



Article Satellitome Analysis on *Talpa aquitania* Genome and Inferences about the satDNAs Evolution on Some Talpidae

Juana Gutiérrez¹, Gaël Aleix-Mata¹, Eugenia E. Montiel¹, Diogo C. Cabral-de-Mello^{1,2}, Juan Alberto Marchal¹ and Antonio Sánchez^{1,*}

- ¹ Departamento de Biología Experimental, Área de Genética, Universidad de Jaén, Paraje de las Lagunillas s/n, 23071 Jaén, Spain
- ² Departamento de Biologia Geral e Aplicada, Instituto de Biociências/IB, UNESP—Universidade Estadual Paulista, Rio Claro, São Paulo 13506-900, Brazil
- * Correspondence: abaca@ujaen.es

Abstract: In the genus Talpa a new species, named Talpa aquitania, has been recently described. Only cytogenetic data are available for the nuclear genome of this species. In this work, we characterize the satellitome of the T. aquitania genome that presents 16 different families, including telomeric sequences, and they represent 1.24% of the genome. The first satellite DNA family (TaquSat1-183) represents 0.558%, and six more abundant families, including TaquSat1-183, comprise 1.13%, while the remaining 11 sat-DNAs represent only 0.11%. The average A + T content of the SatDNA families was 50.43% and the median monomer length was 289.24 bp. The analysis of these SatDNAs indicated that they have different grades of clusterization, homogenization, and degeneration. Most of the satDNA families are present in the genomes of the other Talpa species analyzed, while in the genomes of other more distant species of Talpidae, only some of them are present, in accordance with the library hypothesis. Moreover, chromosomal localization by FISH revealed that some satDNAs are localized preferentially on centromeric and non-centromeric heterochromatin in T. aquitania and also in the sister species T. occidentalis karyotype. The differences observed between T. aquitania and the close relative T. occidentalis and T. europaea suggested that the satellitome is a very dynamic component of the genomes and that the satDNAs could be responsible for chromosomal differences between the species. Finally, in a broad context, these data contribute to the understanding of the evolution of satellitomes on mammals.

Keywords: insectivora; satellite DNAs; satellitome; genus Talpa; Talpidae

1. Introduction

Eukaryotic genomes are divided into two regions, one called euchromatin, where most genes are located [1] and heterochromatin, composed of a large number of repetitive DNA sequences. The repetitive sequences possess high variability in nucleotide base pair composition and number of copies, being involved in the genome size variation and diversity across three of life [2]. As a whole, the repeated DNA content (repeatome) includes dispersed, represented by all classes of transposable elements (TEs) and tandem satellite DNA (satDNA) repeats [3–6]. The satDNA sequences content of a genome, or satellitome [7], are constituted by non-coding repeated sequences organized in tandem, mainly located in constitutive heterochromatin regions, that in multiple species are observed on pericentromeric and centromeric regions of chromosomes and on a lesser extent on telomeres [2,8,9]. Satellite DNAs could represent a large proportion of the whole eukaryotic genomes, with variation between 0.1% and up to 50% (Revised in [2]). Due to the absence of general selective forces, the rate of evolution of the satDNA sequences is very high, and they show great variation in monomer size, nucleotide composition, chromosomal distribution, and abundance in the genome. Those factors can affect genome composition,



Citation: Gutiérrez, J.; Aleix-Mata, G.; Montiel, E.E.; Cabral-de-Mello, D.C.; Marchal, J.A.; Sánchez, A. Satellitome Analysis on *Talpa aquitania* Genome and Inferences about the satDNAs Evolution on Some Talpidae. *Genes* 2023, 14, 117. https://doi.org/ 10.3390/genes14010117

Academic Editors: Manuel A. Garrido-Ramos, Miroslav Plohl and Eva Šatović-Vukšić

Received: 9 December 2022 Revised: 27 December 2022 Accepted: 28 December 2022 Published: 31 December 2022



Copyright: © 2022 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/). structure, or size between species [10]. On many occasions, satDNA can be chromosomeor species-specific and can give rise to large genome variations between closely related species [2,11].

The biological significance of satDNA and heterochromatin itself is a topic that is currently arousing great scientific interest. Traditionally it has been considered as silenced junk DNA lacking any relevant function. However, there are multiple examples in insect, plant, or mammalian species of centromeric or telomeric repeat sequences being transcribed [8,10,12]. Transcription of satDNA could be associated with the heterochromatin assembly mechanism, the regulation of centromere function, the sister chromatid cohesion, chromosome organization, pairing and segregation, control of telomere elongation, gene regulation, karyotypic evolution, and speciation [4,5,13–23].

Massive sequencing techniques combined with computational and molecular cytogenetic analysis have opened a new way to characterize the repetitive DNA present in a genome. The development of powerful graph-based bioinformatics tools such as Repeat-Explorer and TAndem REpeat ANalyzer (TAREAN) using massive sequencing data have enabled repeatome analysis [24–26]. These programs allowed the characterization of the repeated DNA content of the genomes from Next Generation Sequencing data. In this way, the satellitomes have been studied using these approaches in several animal [7,18,27–32] and plant species [33–36], allowing the testing of satDNA evolution models across distinct eukaryote groups. With the use of these new tools, it is possible to gain knowledge of the different families of satDNAs and to obtain information on the size and variability within a particular genome, as well as its evolution and genetic functions. In addition, this methodology allowed the detection of a large number of satDNA families at once. Illustrative examples are *Triatoma infestans*, *Locusta migratoria*, *Rhynchophorus ferrugineus*, and *Pontastacus leptodactylus* with 42, 62, 112, and 258 families identified, respectively [7,18,31,32].

The genus *Talpa* (family Talpidae) includes 11 species of fossorial mammals [37]. The species of the Talpidae family have very stable karyotypes [38,39]. Most of the species of *Talpa* have chromosome number 2n = 34 (exceptions *T. caeca*: 2n = 36 and *T. caucasica*: 2n = 38), and autosomal fundamental number (FNa) varies between 62 and 66 [39–42]. The sex chromosomes are also well conserved, with the X chromosome typically a medium-sized submetacentric or metacentric and the Y chromosome in most species is a dot-like metacentric, submetacentric, or acrocentric [42,43].

Analysis of repetitive DNAs identified heterochromatin on centromeres of most chromosomes of the karyotype of *T. romana* and *T. europaea* [42], while is less abundant in T. aquitania and T. occidentalis [37,40]. In addition, in some species, the Y chromosome is completely heterochromatic, and most of the species also present an autosomal pair with non-centromeric heterochromatin [37]. Although, the specific molecular content of the heterochromatin and chromosomal distribution of repeats is poorly known in the genus. Previous studies investigated only the chromosomal distribution of ribosomal genes, telomeric sequences, and a LINE fragment [37,42–44]. Telomeric sequences are located at chromosome ends but also arranged as ITSs (interstitial telomeric sequences) in some pericentromeric regions in Talpa romana, Talpa europaea, and T. aquitania [37,42]. LINE sequences are widely distributed on all chromosomes, as expected, with a noteworthy enrichment on pericentromeric C-band positive regions and the Y chromosome [43]. Interestingly, chromosome painting analyses demonstrate that heterochromatic blocks from two autosomal pairs and Y-heterochromatin share repeat sequences [43]. Recently, the genome of Talpa occidentalis has been sequenced, revealing that about 30% of the genome is comprised of repetitive sequences, similar data to other species in this group (15% of LINEs; 7% of SINEs; 4% of LTR; about 1% DNA transposon) [45].

In most of the mammalian genomes sequenced, little information is provided about the repetitive DNA sequence composition, the repeatome, and especially about the satDNA sequence composition, the satellitome. Until now, few satellitome analyses have been performed on mammal species [46], including insectivore species. Moreover, in this family, there is scarce information about repeat in general and satDNAs in particular. Data on satDNAs repeats remain almost completely missing for the *Talpa* species genomes. The aim of this work is to characterize the satDNA families, the satellitome, in the genome of the recently described species of this genus *T. aquitania* by using the Next Generation Sequencing approach. The satellitomes of the *T. aquitania* genome present 16 different satDNA families, including the telomeric sequences. Most of the satDNA families are present in the genomes of the other *Talpa* species analyzed, while in the genomes of other more distant species of the Talpidae family only some of them are present. Moreover, chromosomal localization by Fluorescence in situ Hybridization (FISH) revealed that some satDNAs are localized preferentially on centromeric and non-centromeric heterochromatin in *T. aquitania* and also in the sister species *T. occidentalis* karyotype. The differences observed between *T. aquitania* and the close relative *T. occidentalis* suggested that the satellitome is a very dynamic component of the genomes.

2. Materials and Methods

2.1. Talpa Specimens

For this study, we analyzed samples from three *T. aquitania* (captured in Torme (Burgos); northern Spain) and four *T. occidentalis* (captured in Granada; southern Spain) males. Permission for capture was granted by the Servicio Territorial de Medio Ambiente de Burgos (Junta de Castilla-León) and the Delegación Territorial de la Consejería de Agricultura, Pesca y Medio Ambiente de Granada (Junta de Andalucía). All capture and sacrifice protocols were approved by the Junta de Andalucía Ethics Committee for Animal Experimentation (code: 22/05/2018/094).

2.2. DNA Extraction, Genome Sequencing, Sequence Clustering, and Analysis

Liver tissue samples were stored in 100% ethanol at -20 °C and DNA samples for PCR were extracted following the standard phenol-chloroform procedure. For Illumina sequencing, we extracted genomic DNA (gDNA) from one adult male of *T. aquitania* and one adult male of *T. occidentalis* using the Gentra Puregene Tissue Kit (Qiagen). The extracted gDNAs were sequenced with Illumina technology. Briefly, for genome sequencing, approximately 3 µg of genomic DNA was used for the construction of a library of 350-bp-length fragments. This library was based on the Illumina[®] HiseqTM 2500 platform and paired-end reads with 2 × 150 bp were obtained. A total of 4.5 Gb of sequences were obtained from the *T. aquitania* genome, corresponding to a coverage of about 2 × of the genome, considering the genome of similar size to the *T. occidentalis* genome (2.099 Gb) [45], and 2 Gb were obtained from the *T. occidentalis* genome (approximately a coverage about 1×). Graph-based clustering analysis [24–26,47], a method for similarity-based clustering of sequence reads, was performed with 829.778 reads of the *T. aquitania* genome using RepeatExplorer, implemented within the Galaxy environment (http://repeatexplorer.org/, accessed on 19 April 2021).

The genome proportion for the repetitive DNA clusters was calculated as read percentages. Additionally, all the clusters were analyzed with sequence-similarity searches of the assembled contigs against GenBank using BlastN (http://www.ncbi.nlm.nih.gov/, accessed on 5 September 2021) and Repbase using the program RepeatMasker version 4.0.9 (http://repeatmasker.org; https://www.dfam.org/, accessed on 1 April 2022, [48] and CENSOR (http://www.girinst.org/, accessed on 1 April 2022). In addition, to identify satDNA repeats, contigs were analyzed using Dotmatcher (http://emboss.bioinformatics. nl/cgi-bin/emboss/dotmatcher/, accessed on 5 September 2021). For the analysis of the characteristics of sequences of the clusters, we used the software Bioedit (version 7.2.5) (http://www.mbio.ncsu.edu/BioEdit/bioedit.html, accessed on 5 September 2021) [49] and Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/, accessed on 5 September 2021) [50].

Estimates of evolutionary divergence between nucleotide sequences were conducted with MEGA 10 using p-distance. The satDNA family abundance and divergence were calculated with RepeatMasker with the "-a" option and the RMBlast search engine. Four

million reads of *T. aquitania* and *T. occidentalis* samples were selected and aligned to the total collection of satDNA dimers, or monomer concatenations of approximately 200 bp in length in satDNAs with small size monomer. A satellite landscape was generated considering distances from the sequences applying the Kimura 2-parameter model with the Perl script calcDivergenceFromAlign.pl and createRepeatLandscape.pl from the RepeatMasker suite. The tandem structure index (TSI) was calculated from the out file of RepeatMasker output as in Montiel et al. [31]. We calculated the number of reads with at least 89% read length aligned to the dimer (internal/pure satDNA reads) divided by the total number of aligned reads (internal/pure + external/mixed satDNA reads). The satellite landscape created was used to calculate the divergence peak (DivPeak) and the relative abundance size of the peak (RSP) for each satDNA family following the method in [51]. Briefly, the divergence value where the landscape had a maximum (DivPeak) was calculated and the relative peak size as the sum of satDNA abundances at $\pm 2\%$ divergence from the divergence peak (peak size) was divided by the total abundance of that satDNA family. The relative abundance among samples was calculated as the log_2 of the ratio of the compared sample divided by the T. aquitania sample in order to facilitate the understanding of the foldchange between samples. Then, to check the presence of *T. aquitania* satDNAs on the closest relative species the consensus sequences were also masked on the chromosomescale genome assembly of *T. occidentalis*, based on long- (PacBio) and short-read (Illumina) sequencing and scaffolded using Hi-C data, published by Real et al. [45] (GCA_014898055.1). Moreover, aiming to check sharing of satDNAs among representatives of family Talpidae abundance and divergence were calculated from low-coverage sequencing data of other species that were available on NCBI SRA database: Talpa europaea (SRX8240408), Scalopus aquaticus (SRX4562103), Condylura cristata (SRX101046), Galemys pyrenaicus (SRX10243621) and *Uropsilus gracilis* (SRX4562112). Calculations and statistical analysis as a correlation between variables using Spearman's rank correlation rho and comparison between paired samples using Wilcoxon signed rank exact test were performed in R base v.4.0.1 [52]. Figures were also plotted in R with ggplot2 and Viridis packages [53,54].

2.3. Satellite DNA Probes, Chromosome Preparations, and FISH

The consensus sequence of each satDNA family (Table S1) was used to design one (small consensus sequences) or two oligonucleotides (large consensus sequences) by using the Primer3 software [55] (Table S2). Only three SatDNA families with large monomers (TaquSat4-437-466; TaquSat5-3102 and TaquSat11-71) were amplified and labeled by PCR using the specifically designed primer pairs. PCR was performed in 50 µL of reaction mixture containing 100 ng of *T. aquitania* genomic DNA, 10 pmol of each primer, 1 µL 10 mM dNTPs, 5 µL of $10 \times NH_4$ Reaction Buffer, 2.5 µL of 50 mM MgCl₂, and 5 U of BIOTAQTM DNA Polymerase (Bioline). The PCR program used was 5 min at 95 °C and 30 cycles: 30 s at 95 °C, 30 s at 55 °C, and 60 s at 72 °C, with a final extension of 5 min at 72 °C. PCR amplicons were resolved in ethidium bromide-stained 1% agarose gels; the appropriate bands were isolated from the gel with a QIAquick gel extraction kit (Qiagen), and labeled with biotin-16-dUTP or digoxigenin-11-dUTP by PCR amplification, with the specific primers pairs. For satDNA families with small monomers, the oligonucleotides based on the most conserved regions were directly labeled with biotin-16-dUTP or digoxigenin-11-dUTP using Terminal Transferase (Roche) as described by Pita et al. [18].

Chromosomes were prepared from bone marrow cells following the method described by Burgos et al. [56]. The location of satDNAs sequences by Fluorescence in situ Hybridization (FISH) was performed as previously described by Pita et al. [18], Fernández et al. [57], and Cabral-de-Mello and Marec [58]. Briefly, the hybridized probes, labeled with biotin-16-dUTP, were detected by avidin-based indirect fluorescence techniques, with two rounds of amplification [18,57] or with Alexa Fluor 488 streptavidin conjugate, while the probes labeled with digoxigenin-11-dUTP were detected with Anti-Digoxigenin-Rhodamine (Roche) [58]. Slides were mounted with Vectashield (Vector Laboratories, Newark, CA, USA) containing DAPI to counterstain the chromosomes. The images were captured and analyzed using a fluorescence microscope (Olympus BX51) equipped with a CCD camera (Olympus DP70), and processed with Adobe Photoshop software.

3. Results and Discussion

3.1. General Characterization of T. aquitania Satellitome

Sequencing data (4.5 Gb) of the *T. aquitania* genome produced 30,149,458 reads. Nucleotide analyses showed that A + T genome content was 59.06% (GC content 40.94%). A high A + T content is a common feature within the Talpidae genomes (58.09% *T. occidentalis*, 59.5% *T. europaea*, 58.3% *G. pyrenaicus*, 58% *S. aquaticus*, 58.1% *C. cristata*, and 58.9% *U. gracilis*) (GenBanK data; [45,59]).

From the total reads, a subset of 8 million pair-end reads was randomly selected and processed in RepeatExplorer. From that subset, 829,778 sequences were selected in the pipeline. Those correspond to about 6% of the genome of *T. aquitania*, considering the genome of similar size as the *T. occidentalis* genome (2.099 Gb) [45].

After RepeatExplorer clustering, 263,881 reads were grouped into 38,295 clusters, with 20% of the *T. aquitania* genome composed of repeated sequences. All the remaining reads (565,897) were classified as singletons. From those clusters, 425 represented at least 0.0024% of the genome and those that featured a star-like or circular graph topology, typically observed in the satDNA families, were deeply analyzed. For each candidate cluster, we examined the contigs assembled by RepeatExplorer to search tandem repeated structures using the Dotmatcher and multiple-sequence alignments and manual inspection to determine the consensus sequences. After the computational analysis, we identified 15 satDNA families (Tables 1 and S1). The satDNA families were named according to Ruiz-Ruano et al. [7], including the species name abbreviation, a number in decreasing order of abundance and the length of the repeat sequence monomer, starting from TaquSat1-183 (the most abundant) to TaquSat15-64 (the least abundant). For the TaquSat4 family, two variants were found with 437 and 466 bp. Telomeric sequence repeat TTAGGG was not found through this analysis. Nevertheless, the (TTAGGG)₅₀ repeat was included in the RepeatMasker analysis showing that the abundance of telomeric repeats was 0.13% (Table 1). In this way, including the telomeric sequences, the *T. aquitania* satellitome is composed of at least 16 satDNA families, corresponding to 1.24% of the genome (Table 1). This percentage of satDNA is in accordance with observed in other Talpidae species as G. pyrenaicus (1.08%) [59]. On other species of mammals in which RepeatExplorer analysis was carried out, Valeri et al. [46] found only one satDNA family, corresponding to 0.87% of the genome of *Trichechus manatus*, revealing that the picture of abundance and number of families could be highly divergent on distinct groups, deserving more analysis.

The first satDNA family TaquSat1-183 represents 0.558% of the genome. Together with the next five families in the list and the telomeric sequences, they all comprise 1.13% of the genome, while the remaining 11 satDNAs represent only 0.11%. The mean A + T content is 50.43% (with variation between 67.2% and 30.4%), lower than the overall genome A + T content (59.06%). The mean of the monomer length is 289.24, with variation between the largest one with 3102 bp (TaquSat5-3102) and the smallest one with 6 bp (TaquSat3-6 and telomeric sequences), being the monomer length of the most abundant satellite DNA of 183 bp (TaquSat1-183). Most of the satDNA families have repeat units smaller than 183 bp, with the exception of the TaquSat4-437-466 and TaquSat5-3102 in global the mean of monomer length is 289.24 bp (median 71.00).

Name	Genome Proportion	Repeat Unit Length (bp)	A + T Percentage	Kimura Divergence	TSI	DivP	RSP
TaquSat1-183	0.55864	183	67.2	12.85	0.85	10	0.36
TaquSat2-107	0.16776	107	55.1	13.63	0.93	13	0.56
TaquSat3-6	0.12463	6	66.7	15.61	0.76	24	0.21
TaquSat4-437-466	0.146 (0.08–0.066)	437–466	30.4–32.6	13.3–18.5	0.83–0.84	12–19	0.44-0.42
TaquSat5-3102	0.04003	3102	53	0.14	0.99	0	1.00
TaquSat6-84	0.04000	84	65.5	20.63	0.53	21	0.28
TaquSat7-60	0.00993	60	48.3	16.72	0.41	17	0.28
TaquSat8-45	0.00532	45	51.1	3.01	0.45	2	0.81
TaquSat9-90	0.00425	90	56.7	16.28	0.40	14	0.34
TaquSat10-24	0.00418	24	50	18.16	0.48	15	0.48
TaquSat11-71	0.00233	71	40.3	2.1	0.09	1	0.98
TaquSat12-101	0.00180	101	50.5	7.73	0.01	6	0.43
TaquSat13-54	0.00175	54	59.3	18.7	0.20	24	0.24
TaquSat14-17	0.00129	17	35.3	21.69	0.36	18	0.33
TaquSat15-64	0.00045	64	45.3	4.89	0.00	8	0.38
Telomeric	0.13160	6	50	18.29	0.91	18	0.57
Total	1.2422						
Mean		289.24	50.43	13.07	0.53	13.06	0.48
SD		737.28	11.19	6.89	0.33	7.63	0.24
Median		71.00	50.50	15.61	0.48	14.00	0.42

Table 1. Data of the satDNA families found in *T. aquitania*: genome proportion (%), the repeat unit length, A + T content and divergence (%), tandem structure index (TSI), divergence peak (DivP) and relative abundance size of the peak (RSP).

The relationships between satDNA families were analyzed by comparison of the consensus sequences. Most of the satDNA families did not show similarity with the sequences of other families. However, in the satDNA TaquSat4-437-466, analyzing the contigs alignments, and manually, we identified two different monomer lengths 437 and 466, both monomers shared a fragment with an identity of 78.21%, and the remaining portion of the longer monomer corresponded with the first portion of a new the monomer repetition. On some posterior analysis, they are considered as TaquSat4-437 and TaquSat4-466.

Nucleotide divergence of satDNA families in T. aquitania ranged between 0.14% (TaquSat5-3102) and 21.69% (TaquSat14-17). Altogether, the satellitome of *T. aquitania* shows a mean nucleotide divergence value of 13.07% (Table 1). Additionally, we calculated the tandem structure index (TSI) of each satDNA [31]. This index represents the proportion of internal reads of satDNA arrays (reads containing only satDNA sequence) with respect to the total of reads with satDNA sequence. The TSI varies between 0 and 1 and informs about the clusterization level of the satDNA family. Therefore, dispersed families with arrays formed by a low number of repeat units would present a lower value of TSI in comparison to families with their repeat units concentrated in longer arrays, which would present TSI values closer to 1. The satDNAs of T. aquitania have a TSI mean of 0.53. The first six most abundant satDNAs (including the telomeric) have the highest TSI values, ranging between 0.99 for TaquSat5-3102 and 0.76 for TaquSat3-6. However, in satDNA families with large monomer sizes, high values of the TSI index could be related more to the size than the structure since short reads used in that calculation could correspond to the same monomers and not to tandem monomers. The remaining ones have the lowest TSI values ranging between 0.53 for TaquSat6-84 and 0.00 for TaquSat15-64.

The satellite landscape is a good representation of both the homogenization or degeneration status of each satDNA family. Those landscapes were obtained by plotting the abundance versus divergence of satDNA sequences against their consensus sequence. SatDNA families with a recent expansion would be formed by nearly identical monomers and almost the total abundance of the family would be contained in the landscape peak. These families would show leptokurtic satellite landscape distributions, with their peaks on low divergence values, as the families TaquSat5-3102, TaquSat8-45, and TaquSat11-71 (Figure 1). On the contrary, satDNA families with high divergence would contain monomers with a great accumulation of mutations in their sequence and the satDNA family abundance would not be contained in the peak. These divergent families would show platykurtic satellite landscape distribution with the peak on high divergence values. Platykurtic landscapes were shown by the family TaquSat3-6, TaquSat7-60, and TaquSat13-54 (Figure 1). The rest of the satDNA families displayed mesokurtic satellite landscapes indicating families amplified time ago and in process of degeneration (Figure 1). In addition, we calculated two indices, the divergence value where the landscape had a maximum (DivPeak) (Table 1), as a measure of the family degeneration, and the relative abundance size of the peak (RSP) (Table 1), as a measure of the family homogenization [51]. Both indices were inversely correlated in the T. aquitania satellitome (Spearman's rank correlation $r_s = -0.75$, p < 0.001) where the recently expanded families showed low DivPeak and high RSP values, and the highly divergent and degenerated families presented high DivPeak and low RSP values (Figure 2).



Figure 1. Satellite landscapes (abundance versus divergence) of T. aquitania satDNA families.



Figure 2. Correlation between the homogenization (RSP) and degeneration (DivP) indices of *T. aquitania* satDNA families. Recently expanded families are in blue and highly divergent and degenerated families are in yellow.

3.2. Sister Species Genomic Analysis Reveals Insight about satDNAs Diversification on Talpidae

The satellitome is a very dynamic genome component that could vary between closely related species or even among populations of the same species [15,18,31]. We were interested in studying the putative conservation of satDNAs in *T. occidentalis*, the species considered as the closely related species of *T. aquitania* [37,60,61] and other Talpidae representatives. Therefore, the consensus sequences of *T. aquitania* satDNA families were used to investigate their presence in the species *T. occidentalis* and other Talpidae.

The abundance and divergence of each satDNA family sequence were analyzed with RepeatMasker using a sample of 4 million reads of T. occidentalis (aprox. 400 Mb, $0.2 \times$ genome coverage). Fifteen of the sixteen satDNAs found on *T. aquitania* genome were shared with *T. occidentalis*, with the unique exception of TaguSat8-45. The satDNA sequences accounted for 0.69% of the T. occidentalis genome, corresponding to almost half of the satellitome in *T. aquitania*. The variation of abundance among samples was calculated with the ratio between samples (*T. occidentalis/T. aquitania*), which was represented as log₂(*T. occidentalis/T. aquitania*). Most satDNA families (10 out of 15) were less abundant on the *T. occidentalis* genome (\log_2 ratio < -0.6). The most abundant family of *T. aquitania*, TaquSat1-183, only accounted for 0.13% of T. occidentalis, being the third satDNA family in terms of abundance on *T. occidentalis*. In contrast, the two variants of TaquSat4 (TaquSat4-437 and TaquSat4-466) were more abundant on *T. occidentalis* (log₂ ratio > 0.6), with TaquSat4-437 being the most abundant satDNA family on *T. occidentalis* accounting for approximately the 0.20% of the genome (Table S3). As fast-evolving components of the genome, the changes in satDNA abundance among closely related species are not uncommon. The expansion or divergence of these repeated sequences configures specie-specific genetic backgrounds, which could contribute to the speciation process [15,62].

The satellite landscapes for each family were also compared between species, including legeneration (DivPeak) and homogenization (RSP) indices. The landscape distribution atDNA families did not show significant differences, besides abundance, between

the degeneration (DivPeak) and homogenization (RSP) indices. The landscape distribution of satDNA families did not show significant differences, besides abundance, between species, and also indices, DivPeak and RSP, were similar among samples (Tables 1 and S3, Figure S1). It was interesting, however, that satDNA families with a leptokurtic distribution on their satellite landscape on *T. aquitania*, recently expanded families, were low abundance with larger degeneration value (TaquSat11-71) and low (TaquSat5-3102) or even not present (TaquSat8-45) on the homogenization index on *T. occidentalis*. In fact, one would expect that between closely related species, the satellitome would be similar, and between species that are more distant, the satellitome might show significant differences. However, in a comprehensive analysis of the satellitomes of bird species, the observed pattern was not in agreement with this assumption. In a group of birds with more deeply diverged species, satellites were found to be more similar and were less similar among more recently diverged species [63].

Additionally, as the first genome assembly of *T. occidentalis* was published by Real et al. [45], satDNA consensus was mapped to the assembly to test their presence on it. The proportion of satDNA sequences in the assembly was 0.20%, a lower value than the estimation with T. occidentalis low-coverage sequencing data (0.69%). Studies identifying satDNA sequences on the assembled genomes are scarce but their underestimation of the assemblies has been a common theme. For instance, the satellitome of *Rhodnius prolixus*, which accounts for 8% of the genome, only makes up 5.6% of the genome assembly [30]. Another example is the satellitome of the red palm weevil, R. ferrugineus, which is 10 times underrepresented in the genome assemblies [31]. The abundance variation of each family among *T. occidentalis* sequencing data and the assembly was calculated as above (log₂(*T. occidentalis* assembled genome/*T. occidentalis* genome)) (Table S3). Even though no significant differences were found between the two samples (p = 0.1167 two-sided Wilcoxon signed-rank test), there was a negative correlation between the TSI value and the variation of abundance among samples (Spearman's rank correlation $r_{\rm S} = -0.63 p = 0.008$). Thereupon, satDNA families with larger TSI values, families with higher clusterization forming large arrays, were less abundant in the assembly, while satDNA families with lower TSI values, dispersed families with arrays made by a low number of monomers, were more abundant (Figure S2). Repeated DNA sequences tend to be discarded in the assembly process, due to the challenge that their assembly supposes. Therefore, satDNA families dispersed in small clusters scattered in the genome were better assembled than satDNA families composed of large clusters that were likely to be collapsed just in a few monomers. The correlation between TSI and the satDNA estimation in genome assemblies has been reported before. In R. ferrugineus, the high-TSI families showed a higher underestimation than low-TIS families, for example, the family RferSat01-169, which accounts for 0.01–0.06% of the assemblies whereas it was estimated at 20% of the genome [31].

The presence of *T. aquitania* satDNA families was tested on the genomes of other species of the family Talpidae, including *T. europaea, S. aquaticus, C. cristata, G. pyrenaicus,* and *U. gracilis*. Raw data of these species were available in the NCBI SRA database and the analysis was performed as above. The number of satDNA families that present those species was variable. On *T. europaea,* the other member of the subfamily Talpinae, it was identified 14 satDNA families, one less than in *T. occidentalis* (Table S3, Figure S3). The species where fewer families were identified were *U. gracilis* (7 families), the most distant species from *T. aquitania*. Interestingly, those seven families are also presented in the other species and may be conserved in the family Talpidae. The abundance and divergence of satDNA families also seem to be related to the phylogenetic distance to *T. aquitania*. Therefore, for *T. aquitania* satDNA sequences, *T.* europaea displayed a larger abundance (2.27%) and lower Kimura divergence (27.21%). Analysis of Dugongidae mammals revealed also the conservation of one satDNA family found initially in *T. manatus* in other species of the family with differences in abundance [46]. Those differences would be consistent with

the library hypothesis proposed by Fry and Salser [64]. From the same pool of satDNA sequences, certain families could eventually expand and fix in one species, whereas other families could expand in other species. Consequently, at the same time species distance themselves, their satellitomes also do so, contributing to differences among species or even favoring new speciation events [8,65].

3.3. SatDNAs Chromosome Localization on T. aquitania and T. occidentalis Reveals Information about Karyotypic Evolution

T. aquitania karyotype has a diploid number of 2n = 34 (NFa = 64) and all chromosomes including the sex chromosomes are biarmed, either metacentric or submetacentric. The X chromosome is medium-sized submetacentric and the Y chromosome is submetacentric and the smallest of the karyotype. As in all the karyotypes of all *Talpa* species studies on *T. aquitania*, one of the largest autosomes (pair 3) harbors a secondary constriction [35,38,40]. In *T. aquitania* karyotype C-banding revealed heterochromatin in centromeric regions of a few pairs of chromosomes, on the proximal part to the centromere of the short arm of pair 14, on the entire short arm of pair 16, and in the entire Y chromosome [35]. The *T. occidentalis* karyotype is very similar to the *T. aquitania*, with the main differences due to the size of the Y chromosome, which is a medium size biarmed submetacentric on *T. aquitania* and a dot-like probably metacentric, in *T. occidentalis*; and in the size of the small arm of the autosome pair 16, that is large on *T. aquitania* and very small in *T. occidentalis* [35,41].

Only seven of the satDNA families of *T. aquitania* were visualized by FISH in the chromosomes of this species and only six on *T. occidentalis* chromosomes. These satDNAs presented very similar distribution on both species, and most of the variations are due to the above-commented differences between the karyotypes (Figure 3).

Three satDNAs (TaquSat1-183, TaquSat3-6, and TaquSat4-437-466) are located on pericentromeric regions, including the X chromosome, however, only TaquSat4-437-466 is also located on the Y chromosome, and TaquSat3-6 has one interstitial band in the arm of one small autosome in *T. aquitania* not present in *T. occidentalis* (Figure 3a–f).

SatDNA TaquSat2-107 is located on subtelomeric regions of the long arm of the Y chromosome and on the subtelomeric regions of the heterochromatic short arm or the pair 16 in *T. aquitania* (Figure 3g), while in *T. occidentalis* it is probably on the small arm of the minute Y chromosome and occupying all the small arm of the pair 16 (Figure 3h). These localizations suggested that the long arm of the Y chromosome of *T. aquitania* arises by the enlargement of the small arm of the Y chromosomes of *T. occidentalis*. In the same way, the enlargement of the heterochromatic small arm of a pair 16 of *T. occidentalis* arises the big small arm of this pair in *T. aquitania*. Heterochromatin enlargements in both chromosomes amplified the satDNA TaquSat6-84. In fact, TaquSat6-84 is accumulated on the heterochromatin of both, the Y chromosome and the short arm of the pair 16 in *T. aquitania*, but was undetectable in *T. occidentalis* (Figure 3i,j). The recent amplification of pair 16 heterochromatic small arm in *T. aquitania* could explain the different composition with chromosome 14 heterochromatin.

SatDNA TaquSat5-3102 is distributed along all the chromosomes in both species, with enrichment on heterochromatin regions of pair 14, the centromeric region of some pairs, and on the Y chromosome; however, is not accumulated on the heterochromatin region of pair 16 (Figure 3k,l). The satDNA TaquSat11-71 is also distributed in both species along all the chromosomes, with enrichment on heterochromatin short arms of pairs 14 and 16, and on the Y chromosome (Figure 3m,n). Hence, the heterochromatic short arm of pairs 14 and 16 on both species shared satellite sequences, as the TaquSat2-107 and TaquSat5-3102, but also present different compositions on sequences, such as TaquSat6-84 and TaquSat5-3102.

Our data indicate that in *Talpa* genomes satDNA families are preferentially accumulated in the centromeric regions and in the C-band positive heterochromatic regions. Overall, the results demonstrate that the satellitome of *T. occidentalis* and *T. europaea* is very similar to that of *T. aquitania*, but differed significantly in the presence and amount of several satDNA families. As the estimated divergence time between *T. aquitania* and *T. occi-*

dentalis is 2.47 ± 0.12 million of years ago (Mya) and between these species and *T. europaea* is 2.82 ± 0.10 Mya [61], the results demonstrated that the satellitome is a very dynamic component of the genomes due to the differences observed in these closely related species.



Figure 3. Male metaphases of *T. aquitania* (**a**,**c**,**e**,**g**,**i**,**k**,**m**) and *T. occidentalis* (**b**,**d**,**f**,**h**,**j**,**l**,**n**) hybridized with satDNA probes: (**a**,**b**): TaquSat1-183; (**c**,**d**): TaquSat3-6; (**e**,**f**): TaquSat4-437-466; (**g**,**h**): TaquSat2-107; (**i**,**j**): TaquSat6-84; (**k**,**l**): TaquSat5-3102; (**m**,**n**): TaquSat11-71. Arrowhead denotes the interstitial signal on the small autosome pair in (**c**), accumulation of signal on the centromeric region of one autosome pair in (**k**). Insert in (**h**), pair 16 chromosomes from other metaphase hybridized with the same probe.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/genes14010117/s1, Figure S1: Satellitome landscapes (abundance versus divergence) of *T. aquitania* satDNA families on *T. aquitania* and *T. occidentalis* genomes; Figure S2: Correlation between the tandem structure index and the fold-change of satDNA family abundance among *T. occidentalis* assembled genome and *T. occidentalis* genome; Figure S3: Presence and abundance of *T. aquitania* satDNA families across species of the family Talpaidea; Table S1: *T. aquitania* satDNA families consensus sequences; Table S2: List of oligos used to localize by FISH the repeat DNA sequences of *T. aquitania*; Table S3: Excel file containing data about *T. aquitania* satDNA families abundance in Talpidae species.

Author Contributions: J.G., G.A.-M., E.E.M., D.C.C.-d.-M., J.A.M. and A.S., conceived, designed, and performed the experiments, analyzed the data, and wrote the draft. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Consejería de Innovación, Ciencia y Empresa of the Junta de Andalucía (Group: RNM924), and by the Universidad de Jaén (as part of the program Plan de Apoyo a la Investigación 2019–2020, Acción 1). DCC-d-M. is supported by the María Zambrano grant (E-23-2022-0219886).

Institutional Review Board Statement: Permission for captures was granted by the Servicio Territorial de Medio Ambiente de Burgos (Junta de Castilla-León) and the Delegación Territorial de la Consejería de Agricultura, Pesca y Medio Ambiente de Granada (Junta de Andalucía). All capture and sacrifice protocols were approved by the Junta de Andalucía Ethics Committee for Animal Experimentation (code: 22/05/2018/094).

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and analyzed during the current study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Charlesworth, B.; Sniegowski, P.; Stephan, W. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 1994, 371, 215–220. [CrossRef] [PubMed]
- 2. Garrido-Ramos, M.A. Satellite DNA: An Evolving Topic. Genes 2017, 8, 230. [CrossRef] [PubMed]
- 3. López-Flores, I.; Garrido-Ramos, M.A. The repetitive DNA content of eukaryotic genomes. Genome Dyn. 2012, 7, 1–28. [PubMed]
- Biscotti, M.A.; Canapa, A.; Forconi, M.; Olmo, E.; Barucca, M. Transcription of tandemly repetitive DNA: Functional roles. Chromosome Res. 2015, 23, 463–477. [CrossRef] [PubMed]
- Biscotti, M.A.; Olmo, E.; Heslop-Harrison, J.S. Repetitive DNA in eukaryotic genomes. *Chromosome Res.* 2015, 23, 415–420. [CrossRef] [PubMed]
- 6. Choi, I.Y.; Kwon, E.C.; Kim, N.S. The C- and G-value paradox with polyploidy, repeatomes, introns, phenomes and cell economy. *Genes Genom.* 2020, 42, 699–714. [CrossRef] [PubMed]
- Ruiz-Ruano, F.J.; López-León, M.D.; Cabrero, J.; Camacho, J.P. High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci. Rep.* 2016, *6*, 28333. [CrossRef] [PubMed]
- 8. Ugarkovic, D.; Plohl, M. Variation in satellite DNA profiles-causes and effects. EMBO J. 2002, 21, 5955–5959. [CrossRef]
- 9. Palomeque, T.; Lorite, P. Satellite DNA in insects: A review. *Heredity* 2008, 100, 564–573. [CrossRef]
- Plohl, M.; Luchetti, A.; Mestrovic, N.; Mantovani, B. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 2008, 409, 72–82. [CrossRef]
- 11. Mehrotra, S.; Goyal, V. Repetitive sequences in plant nuclear DNA: Types, Distribution, Evolution and Function. *Genom. Proteom. Bioinform.* **2014**, *12*, 164–171. [CrossRef] [PubMed]
- 12. Dimitri, P.; Corradini, N.; Rossi, F.; Vernì, F. The paradox of functional heterochromatin. *Bioessays* **2005**, 27, 29–41. [CrossRef] [PubMed]
- Grewal, S.I.; Elgin, S.C. Transcription and RNA interference in the formation of heterochromatin. *Nature* 2007, 447, 399–406. [CrossRef] [PubMed]
- 14. Josse, T.; Teysset, L.; Todeschini, A.-L.; Sidor, C.M.; Anxolabehere, D.; Ronsseray, S. Telomeric trans-silencing: An epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet.* **2007**, *3*, 158. [CrossRef] [PubMed]
- 15. Ferree, P.M.; Prasad, S. How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. *Genet. Res. Int.* **2012**, 2012, 430136. [CrossRef]
- Pezer, Z.; Brajkovic, J.; Feliciello, I.; Ugarkovic, D. Satellite DNA-Mediated Effects on Genome Regulation. *Genome Dyn.* 2012, 7, 153–169. [PubMed]

- 17. Plohl, M.; Mestrovic, N.; Mravinac, B. Centromere identity from the DNA point of view. *Chromosoma* **2014**, *123*, 313–325. [CrossRef] [PubMed]
- 18. Pita, S.; Panzera, F.; Mora, P.; Vela, J.; Cuadrado, A.; Sánchez, A.; Palomeque, T.; Lorite, P. Comparative repeatome analysis on *Triatoma infestans* Andean and Non-Andean lineages, main vector of Chagas disease. *PLoS ONE* **2017**, *12*, e0181635. [CrossRef]
- 19. Pita, S.; Mora, P.; Vela, J.; Palomeque, T.; Sánchez, A.; Panzera, F.; Lorite, P. Comparative analysis of repetitive DNA between the main vectors of Chagas disease: *Triatoma infestans* and *Rhodnius prolixus*. *Int. J. Mol. Sci.* **2018**, 24, 1277. [CrossRef]
- 20. Fingerhut, J.M.; Moran, J.V.; Yamashita, Y.M. Satellite DNA-containing gigantic introns in a unique gene expression program during *Drosophila* spermatogenesis. *PLoS Genet.* **2019**, 15, e1008028. [CrossRef]
- Shatskikh, A.S.; Kotov, A.A.; Adashev, V.E.; Bazylev, S.S.; Olenina, L.V. Functional Significance of Satellite DNAs: Insights from Drosophila. Front. Cell Dev. Biol. 2020, 8, 312. [CrossRef] [PubMed]
- 22. Jagannathan, M.; Yamashita, Y.M. Defective satellite DNA clustering into chromocenters underlies hybrid incompatibility in *Drosophila*. *Mol. Biol. Evol.* **2021**, *38*, 4977–4986. [CrossRef] [PubMed]
- Wei, X.; Eickbush, D.G.; Speece, I.; Larracuente, A.M. Heterochromatin-dependent transcription of satellite DNAs in the *Drosophila* melanogaster female germline. *eLife* 2021, 10, e62375. [CrossRef] [PubMed]
- Novák, P.; Neumann, P.; Macas, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinform.* 2010, 11, 378. [CrossRef] [PubMed]
- Novák, P.; Neumann, P.; Pech, J.; Steinhaisl, J.; Macas, J. RepeatExplorer: A Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 2013, 29, 792–793. [CrossRef] [PubMed]
- Novák, P.; Ávila Robledillo, L.; 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, e111. [CrossRef]
- Utsunomia, R.; Silva, D.M.Z.A.; Ruiz-Ruano, F.J.; Goes, C.A.G.; Melo, S.; Ramos, L.P.; Oliveira, C.; Porto-Foresti, F.; Foresti, F.; Hashimoto, D.T. Satellitome landscape analysis of *Megaleporinus macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. *Sci. Rep.* 2019, *9*, 5856. [CrossRef]
- da Silva, M.J.; Fogarin Destro, R.; Gazoni, T.; Narimatsu, H.; Pereira Dos Santos, P.S.; Haddad, C.F.B.; Parise-Maltempi, P.P. Great abundance of satellite DNA in *Proceratophrys* (Anura, Odontophrynidae) revealed by genome sequencing. *Cytogenet. Genome Res.* 2020, 160, 141–147. [CrossRef]
- 29. Mora, P.; Vela, J.; Ruiz-Ruano, F.J.; Ruiz-Mena, A.; Montiel, E.E.; Palomeque, T.; Lorite, P. Satellitome Analysis in the Ladybird Beetle *Hippodamia variegata* (Coleoptera, Coccinellidae). *Genes* **2020**, *11*, 783. [CrossRef]
- 30. Montiel, E.E.; Mora, P.; Rico-Porras, J.M.; Palomeque, T.; Lorite, P. Satellitome of the red palm weevil, *Rhynchophorus ferrugineus* (Coleoptera: Curculionidae), the most divers among insects. *Front. Ecol. Evol.* **2022**, *10*, 826808. [CrossRef]
- Montiel, E.E.; Panzera, F.; Palomeque, T.; Lorite, P.; Pita, S. Satellitome Analysis of *Rhodnius prolixus*, one of the Main Chagas Disease Vector Species. *Int. J. Mol. Sci.* 2021, 22, 6052. [CrossRef] [PubMed]
- Boštjančić, L.L.; Bonassin, L.; Anušić, L.; Lovrenčić, L.; Besendorfer, V.; Maguire, I.; Grandjean, F.; Austin, C.M.; Greve, C.; Hamadou, A.B.; et al. The *Pontastacus leptodactylus* (Astacidae) repeatome provides insight into genome evolution and reveals remarkable diversity of satellite DNA. *Front. Genet.* 2021, *11*, 611745. [CrossRef] [PubMed]
- Weiss-Schneeweiss, H.; Leitch, A.R.; McCann, J.; Jang, T.S.; Macas, J. Employing next generation sequencing to explore the repeat landscape of the plant genome. In *Next Generation Sequencing in Plant Systematics Regnum Vegetabile*; Hörandl, E., Appelhans, M., Eds.; Koeltz Scientific Books: Königstein, Germany, 2015; pp. 155–179.
- Ruiz-Ruano, F.J.; Navarro-Domínguez, B.; Camacho, J.P.M.; Garrido-Ramos, M.A. Characterization of the satellitome in lower vascular plants: The case of the endangered fern *Vandenboschia speciose*. Ann. Bot. 2019, 123, 578–599. [CrossRef] [PubMed]
- 35. Heitkam, T.; Schulte, L.; Weber, B.; Liedtke, S.; Breitenbach, S.; Kögler, A.; Morgenstern, K.; Brückner, M.; Tröber, U.; Wolf, H.; et al. Comparative Repeat Profiling of Two Closely Related Conifers (*Larix decidua* and *Larix kaempferi*) Reveals High Genome Similarity With Only Few Fast-Evolving Satellite DNAs. *Front. Genet.* 2021, *12*, 683668. [CrossRef] [PubMed]
- 36. Sader, M.; Vaio, M.; Cauz-Santos, L.A.; Dornelas, M.C.; Vieira, M.L.C.; Melo, N.; Pedrosa-Harand, A. Large vs. small genomes in Passiflora: The influence of the mobilome and the satellitome. *Planta* **2021**, *253*, 86. [CrossRef]
- 37. Gutiérrez, J.; Aleix-Mata, G.; Lamelas, L.; Arroyo, M.; Marchal, J.A.; Sánchez, A. Karyotype analysis of the new *Talpa* species *Talpa aquitania* (Talpidae; Insectivora) from northern Spain. *Cytogenet. Genome Res.* **2019**, *159*, 26–31. [CrossRef]
- 38. Biltueva, L.; Vorobieva, N. Chromosome evolution in Eulipotyphla. Cytogenet. Genome Res. 2012, 137, 154–164. [CrossRef]
- 39. Ye, J.; Biltueva, L.; Huang, L.; Nie, W.; Wang, J.; Jing, M.; Su, W.; Vorobieva, N.V.; Jiang, X.; Graphodatsky, A.S.; et al. Crossspecies chromosome painting unveils cytogenetic signatures for the Eulipotyphla and evidence for the polyphyly of Insectivora. *Chromosome Res.* **2006**, *14*, 151–159. [CrossRef]
- 40. Jimenez, R.; Burgos, M.; Díaz de la Guardia, R. Karyotype and chromosome banding in the mole (*Talpa occidentalis*) from the south-east of the Iberian Peninsula. Implications on its taxonomic position. *Caryologia* **1984**, *37*, 253–258. [CrossRef]
- Kawada, S.; Harada, M.; Grafodatsky, A.S.; Oda, S. Cytogenetic study of the Siberian mole, *Talpa altaica* (Insectivora: Talpidae) and karyological relationships within the genus *Talpa*. *Mammalia* 2002, 66, 53–62. [CrossRef]
- 42. Gornung, E.; Volleth, M.; Capanna, E.; Castiglia, R. Comparative cytogenetics of moles (Eulipotyphla, Talpidae): Chromosomal differences in *Talpa romana* and *Talpa europaea*. *Cytogenet*. *Genome Res.* **2008**, 121, 249–254. [CrossRef] [PubMed]

- Gutiérrez, J.; Aleix-Mata, G.; Marchal, J.A.; Arroyo, M.; Castiglia, R.; Sánchez, A. Molecular Cytogenetic Analysis of Karyotype and Y Chromosome Conservation in Species of the Genus Talpa (Insectivora). *Cytogenet. Genome Res.* 2020, 160, 264–271. [CrossRef] [PubMed]
- 44. Zurita, F.; Sánchez, A.; Burgos, M.; Jiménez, R.; Díaz de la Guardia, R. Interchromosomal, intercellular and interindividual variability of NORs studied with silver staining and in situ hybridization. *Heredity* **1997**, *78*, 229–234. [CrossRef] [PubMed]
- Real, F.M.; Haas, S.A.; Franchini, P.; Xiong, P.; Simakov, O.; Kuhl, H.; Schöpflin, R.; Heller, D.; Moeinzadeh, M.H.; Heinrich, V.; et al. The mole genome reveals regulatory rearrangements associated with adaptive intersexuality. *Science* 2020, 370, 208–214. [CrossRef] [PubMed]
- 46. Valeri, M.P.; Dias, G.B.; do Espírito Santo, A.A.; Moreira, C.N.; Yonenaga-Yassuda, Y.; Sommer, I.B.; Kuhn, G.C.S.; Svartman, M. First Description of a Satellite DNA in Manatees' Centromeric Regions. *Front. Genet.* **2021**, *12*, 694866. [CrossRef]
- 47. Novák, P.; Neumann, P.; Macas, J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat. Protoc.* **2020**, *15*, 3745–3776. [CrossRef]
- Storer, J.; Hubley, R.; Rosen, J.; Wheeler, T.J.; Smit, A.F. The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mob. DNA* 2021, 12, 2. [CrossRef]
- 49. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **1999**, *41*, 95–98.
- 50. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.N.; Potter, S.C.; Finn, R.D.; et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [CrossRef]
- Camacho, J.P.M.; Cabrero, J.; López-León, M.D.; Martín-Peciña, M.; Perfectti, F.; Garrido-Ramos, M.A.; Ruiz-Ruano, F.J. Satellitome comparison of two oedipodine grasshoppers highlights the contingent nature of satellite DNA evolution. *BMC Biol.* 2022, 20, 36. [CrossRef]
- 52. R Core Team. A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2019.
- 53. Warnes, G.R.; Bolker, B.; Bonebakker, L.; Gentleman, R.; Huber, W.; Liaw, A.; Lumley, T.; Maechler, M.; Magnusson, A.; Moeller, S.; et al. Gplots: Various R Programming Tools for Plotting Data. R Package Version 3.1.1. 2020. Available online: https://CRAN.R-project.org/package=gplots (accessed on 15 June 2022).
- 54. Garnier, S.; Ross, N.; Rudis, R.; Camargo, A.P.; Sciaini, M.; Scherer, C. Rvision—Colorblind-Friendly Color Maps for R. R Package Version 0.6.1. 2021. Available online: https://cran.r-project.org/package=viridisLite (accessed on 15 June 2022).
- 55. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—New capabilities and interfaces. *Nucleic Acids Res.* 2012, 40, e115. [CrossRef] [PubMed]
- 56. Burgos, M.; Jiménez, R.; Díaz de la Guardia, R. A rapid, simple and reliable combined method for G-banding mammalian and human chromosomes. *Stain Technol.* **1986**, *61*, 257–260. [CrossRef] [PubMed]
- Fernández, R.; Barragán, M.J.; Bullejos, M.; Marchal, J.A.; Martínez, S.; Díaz de la Guardia, R.; Sánchez, A. Molecular and cytogenetic characterization of highly repeated DNA sequences in the vole *Microtus cabrerae*. *Heredity* 2001, 87, 637–646. [CrossRef] [PubMed]
- Cabral-de-Mello, D.C.; Marec, F. Universal fluorescence in situ hybridization (FISH) protocol for mapping repetitive DNAs in insects and other arthropods. *Mol. Genet. Genom.* 2021, 296, 513–526. [CrossRef]
- 59. Escoda, L.; Castresana, J. The genome of the Pyrenean desman and the effects of bottlenecks and inbreeding on the genomic landscape of an endangered species. *Evol. Appl.* **2021**, *14*, 1898–1913. [CrossRef]
- 60. Nicolas, V.; Martínez-Vargas, J.; Hugot, J.P. Preliminary note: *Talpa aquitania* nov. sp. (Talpidae, Soricomorpha) a new mole species from southwest France and north Spain. *Bul. Acad. Vet. Fr.* **2015**, *168*, 329–334. [CrossRef]
- 61. Nicolas, V.; Martínez-Vargas, J.; Hugot, J.P. Molecular data and ecological niche modelling reveal the evolutionary history of the common and Iberian blind moles (Talpidae) in Europe. *Zool. Scr.* 2017, *46*, 12–26. [CrossRef]
- 62. Hughes, S.E.; Hawley, R.S. Heterochromatin: A rapidly evolving species barrier. PLoS Biol. 2009, 7, e1000233. [CrossRef]
- 63. Peona, V.; Kutschera, V.E.; Blom, M.P.K.; Irestedt, M.; Suh, A. Satellite DNA evolution in Corvoidea inferred from short and long reads. *Mol. Ecol.* 2022. [CrossRef]
- 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, 1069–1084. [CrossRef]
- 65. Mestrović, N.; Plohl, M.; Mravinac, B.; Ugarković, D. Evolution of satellite DNAs from the genus Palorus—Experimental evidence for the "library" hypothesis. *Mol. Biol. Evol.* **1998**, *15*, 1062–1068. [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.