/or assembly challenges, particularly for long tandem arrays

It should be noted that satDNA repeats are often integral components of transposable elements [74–78]. In line with this, the conserved regions flanking the TmSat06 and TmSat07 arrays, although currently unclassified, may represent segments of transposable elements. Additionally, eccDNAs have been shown to act as vehicles for the transport of genetic material [79, 80]. In this context, the patchy phylogenetic distribution of these satDNAs, combined with higher sequence similarity than expected from vertical inheritance alone (Supplementary Table S3, Fig. 7), could suggest even the occurrence of horizontal transfer events, potentially mediated by transposable elements and/or eccDNAs.

Conclusion

This study provides an integrative evolutionary framework for understanding satDNA dynamics in the yellow mealworm *Tenebrio molitor*. By combining cytogenetics, comparative genomics, phylogenetics, divergence profiling, eccDNA screening, and transcriptional analyses, we demonstrate that individual satDNA families follow distinct evolutionary trajectories shaped by differential homogenization, turnover rates, chromosomal context, and lineage history. Our results reveal that satDNAs range from deeply conserved families spanning multiple insect orders to lineage-specific satDNAs restricted to *T. molitor*, reflecting both ancient origins and recent emergence. The detection of multiple satDNAs in eccDNA fractions supports a role for eccDNA in satDNA turnover and genome (re)shaping, while coordinated, stage-specific transcription highlights their potential functional relevance during development. Together, these findings emphasize that satDNAs are dynamic, functionally relevant components of eukaryotic genomes and underscore the importance of such comprehensive, multi-level analyses for elucidating their evolutionary significance.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13062-026-00808-1>.

Supplementary material 1

Acknowledgements

We thank Lucija Horvat for her support in confocal microscopy.

Author contributions

Conceptualization: E.Š.-V., M.P., Đ.U.; Experimental work: P.M., E.Š.-V.; Bioinformatic analyses: P.M., A.S., E.Š.-V.; Data Curation: P.M., A.S., Đ.U., B.O., M.P., E.Š.-V.; Visualization: P.M., A.S., E.Š.-V.; Resources: E.Š.-V., B.O.; Writing - Original Draft: E.Š.-V.; Writing - Review & Editing: P.M., A.S., Đ.U., B.O., M.P., E.Š.-V.; Supervision: E.Š.-V.; Project Administration: E.Š.-V.; Funding Acquisition: E.Š.-V.

Funding

This work was supported by the Adris foundation grant.

Data availability

The data underlying this article are available in the article and in its online supplementary material. Consensus sequences of the 11 *T. molitor* satDNAs are deposited in NCBI under accession numbers PX395431–PX395441.

Declarations

Ethics declaration

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Received: 10 February 2026 / Accepted: 16 April 2026

Published online: 08 May 2026

References

- López-Flores I, Garrido-Ramos MA. The repetitive DNA content of eukaryotic genomes. *Genome Dyn*. 2012;1–28.
- Biscotti MA, Olmo E, Heslop-Harrison JS. Repetitive DNA in eukaryotic genomes. *Chromosome Res*. 2015;23(3):415–20. <https://doi.org/10.1007/s10077-015-9499-z>.
- Garrido-Ramos MA. Satellite DNA: an evolving topic. *Genes (Basel)*. 2017;8(9):230–41. <https://doi.org/10.3390/genes8090230>.
- Garrido-Ramos MA, Plohl M, Šatović-Vukšić E. Satellite DNA genomics: the ongoing story. *IJMS*. 2025;26(23):1–23. <https://doi.org/10.3390/ijms262311291>.
- Hartley G, O'Neill R. Centromere repeats: hidden gems of the genome. *Genes (Basel)*. 2019;10(3):223. <https://doi.org/10.3390/genes10030223>.
- Plohl M, Meštrović N, Mravinac B. Satellite DNA evolution. *Genome Dyn*. 2012;126–52.
- Thakur J, Packiaraj J, Henikoff S. Sequence, chromatin and evolution of satellite DNA. *IJMS*. 2021;22(9):4309. <https://doi.org/10.3390/ijms22094309>.
- Ugarković Đ, Plohl M. Variation in satellite DNA profiles—causes and effects. *Embo J*. 2002;21(22):5955–59. <https://doi.org/10.1093/emboj/cdf612>.
- Šatović-Vukšić E, Plohl M. Satellite DNAs—from localized to highly dispersed genome components. *Genes (Basel)*. 2023;14(3):742. <https://doi.org/10.3390/genes14030742>.
- Šatović-Vukšić E, Majcen P, Plohl M. Satellite DNAs rising from the transposon graveyards. *DNA Res*. 2025;32(5):dsaf026. <https://doi.org/10.1093/dnares/dsaf026>.
- Ugarković Đ, Sermek A, Ljubić S, Felicello I. Satellite DNAs in Health and Disease. *Genes (Basel)* 13. 2022;13(7):1154. <https://doi.org/10.3390/genes13071154>.
- de Lima LG, Svartman M, Kuhn GCS. Dissecting the satellite DNA landscape in three cactophilic *Drosophila* sequenced genomes. *G3 Genes, Genomes, Genet*. 2017;7(8):2831–43. <https://doi.org/10.1534/g3.117.042093>.
- Montiel EE, Panzera F, Palomeque T, Lorite P, Pita S. Satellitome analysis of *Rhodnius prolixus*, one of the main Chagas disease vector species. *IJMS*. 2021;22(11):6052. <https://doi.org/10.3390/ijms22116052>.
- Dover G. Molecular drive. *Trends Genet*. 2002;18(11):587–89. [https://doi.org/10.1016/S0168-9525\(02\)02789-0](https://doi.org/10.1016/S0168-9525(02)02789-0).
- Cohen Z, Bacharach E, Lavi S. Mouse major satellite DNA is prone to eccDNA formation via DNA ligase IV-dependent pathway. *Oncogene*. 2006;25(33):4515–24. <https://doi.org/10.1038/sj.onc.1209485>.
- Cohen S, Agmon N, Sobol O, Segal D. Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. *Mob DNA*. 2010;1(1):1–11. <https://doi.org/10.1186/1759-8753-1-11>.
- Adamski Z, Bufo SA, Chowański S, Falabella P, Lubawy J, Marciniak P, et al. Beetles as model organisms in physiological, biomedical and environmental studies – a review. *Front. Physiol*. 2019;10:1–22. <https://doi.org/10.3389/fphys.2019.00319>.
- Dobson AJ, Johnston PR, Vilcinskas A, Rolff J. Identification of immunological expressed sequence tags in the mealworm beetle *Tenebrio molitor*. *J Insect Physiol*. 2012;58(12):1556–61. <https://doi.org/10.1016/J.JINSPHYS.2012.09.009>.
- Goptar IA, Shagin DA, Shagina IA, Mudrik ES, Smirnova YA, Zhuzhikov DP, et al. A digestive prolyl carboxypeptidase in *Tenebrio molitor* larvae. *Insect Biochem Mol Biol*. 2013;43(6):501–09. <https://doi.org/10.1016/J.IJMB.2013.02.009>.
- Jeong JE, Kang SW, Hwang HJ, Chae S-H, Patnaik BB, Han YS, et al. Expressed sequence tags (ESTs) analysis of *Tenebrio molitor* larvae. *Entomological Res*. 2013;43(3):168–76. <https://doi.org/10.1111/1748-5967.12019>.
- Ohde T, Yaginuma T, Niimi T. Insect morphological diversification through the modification of wing serial homologs. *Science*. 2013;340(6131):495–98. <https://doi.org/10.1126/science.1234219>.
- Prokkola J, Roff D, Kärkkäinen T, Krams I, Rantala MJ. Genetic and phenotypic relationships between immune defense, melanism and life-history traits at different temperatures and sexes in *Tenebrio molitor*. *Heredity (Edinb)*. 2013;111(2):89–96. <https://doi.org/10.1038/hdy.2013.20>.
- Tereshchenkova VF, Goptar IA, Zhuzhikov DP, Belozersky MA, Dunaevsky YE, Oppert B, et al. Prolidase is a critical enzyme for complete gliadin digestion in *Tenebrio molitor* larvae. *Arch Insect Biochem Physiol*. 2017;95(4):1–12. <https://doi.org/10.1002/arch.21395>.
- Zhu JY, Yang P, Zhang Z, Wu G-X, Yang B. Transcriptomic immune response of *Tenebrio molitor* pupae to parasitization by *Scleroderma guani*. *PLoS One*. 2013;8(1):1–15. <https://doi.org/10.1371/journal.pone.0054411>.
- Ding Q, Wu RA, Shi T, Yu Y, Yan Y, Sun N, et al. Antiproliferative effects of mealworm larvae (*Tenebrio molitor*) aqueous extract on human colorectal adenocarcinoma (caco-2) and hepatocellular carcinoma (HepG2) cancer cell lines. *J Food Biochem*. 2021;45(7):e13778. <https://doi.org/10.1111/jfbc.13778>.
- Hwang D, Lee SH, Goo T-W, Yun Y. Potential of antimicrobial peptide-overexpressed *Tenebrio molitor* larvae extract as a natural preservative for Korean traditional sauces. *Insects*. 2022;13(4):381. <https://doi.org/10.3390/insects13040381>.
- Klost M, Ramirez-Huerta MI, Drusch S. Food hydrocolloids for health heat-induced gelation of protein from mealworm (*Tenebrio molitor*): influence of pH and zinc concentration. *Food Hydrocoll Health*. 2022;2:100105. <https://doi.org/10.1016/j.fhfh.2022.100105>.
- Song DH, Kim M, Jin E-S, Sim D-W, Won H-S, Kim EK, et al. Cryoprotective effect of an antifreeze protein purified from *Tenebrio molitor* larvae on vegetables. *Food Hydrocolloids*. 2019;94:585–91. <https://doi.org/10.1016/j.foodhyd.2019.04.007>.
- Morales-Ramos JA, Rojas MG, Shelby KS, Coudron TA. Nutritional value of pupae versus larvae of *Tenebrio molitor* (Coleoptera: Tenebrionidae) as food for rearing podisus maculiventris (Heteroptera: Pentatomidae). *J Econ Entomol*. 2016;109(2):564–71. <https://doi.org/10.1093/jee/tov338>.
- Costa S, Pedro S, Lourenço H, et al. Evaluation of *Tenebrio molitor* larvae as an alternative food source. *NFS J* 2020;21:57–64. <https://doi.org/10.1016/j.nfs.2020.10.001>.
- Leierer D, Olmstead M, Oppert B. Sequencing to identify pathogens in *Tenebrio molitor*: implications in insects farmed for food and feed. *Front. Insect Sci*. 2023;3:1–7. <https://doi.org/10.3389/finsc.2023.1059046>.
- Eleftheriou E, Aury J-M, Vacherie B, Istace B, Belser C, Noel B, et al. Suppl chromosome-scale assembly of the yellow mealworm genome. *Open Res Eur*. 2021;1:94. <https://doi.org/10.12688/openreseurope.13987.1>.
- Kaur S, Stinson SA, DiCenzo GC. Whole genome assemblies of *Zophobas morio* and *Tenebrio molitor*. *G3 (Bethesda)* 13: G3: Genes, Genomes, Genet. 2023;13(6):1–11. <https://doi.org/10.1093/g3journal/jkad079>.
- Oppert B, Dossey AT, Chu F-C, Šatović-Vukšić E, Plohl M, Smith TPL, et al. The genome of the yellow mealworm, *Tenebrio molitor*: It's bigger than you think. *Genes (Basel)*. 2023;14(12):2209. <https://doi.org/10.3390/genes14122209>.
- Sedlazeck FJ, Lee H, Darby CA, Schatz MC. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat Rev Genet*. 2018;19(6):329–46. <https://doi.org/10.1038/s41576-018-0003-4>.
- Davis CA, Wyatt GR. Distribution and sequence homogeneity of an abundant satellite DNA in the beetle, *Tenebrio molitor*. *Nucl Acids Res*. 1989;17(14):5579–86. <https://doi.org/10.1093/nar/17.14.5579>.
- Juan C, Gosálvez J, Petitpierre E. Improving beetle karyotype analysis: restriction endonuclease banding of *Tenebrio molitor* chromosomes. *Heredity (Edinb)*. 1990;65(2):157–62. <https://doi.org/10.1038/hdy.1990.83>.
- Ugarković Đ, Plohl M, Gamulin V. Sequence variability of satellite DNA from the mealworm *Tenebrio molitor*. *Gene*. 1989;83(1):181–83. [https://doi.org/10.1016/0378-1119\(89\)90417-4](https://doi.org/10.1016/0378-1119(89)90417-4).
- Gao Y, Liu B, Wang Y, Xing Y. TideHunter: efficient and sensitive tandem repeat detection from noisy long-reads using seed-and-chain. *Bioinformatics*. 2019;35(14):i200–07. <https://doi.org/10.1093/bioinformatics/btz376>.
- Novák P, Robledillo LÁ, Koblížková A, Vrbová I, Neumann P, Macas J. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res*. 2017;45(12):e111. <https://doi.org/10.1093/nar/gkx257>.

41. Bhattacharya AK, Ameel JJ, Waldbauer GP. A method for sexing living pupal and adult yellow mealworms. *Ann Entomol Soc Am.* 1970;63(6):1783. <https://doi.org/10.1093/aesa/63.6.1783>.
42. Pérez-García C, Morán P, Pasantes JJ. Cytogenetic characterization of the invasive mussel species *Xenostrobus securis* Lmk. (Bivalvia: mytilidae). *Genome.* 2011;54(9):771–78. <https://doi.org/10.1139/g11-040>.
43. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37(8):907–15. <https://doi.org/10.1038/s41587-019-0201-4>.
44. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–30. <https://doi.org/10.1093/bioinformatics/btt656>.
45. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.
46. Hardin J, Rocke DM. Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator. *Comput Stat Data Anal.* 2004;44(4):625–38. [https://doi.org/10.1016/S0167-9473\(02\)00280-3](https://doi.org/10.1016/S0167-9473(02)00280-3).
47. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012;9(4):357–59. <https://doi.org/10.1038/nmeth.1923>.
48. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–79. <https://doi.org/10.1093/bioinformatics/btp352>.
49. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5(7):621–28. <https://doi.org/10.1038/NMETH.1226>.
50. Wickham H. *Elegant graphics for data analysis.* 2nd edn. Springer Cham; 2016) ggplot2.
51. Kumar S, Suleski M, Craig JM, Kasprowitz AE, Sanderford M, Li M, et al. Time-Tree 5: an Expanded Resource for species divergence times. *Mol Biol Evol.* 2022;39(8):1–6. <https://doi.org/10.1093/molbev/msac174>.
52. Šatović E, Tunjić-Cvitančić M, Plohl M. Tools and databases for solving problems in detection and identification of repetitive DNA sequences. *Period Biol.* 2020;121-122(1–2):7–14. <https://doi.org/10.18054/pb.v121-122i1-2.10571>.
53. De BA, Cuadrado A, Jouve N. Sequencing of long stretches of repetitive DNA. *Sci Rep.* 2016;6(1):36665. <https://doi.org/10.1038/srep36665>.
54. Juan C, Pons J, Petitpierre E. Localization of tandemly repeated DNA sequences in beetle chromosomes by fluorescent *in situ* hybridization. *Chromosome Res.* 1993;1(3):167–174. <https://doi.org/10.1007/BF00710770>.
55. Mora P, Rico-Porras JM, Palomeque T, Valdivia A, Cabral-DE-Mello DC, Lorite P. Cytogenetic studies on three tenebrionid beetles, *Tenebrio molitor*, *Alphitobius diaperinus* and *Zophobas morio* (Coleoptera: Tenebrionidae): an overview and new data. *Eur J Entomol.* 2024;121:206–18. <https://doi.org/10.14411/eje.2024.023>.
56. Kuhn GCS, Franco FF, Manfrin MH, Moreira-Filho O, Sene FM. Low rates of homogenization of the DBC-150 satellite DNA family restricted to a single pair of microchromosomes in species from the *Drosophila buzzatii* cluster. *Chromosome Res.* 2007;15(4):457–70. <https://doi.org/10.1007/s10577-007-1138-x>.
57. De La Herrán R, Robles F, Cuñado N et al. A heterochromatic satellite DNA is highly amplified in a single chromosome of *Muscari* (Hyacinthaceae). *Chromosoma.* 2001;110(3):197–202. <https://doi.org/10.1007/s004120000115>.
58. Vidal JAD, Charlesworth D, Utsunomia R, Garrido-Ramos MA, dos Santos RZ, Porto-Foresti F, et al. Unraveling the role of satellite DNAs in the evolution of the giant XY sex chromosomes of the flea beetle *Omophoita octoguttata* (Coleoptera, chrysomelidae). *BMC Biol.* 2025;23(1). <https://doi.org/10.1186/s12915-025-02155-5>.
59. Mora P, Vela J, Ruiz-Mena A, Palomeque T, Lorite P. Characterization and transcriptional analysis of a subtelomeric satellite DNA family in the ladybird beetle *Henosepilachna argus* (Coleoptera: Coccinellidae). *Eur J Entomol.* 2017;114:481–87. <https://doi.org/10.14411/eje.2017.061>.
60. Dover G. Molecular drive: a cohesive mode of species evolution. *Nature.* 1982;299(5879):111–117. <https://doi.org/10.1038/299111a0>.
61. Navrátilová A, Koblížková A, Macas J. Survey of extrachromosomal circular DNA derived from plant satellite repeats. *BMC Plant Biol.* 2008;8(1):1–13. <https://doi.org/10.1186/1471-2229-8-90>.
62. Cohen S, Yacobi K, Segal D. Extrachromosomal circular DNA of tandemly repeated genomic sequences in *Drosophila*. *Genome Res.* 2003;13(6a):1133–45. <https://doi.org/10.1101/gr.907603>.
63. Feliciello I, Akrap I, Ugarković Đ. Correction: satellite DNA modulates gene expression in the beetle *Tribolium castaneum* after Heat stress. *PLoS Genet.* 2015;11(9):1–18. <https://doi.org/10.1371/journal.pgen.1005547>.
64. Lorite P, Renault S, Rouleux-Bonnin F, Bigot S, Periquet G, Palomeque T. Genomic organization and transcription of satellite DNA in the ant *Aphaenogaster subterranea* (Hymenoptera, Formicidae). *Genome.* 2002;45(4):609–16. <https://doi.org/10.1139/g02-022>.
65. Pezer Z, Brajković J, Feliciello I, Ugarković Đ. Transcription of satellite DNAs in insects. *Prog Mol Subcell Biol.* 2011;161–78.
66. Renault S, Rouleux-Bonnin F, Periquet G, Bigot Y. Satellite DNA transcription in *Diadromus pulchellus* (Hymenoptera). *Insect Biochem Mol Biol.* 1999;29(2):103–11. [https://doi.org/10.1016/S0965-1748\(98\)00113-1](https://doi.org/10.1016/S0965-1748(98)00113-1).
67. Rouleux-Bonnin F, Renault S, Bigot Y, Periquet G. Transcription of four satellite DNA subfamilies in *Diprion pini* (Hymenoptera, Symphyta, Diprionidae). *Eur J Biochem.* 1996;238(3):752–59. <https://doi.org/10.1111/j.1432-1033.1996.0752.w.x>.
68. Sermek A, Feliciello I, Ugarković Đ. Distinct regulation of the expression of satellite DNAs in the beetle *Tribolium castaneum*. *IJMS.* 2020;22(1):1–16. <https://doi.org/10.3390/ijms22010296>.
69. Věchtová P, Dalíková M, Sýkorová M, Žurovcová M, Füssy Z, Zrzavá M. CpSAT-1, a transcribed satellite sequence from the codling moth, *Cydia pomonella*. *Genetica.* 2016;144(4):385–95. <https://doi.org/10.1007/s10709-016-9907-0>.
70. Plohl M, Petrović V, Luchetti A, Ricci A, Šatović E, Passamonti M, et al. Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusks. *Heredity (Edinb).* 2010;104(6):543–51. <https://doi.org/10.1038/hdy.2009.141>.
71. Sales-Oliveira VC, Santos RZ, Goes CAG, Calegari RM, Garrido-Ramos MA, Altmanová M, et al. Evolution of ancient satellite DNAs in extant alligators and caimans (Crocodylia, Reptilia). *BMC Biol.* 2024;22(1):1–16. <https://doi.org/10.1186/s12915-024-01847-8>.
72. Wang B, Zhang H. The oldest Tenebrionidae (Coleoptera) from the middle Jurassic of China. *J Paleontol.* 2011;85(2):266–70. <https://doi.org/10.1666/09-088.1>.
73. Fry K, Salser W. Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell.* 1977;12(4):1069–84. [https://doi.org/10.1016/0092-8674\(77\)90170-2](https://doi.org/10.1016/0092-8674(77)90170-2).
74. Scalvenzi T, Pollet N. Insights on genome size evolution from a miniature inverted repeat transposon driving a satellite DNA. *Mol Phylogenet Evol.* 2014;81:1–9. <https://doi.org/10.1016/j.jympev.2014.08.014>.
75. Šatović E, Plohl M. Distribution of DTH53 satellite DNA across 12 bivalve species. *J Genet.* 2018;97(2):575–80. <https://doi.org/10.1007/s12041-018-0940-x>.
76. Šatović E, Plohl M. Tandem repeat-containing MITES in the clam *Donax trunculus*. *Genome Biol Evol.* 2013;5(12):2549–59. <https://doi.org/10.1093/gbe/evt202>.
77. Tunjić-Cvitančić M, García-Souto D, Pasantes JJ, Šatović-Vukšić E. Dominance of transposable element-related satDNAs results in great complexity of “satDNA library” and invokes the extension towards “repetitive DNA library”. *Mar Life Sci Technol.* 2024;6(2):236–51. <https://doi.org/10.1007/s42995-024-00218-0>.
78. Tunjić-Cvitančić M, Pasantes JJ, García-Souto D, Cvitančić T, Plohl M, Šatović-Vukšić E. Satellitome analysis of the Pacific Oyster *Crassostrea gigas* reveals new pattern of satellite DNA organization, highly scattered across the genome. *IJMS.* 2021;22(13):6798. <https://doi.org/10.3390/ijms22136798>.
79. Pereira L, Dunning LT. Extrachromosomal circular DNA as a vehicle to gene transfer in plants news and views. *Plant Physiol.* 2023;193(1):172–73. <https://doi.org/10.1093/plphys/kiad380>.
80. Zhang A, Wang T, Yuan L, Shen Y, Liu K, Liu B, et al. Horizontal transfer of plasmid-like extrachromosomal circular DNAs across graft junctions in Solanaceae. *Mol Horticulture.* 2024;4(1):1–21. <https://doi.org/10.1186/s43897-024-00124-0>.

Publisher's Note

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