Check for updates

Received: 31 August 2023 | Revised: 12 December 2023 | Accepted: 30 January 2024

DOI: 10.1002/jez.b.23244

RESEARCH ARTICLE

WILEY

Variants of a major DNA satellite discriminate parental subgenomes in a hybrid parthenogenetic lizard *Darevskia unisexualis* (Darevsky, 1966)

Pavel Nikitin[1,2] | Sviatoslav Sidorov[3] | Thomas Liehr[4] | Ksenia Klimina[5] | Ahmed Al-Rikabi[4] | Vitaly Korchagin[6] | Oxana Kolomiets[7] | Marine Arakelyan[8] | Victor Spangenberg[7]

[1]Laboratory of Comparative Ethology and Biocommunication, Severtsov Institute of Ecology and Evolution RAS, Moscow, Russia

[2]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia

[3]Computational Regulatory Genomics, MRC Laboratory of Medical Sciences, Hammersmith Hospital Campus, London, UK

[4]Jena University Hospital, Friedrich Schiller University, Institute of Human Genetics, Jena, Germany

[5]Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia

[6]Institute of Gene Biology RAS, Moscow, Russia

[7]Laboratory of Cytogenetics, Vavilov Institute of General Genetics RAS, Moscow, Russia

[8]Department of Zoology, Yerevan State University, Yerevan, Armenia

Correspondence

Victor Spangenberg, Vavilov Institute of General Genetics RAS, Gubkin str. 3, Moscow 119991, Russia.
Email: v.spangenberg@gmail.com

Funding information

State contract of VIGG RAS (cytogenetic experiments), Grant/Award Number: 0092-2022-0002; Science Committee of Republic of Armenia (microscopy), Grant/Award Number: 23IRF-1F06; Russian National Foundation (expeditions and keeping the animals), Grant/Award Number: 22-14-00227

Abstract

Hybrid parthenogenetic animals are an exceptionally interesting model for studying the mechanisms and evolution of sexual and asexual reproduction. A diploid parthenogenetic lizard *Darevskia unisexualis* is a result of an ancestral cross between a maternal species *Darevskia raddei nairensis* and a paternal species *Darevskia valentini* and presents a unique opportunity for a cytogenetic and computational analysis of a hybrid karyotype. Our previous results demonstrated a significant divergence between the pericentromeric DNA sequences of the parental *Darevskia* species; however, an in-depth comparative study of their pericentromeres is still lacking. Here, using target sequencing of microdissected pericentromeric regions, we reveal and compare the repertoires of the pericentromeric tandem repeats of the parental *Darevskia* lizards. We found species-specific sequences of the major pericentromeric tandem repeat CLsat, which allowed computational prediction and experimental validation of fluorescent DNA probes discriminating parental chromosomes within the hybrid karyotype of *D. unisexualis*. Moreover, we have implemented a generalizable computational method, based on the optimization of the Levenshtein distance between tandem repeat monomers, for finding species-specific fluorescent probes for pericentromere staining. In total, we anticipate that our comparative analysis of *Darevskia* pericentromeric repeats, the species-specific fluorescent probes that we found and the pipeline that we developed will form a basis for the future detailed cytogenomic studies of a wide range of natural and laboratory hybrids.

KEYWORDS

centromere, chromosome, fluorescence in situ hybridization, hybrid, parthenogenesis, satellite DNA, tandem repeats

Pavel Nikitin, Sviatoslav Sidorov, Thomas Liehr, Oxana Kolomiets, Marine Arakelyan, and Victor Spangenberg contributed equally to this study.

 © 2024 Wiley Periodicals LLC.

1 | INTRODUCTION

The formation of hybrid parthenogenetic lineages of animals has been an area of active research in modern genetics (Barley et al., 2021; Dedukh et al., 2022; Newton et al., 2016). A classical model for these studies is a lizard genus *Darevskia* which includes several unisexual species of hybrid origin (Freitas et al., 2016; Fu et al., 2000; Grechko et al., 2006; Tokarskaya et al., 2001). On the other hand, pairs of species that give rise to unisexually reproducing hybrids have been proposed to be within certain limits of genetic divergence (Murphy et al., 2000). Therefore, differences in the satellite composition of pericentromeres and the overall divergent evolution of these regions may play a role in the formation of hybrid unisexual species.

Previously, we studied the somatic karyotype of a hybrid parthenogenetic species *Darevskia unisexualis* (Spangenberg, Arakelyan, et al., 2020, 2021). The hybrid emerged as a result of an ancestral cross between a maternal species *Darevskia raddei nairensis* and a paternal species *Darevskia valentini* (Murphy et al., 2000; Uzzell & Darevsky, 1975) that diverged ~18.5 million years ago (Murtskhvaladze et al., 2020). *D. unisexualis* has a diploid karyotype (2N = 38) composed of two sets of homeologous (but not homologous) chromosomes inherited from the two diploid parental species (Spangenberg, Arakelyan, et al., 2020). Using comparative genomic hybridization (CGH), we previously showed that chromosomes in this hybrid karyotype preserved their parental species-specific composition of satellite DNA (Spangenberg, Arakelyan, et al., 2020). However, a comparative analysis of the pericentromeric tandem repeats in the parental species is still lacking. Additionally, the karyological studies of nonmammalian animals are often complicated by the fact that the standard antibodies for the immunostaining of the centromeric proteins (primarily from the mammalian CENP family) do not work in many other animal taxa. Moreover, the CGH approach that we used to differentially stain the chromosomes of the hybrid *D. unisexualis* karyotype (Spangenberg, Arakelyan, et al., 2020) is expensive and laborious. Therefore, the development of a species-agnostic computational method for designing DNA probes for fluorescence in situ hybridization (FISH) to differentially stain the pericentromeres of phylogenetically distant species would allow replacing CGH and facilitate future studies of hybrid karyotypes.

In recent years, the combination of molecular cytogenetics and bioinformatic methods provided new insights on the localization, size, and evolution of pericentromeres (Altemose et al., 2022; Biscotti et al., 2015; Nurk et al., 2022). Importantly, recent studies accumulated a vast amount of data on the chromosomal localization of various families of satellite repeats in reptiles (Ahmad et al., 2020; Giovannotti et al., 2009, 2013, 2018; Lisachov et al., 2023; Matsubara et al., 2015; Rojo et al., 2015). Methods for the identification of satellite DNA monomers have diversified significantly, ranging from classical in vitro approaches for the isolation of a fraction of DNA repeats to in silico methods of genomic data analysis (Liehr, 2022; Lisachov et al., 2023; Spangenberg, Losev, et al., 2021).

Here, using the microdissection and target sequencing of the pericentromeric regions of the parental *Darevskia* species, we compared their pericentromeric tandem repeats and found species-specific oligo-DNA probes for the rapid chromosome identification in the hybrid karyotype. Additionally, we found a universal pericentromeric marker for the two parental and the hybrid *Darevskia* species and established the pericentromeric localization of the CLsat repeats from which we derived the probes. Finally, to make our computational analysis reproducible, we implemented it as a publicly available Nextflow pipeline, which, as we anticipate, will also facilitate the development of our method of search for candidate species-specific probes into a standalone software tool.

2 | RESULTS

2.1 | Pericentromeric tandem repeats from *D. raddei nairensis* and *D. valentini* demonstrate distinct sequence characteristics

To predict species-specific DNA FISH probes for *D. raddei nairensis* and *D. valentini*, we sequenced DNA from the microdissected pericentromeric regions of three chromosomes of each of the two species using short paired-end reads. In this way, we took advantage of the abundant tandem repeats located in pericentromeric regions to ensure a strong fluorescent signal. We obtained a total of ~1.9 mln reads for *D. raddei nairensis* and ~1.7 mln reads for *D. valentini*. Next, we assessed the read quality (Supporting Information S1: Figures S1 and S2) and found an asymmetry in the GC-content distribution for *D. valentini* (Supporting Information S1: Figure S2B,D). As our test for viroid, viral, prokaryotic and eukaryotic contamination did not yield plausible contaminants (Supporting Information S3: Table S1), we concluded that the observed bias is an intrinsic property of the genomic DNA from *D. valentini*; hence, we continued the preprocessing of the sequencing data (Figure 1a).

We used Tandem Repeat Analyzer (TAREAN) (Novák et al., 2017) to assemble reads from each species into contigs and to find tandem repeat monomers (TRMs). As, contrary to our expectations, TAREAN did not find tandem repeats, we extracted TRMs from the TAREAN-assembled contigs using Tandem Repeats Finder (TRF) (Benson, 1999). We obtained 963 unique TRMs from ~19,300 contigs of the *D. raddei nairensis* assembly and 8,406 unique TRMs from ~16,400 contigs of the *D. valentini* assembly (Supporting Information S3: Tables S2 and S3). As expected, these assemblies of short reads from highly repetitive pericentromeric regions yielded short contigs with median lengths less than 200 bp (Supporting Information S1: Figure S3A), which limited the maximal lengths of the obtained TRMs.

The majority of TRMs from *D. raddei nairensis* were shorter than 40 bp; additionally, a small group of TRMs had a characteristic length of 140 bp. In contrast, TRMs from *D. valentini* had a characteristic length of ~20 bp, while a small fraction of TRMs were ~40 bp-long (Figure 1c). In addition, TRMs from the two parental species demonstrated drastically different GC-content distributions (Figure 1b). Specifically, while TRMs from *D. raddei nairensis* had a broad GC-content distribution with the highest proportions of TRMs in the range from 35% to 50%, the TRMs from *D. valentini* had the characteristic GC-content within the range from

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

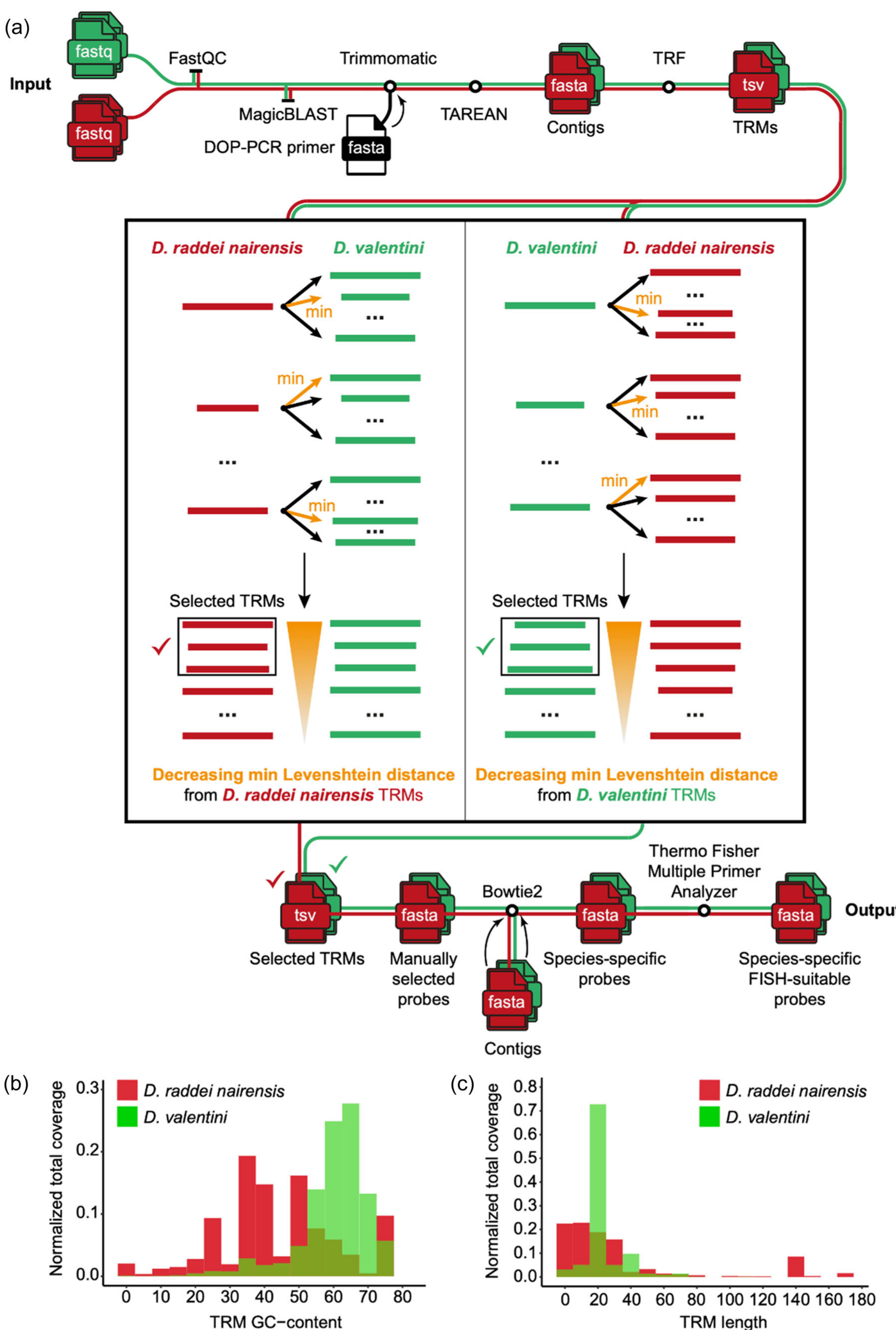

FIGURE 1 (See caption on next page).

55% to 70% (Figure 1b). Of note, the GC-content distribution of the *D. valentini* contigs demonstrated a bimodality that was not reflected by the GC-content distribution of the TRMs obtained for this species (Figure 1b and Supporting Information S1: Figure S3B). Therefore, the peculiar shape of the GC-content distributions of the reads and contigs from *D. valentini* is not caused by different TRM subsets. Overall, the two parental species demonstrated considerable differences in the number of unique TRMs, their lengths and GC-content.

2.2 | Comparative analysis of TRMs reveals a species-specific DNA FISH probe for the maternal species *D. raddei nairensis*

Based on the obtained TRM sets for *D. raddei nairensis* and *D. valentini*, we next predicted species-specific pericentromeric DNA FISH probes for both species. To this end, we selected TRMs from the highest-covered 10% of contigs in each of the two species, then calculated cross-species pairwise edit distances between TRMs selected in *D. raddei nairensis* and *D. valentini* and, finally, in each species we chose several TRMs that have the largest minimal distances to the TRMs of the other species (Figure 1a). By selecting TRMs from the highest-covered contigs, we assumed that these TRMs are the most abundant in pericentromeres and, therefore, should yield the highest fluorescent signal in FISH experiments. Additionally, by choosing the TRMs with the largest minimal distances to TRMs in the other species, we optimized for the species-specificity of the candidate FISH probes that we next extracted from these TRMs. Namely, from the chosen TRMs we manually selected several k-mers and additionally tested their species-specificity as candidate DNA FISH probes by cross-mapping to the contigs of the opposite species. We chose candidate probes that did not map and additionally tested their suitability for FISH experiments by ensuring the absence of self- and cross-hybridization (Figure 1a; see Section 4 for details). In this way, we predicted eight species-specific FISH probes for *D. raddei nairensis* and 8 species-specific probes for *D. valentini*.

We tested the species-specificity of the predicted probes in FISH experiments on the mitotic metaphase plates of the hybrid species *D. unisexualis* and the two parental species, *D. raddei nairensis* and *D. valentini*. One probe for *D. raddei nairensis*, which we later named CLsat30radn (Table 1), demonstrated a strong signal on all of the *D. raddei nairensis* chromosomes (Figure 2b) but no noticeable signal on the *D. valentini* chromosomes (Figure 2c, indicated by an asterisk), confirming species-specificity on the parental karyotypes. Importantly, this probe hybridized only to a half of the *D. unisexualis* chromosomes (19 out of 38, 2N = 38), demonstrating species-specificity on the karyotype of the hybrid species (Figure 2a). Of note, other predicted DNA FISH probes for the pericentromeric regions of *D. raddei nairensis* did not show a strong species-specific fluorescent signal on a half of the *D. unisexualis* chromosomes (data not shown). Additionally, this approach did not allow us to find species-specific FISH probes for *D. valentini*.

2.3 | Variant sequence of a common satellite provides a species-specific FISH probe for the paternal species *D. valentini*

As our approach described above did not yield a species-specific FISH probe for *D. valentini*, we took a different strategy to find it. First, we used the NCBI nucleotide BLAST web-service to characterize the contig from which we obtained the species-specific probe for *D. raddei nairensis*. Using the reptile database, we found that the contig matches known CLsatI sequences of the *Darevskia* genus (Supporting Information: Files S4 and S5).

TABLE 1 Species-specific pericentromeric DNA FISH probes for *Darevskia raddei nairensis* **(CLsat30radn)** and *Darevskia valentini* (CLsat30val) and a DNA FISH probe for the conservative region of the CLsat sequence (CLsat36).

Probe	Probe sequence (5′-3′)	Probe length (nt)	GC-content (%)	Contig of origin	Contig length (nt)	Contig mean coverage	Tm (°C)
CLsat30radn	GACGTACTTGGCTTTTGTAACGT TCAATTT	30	36.7	CL1Contig21	378	1888	69.6
CLsat30val	ACGTTCAGTTTGGCTTACTTGCG TGATTTT	30	40.0	CL107Contig1 (reverse complement)	284	67	73.1
CLsat36	TGGTGTGTTTTCTATGCATTT CGACCTGAAAGAAGC	36	41.7	–	–	–	77.7

Abbreviation: FISH, fluorescence in situ hybridization.

FIGURE 1 (a) Prediction of species-specific DNA FISH probes for *Darevskia raddei nairensis* (red elements) and *Darevskia valentini* (green elements). The inset shows the calculation of the minimal edit distances between TRMs from the two species and the selection of the most species-specific TRMs for each species. The FASTA file with the DOP-PCR primer is an additional input for Trimmomatic; the FASTA files with contigs form an additional input for Bowtie2. See the main text and Section 4 for details. (b) GC-content distributions and (c) length distributions of TRMs from *D. raddei nairensis* and *D. valentini*. To calculate the normalized total coverage, we counted each value of the GC-content (length) in each species the number of times equal to the mean coverage of the corresponding contig and then divided the total count in each bin by the total count across all bins. We hypothesized that in this way we represented the values of the GC-content or length in each bin proportionately to the share of the pericentromere occupied by the corresponding tandem repeats. DOP-PCR, degenerate oligonucleotide-primed polymerase chain reaction; FISH, fluorescence in situ hybridization; TRMs, tandem repeat monomers.

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 2 DNA fluorescence in situ hybridization with oligo-probes CLsat30radn (red) and CLsat30val (green) tested separately on the mitotic metaphase plates of the parental species *Darevskia raddei nairensis* and *Darevskia valentini*. Microchromosomes are marked with "m"; the W chromosome is marked with "w"; an asterisk indicates a lack of a notable fluorescent signal. (a)–(c) Probe CLsat30radn, specific to the karyotype of the maternal species (*D. raddei nairensis*), tested on *D. unisexualis* (a), *D. raddei nairensis* (b) and *D. valentini* (c). (d)–(f) Probe CLsat30val, specific to the karyotype of the paternal species (*D. valentini*), tested on *D. unisexualis* (d), *D. raddei nairensis* (e), and *D. valentini* (f). Chromatin was stained with DAPI (blue). The scale bar is 10 µm.

Therefore, we found that although the CLsat sequence is conserved between the *Darevskia* species (Grechko et al., 2006), it still exhibits enough variation to produce a species-specific probe.

Based on this result, we named the species-specific probe for *D. raddei nairensis* that we described above CLsat30radn (CLsat-derived 30-bp DNA FISH probe for *D. raddei nairensis*) and hypothesized that the CLsatI sequence could also yield a species-specific probe for *D. valentini*. Therefore, we searched for a CLsatI-like sequence among the *D. valentini* contigs. To this end, we mapped a 36-nucleotide sequence CLsat36 from the evolutionarily conserved region of the CLsatI satellite DNA (see fig. 1a in Grechko et al. (2006) and Table 1) to the contigs of *D. valentini* and found a matching contig. This contig had the mean coverage 67, which explains the fact that we missed the TRM derived from it in our first approach, as the lowest mean contig coverage used to calculate pairwise edit distances between TRMs was 155. Next, we selected the CLsat TRM from the obtained *D. valentini* contig and, from this TRM, we selected several 30-nucleotide sequences and tested them as candidate DNA FISH probes.

Using this approach, we obtained a probe CLsat30val (CLsat-derived 30-bp DNA FISH probe for *D. valentini*) that demonstrated a strong species-specific signal in a series of FISH experiments. Namely, CLsat30val efficiently hybridized to all chromosomes of *D. valentini* (Figure 2f) and did not hybridize to the chromosomes of *D. raddei nairensis* (Figure 2e, indicated by an asterisk). Importantly, the probe hybridized only to a half of the chromosomes of *D. unisexualis* (19 out of 2N = 38), demonstrating species-specificity not only on the full parental karyotypes but also on the parental halves of the hybrid karyotype (Figure 2d). In total, we obtained two probes that hybridize specifically to the parental chromosomes of the *D. unisexualis* karyotype. Supporting Information S1: Figure S4 shows their positioning relative to the global alignment of the CLsat monomers found in *D. raddei nairensis* and *D. valentini*.

2.4 | Combined FISH on *D. unisexualis* confirms the compatibility and specificity of the obtained DNA probes

We computationally predicted and experimentally validated species-specific DNA FISH probes for *D. raddei nairensis* (CLsat30radn) and *D. valentini* (CLsat30val) (Table 1). However, it remained unclear if the probes retain species specificity when used simultaneously in FISH

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

experiments on the hybrid karyotype of *D. unisexualis*. To answer this question, we performed a FISH experiment combining the two probes on the mitotic metaphase plates of *D. unisexualis*. We found that each of these probes hybridized to only a half of the *D. unisexualis* chromosomes and that the signal from the two probes did not overlap (Figure 3). Therefore, a combination of probes CLsat30radn and CLsat30val provides clear differentiation of chromosomal sets inherited by *D. unisexualis* from each of the two parental species.

2.5 | CLsat is a pericentromeric repeat and a universal pericentromeric marker of the *Darevskia* species trio

As we have shown above, both of the species-specific probes that we found and validated originate from the CLsat sequence. Previously, this satellite DNA was characterized as tandemly organized repeated elements (Ciobanu et al., 2004; Roudykh et al., 2002). Subsequent studies identified at least four CLsat families (CLsat I-IV) which had 70%–75% sequence similarity between 18 species of the *Darevskia* genus (Grechko et al., 2006). However, the chromosomal localization of the CLsat repeat has not been investigated. Therefore, our results demonstrate that CLsat is a pericentromeric tandem repeat.

Next, we hypothesized that an evolutionarily conserved fragment within the CLsatI sequence may serve as a universal pericentromeric marker of the three *Darevskia* species considered in this study. To test this hypothesis, we did a DNA FISH experiment on the metaphase plate of *D. unisexualis* using the probe CLsat36 (CLsat-derived 36-bp DNA FISH probe) from the evolutionarily conserved CLsat fragment (Table 1). We found pronounced fluorescent signals in the pericentromeric regions of all 38 chromosomes of this species and a particularly intense signal on the W sex chromosome (Figure 4). Thus, we identified CLsat36 as a

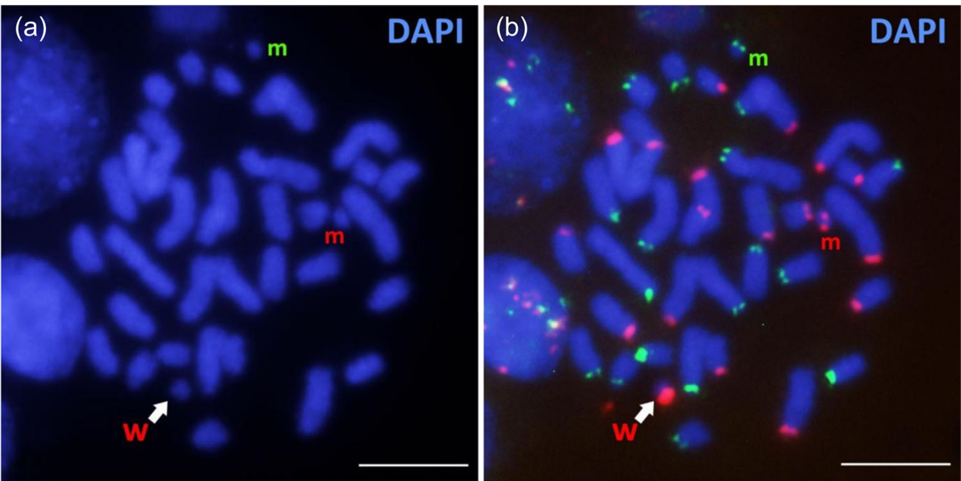

FIGURE 3 DNA fluorescence *in situ* hybridization with oligo-probes CLSat30rad (red) and CLSat30val (green) tested simultaneously on the metaphase plate of the hybrid species *Darevskia unisexualis*. Microchromosomes are marked with “m”; the W chromosome is marked with “w.” (a) Chromatin staining with DAPI. (b) Probes CLSat30rad and CLSat30val together reveal the maternal and paternal chromosomes of the hybrid species *D. unisexualis*. The scale bar is 10 µm.

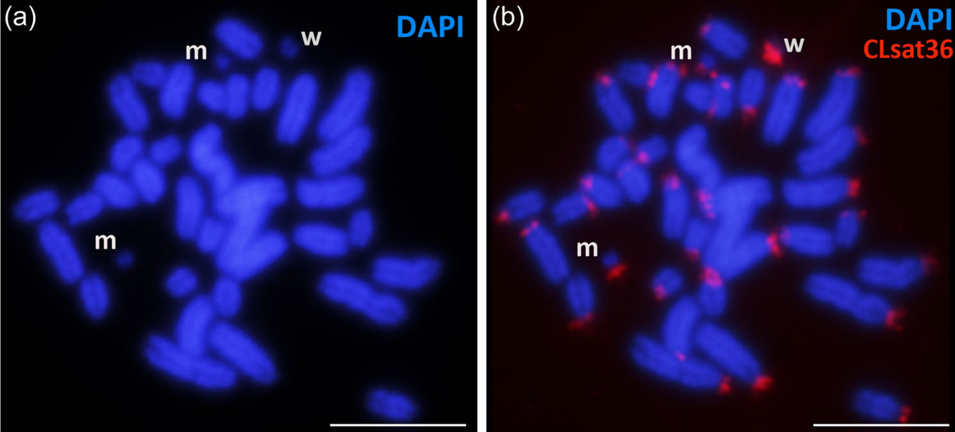

FIGURE 4 DNA fluorescence in situ hybridization with the CLsat36 probe (red) on the metaphase plate of the hybrid species *Darevskia unisexualis*. Chromatin was stained with DAPI (a). All 38 chromosomes demonstrate a pericentromeric fluorescent signal (b). Microchromosomes are marked with “m”; the w sex chromosome is marked with “w”. The scale bar is 10 µm.

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

universal marker of the pericentromeric regions of *D. unisexualis* and, consequently, of the two parental species.

3 | DISCUSSION

Studying composite karyotypes of unisexual hybrid animals holds the potential to advance our understanding of the reproductive mechanisms and evolution of these understudied species. Here, we present the design and verification of the species-specific DNA FISH probes CLsat30radn and CLsat30val for the karyotypes of *D. raddei nairensis* and *D. valentini*, the parental species of a unisexual hybrid lizard *D. unisexualis*. The two probes applied together robustly highlight the corresponding subsets of parental chromosomes within the hybrid karyotype. This fact opens an exciting avenue of studying parental chromosome dynamics. Therefore, the described DNA FISH probes may help to enhance our understanding of the reproductive mechanisms in hybrid animals.

Importantly, these species-specific probes allow rapid chromosome identification in the hybrid karyotype of *D. unisexualis* in a way that is cheaper and less laborious than comparative genomic hybridization (CGH) (Spangenberg, Kolomiets, et al., 2020; Spangenberg, Arakelyan, et al., 2021). Additionally, each of the two probes ensures a clear and species-specific fluorescent signal on the corresponding half of the hybrid karyotype even when used alone, which simplifies DNA FISH experiments aimed to identify the chromosomes of a particular parental species. Overall, the obtained species-specific probes have a great potential to serve as a valuable tool to study the hybrid karyotype of *D. unisexualis*. The fact that the probes CLsat30radn and CLsat30val locate pericentromerically, originate from the CLsat sequence and give a bright fluorescent signal establishes the pericentromeric localization of the CLsat repeats and suggests that CLsat is a major pericentromeric tandem repeat in the *Darevskia* genus, analogous in its abundance and function to the murine major satellite (Komissarov et al., 2011) and the human α satellite (Hasson et al., 2011).

In addition to the species-specific DNA FISH probes, we designed a 36-bp probe, CLsat36, that stains pericentromeric regions of all chromosomes in the hybrid species *Darevskia unisexualis* (Figure 4b). Because CLsat36 locates within a perfectly matching part of a global alignment of the two species-specific CLsat monomers, we speculate that CLsat36 is a universal pericentromeric marker of the karyotypes of the *Darevskia* genus. Of note, both CLsat36 (Figure 4b) and CLsat30radn (Figures 2a and 3b) gave an especially bright signal on the W sex chromosome in the *D. unisexualis* karyotype. This observation is in line with our previous results showing that the W chromosome is easy to identify, in comparison to other chromosomes, by CGH as the chromosome most brightly stained by the maternal genomic DNA (Spangenberg, Arakelyan, et al., 2020, 2021). The higher fluorescent signal on the W chromosome suggests a higher CLsat content, in accord with a well-known high tandem repeat content of nonrecombining sex chromosomes of the Lacertidae lizards (Giovannotti et al., 2018; Suwala et al., 2020) and the Y chromosome in human (Rhie et al., 2023). Importantly, our results suggest that CLsat36 and CLsat30radn could be used as markers of the W sex chromosome, aiding future sex chromosome studies in *Darevskia*.

To predict the species-specific DNA FISH probes, we developed a novel computational method based on the optimization of the Levenshtein distance between the tandem repeat monomers found in the high-coverage contigs of the parental species. This method could help study a wide range of F1 hybrids, both natural and obtained experimentally, including parthenogenetic species of hybrid origin. Importance of such bioinformatic analysis is obvious when molecular biological data is insufficient for deciphering the origin of hybrid karyotypes, and instead chromosome-level studies are necessary. To facilitate the reproduction of our computational results and further development of our method, we implemented it as a Nextflow pipeline, publicly available under the MIT License (https://github.com/nikitin-p/darevskia-pericentromere-analysis). Our method does not require a priori knowledge about the most abundant pericentromeric repetitive sequences in the analyzed genomes and, therefore, could be applied to search for species-specific probes for any pairs of nonmodel species. Additionally, instead of TRMs, one could optimize the Levenshtein distance between all k-mers (candidate probe sequences) derived from the TRMs, in this way eliminating the manual candidate probe selection and making the pipeline fully automatic. Furthermore, although we applied our pipeline to short reads obtained specifically from the pericentromeric chromosomal regions, we believe that it should also be applicable to whole-genome sequencing data. Indeed, we expect that few nonpericentromeric contigs with tandem repeats (satellites) would have a coverage comparable to that of the highest-covered pericentromeric contigs, due to the abundance of tandem repeats in pericentromeres. Consequently, the highly covered contigs with tandem repeats that will serve for the candidate probe selection would come primarily from pericentromeres. Moreover, due to a modular structure of our pipeline, any general-purpose genome assembler could be used instead of TAREAN to obtain contigs for further tandem repeat search. In total, our pipeline could be developed further to turn it into a general and fully automatic standalone software tool.

Finally, we would like to address three irregularities in the results of our data analysis. First of all, based on the available data, we cannot decisively explain the GC-content bias in the *D. valentini* sequencing library. We propose the following possible reasons for it: (1) a subgroup of tandem repeats missed by TRF; (2) a contamination by another species missed by MagicBLAST; (3) dissection of a sex chromosome pericentromere whose sequence properties could considerably diverge from the properties of somatic chromosome pericentromeres due to the lack of recombination in sex chromosomes; (4) accidental dissection of a part of a chromosome arm, in addition to a pericentromeric region. As the options (1) and (2) would be hard to explain, we regard the options (3) and (4) more likely. However, we were not able to reject either of the latter two options, because the annotation of the recently published whole genome assembly of *D. valentini* was not publicly available at the time of the manuscript preparation, precluding the definitive interpretation of any mapping results (Ochkalova et al., 2022).

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Secondly, we found almost nine times more TRMs in the *D. valentini* pericentromeres, than in the *D. raddei nairensis* pericentromeres. The reason for this striking difference could be either technical or biological. Possible technical reasons include (1) incomplete sequencing of the *D. raddei nairensis* pericentromeric DNA; (2) dissection of a sex chromosome in *D. valentini*, as sex chromosomes tend to have an especially large amount of tandem repeats (Giovannotti et al., 2018; Suwala et al., 2020). The biological reason could be a more complex evolutionary history of the *D. valentini* pericentromeres, leading to the proliferation of many different tandem repeats. Further studies are necessary to gain a better understanding of the evolution of these karyotypes and to discriminate between the described possibilities.

Thirdly, we were not able to predict a successful DNA FISH probe for *D. valentini* using the Levenshtein distance optimization and, therefore, had to undertake additional steps to obtain it. As we mentioned in the Results section, this is explained by a low coverage of the CLsat-containing contig in the *D. valentini* assembly. However, the low coverage of a contig that contains a major pericentromeric satellite, in turn, calls for an explanation. As a reason for this discrepancy, we suggest either a bias in the sequenced pericentromeric DNA of *D. valentini* which lead to an underrepresentation of the CLsat-derived reads in this species, or a higher variability of the CLsat sequence in *D. valentini*, than in *D. raddei nairensis*, which precluded the assembly of the majority of CLsat-derived reads in one high-coverage contig.

In total, we present two species-specific DNA FISH probes robustly discriminating parental chromosomes in *D. unisexualis* and a computational pipeline to find such probes and to facilitate future studies of other hybrid karyotypes. In general, whether a significant divergence of pericentromeric repeats triggers incompatibility between chromosomal sets and switches a hybrid animal to unisexual reproduction is an intriguing question. Importantly, evolution of the centromeric regions was proposed as a reason for reproductive isolation more than 20 years ago (Henikoff et al., 2001). The switch to parthenogenetic reproduction in hybrid animals is the way to avoid hybrid sterility (Dedukh et al., 2020). One of the possible reasons for the start of unisexual mode of reproduction could be a large-scale pericentromeric DNA divergence revealed in the *Darevskia* genus (Spangenberg, Arakelyan, et al., 2020, 2021).

4 | MATERIAL AND METHODS

4.1 | Samples

One adult *D. valentini* female was captured in the Sepasar population in June 2019. One adult *D. raddei nairensis* female was captured in the Amberd population in June 2019. One adult *D. unisexualis* individual was captured in the Artsvanist population in October 2019. The specimens were deposited in the research collection of the Yerevan State University (Sepasar population, YSU_291, specimen VS0172; Amberd population, YSU_293 specimen VS0183; Artsvanist population, YSU_294, specimen VS0284).

4.2 | DNA extraction and sequencing

Three pericentromeric regions of mitotic metaphase chromosomes of *D. raddei nairensis* and three pericentromeric regions of mitotic metaphase chromosomes of *D. valentini* were isolated by glass needle-based microdissection with Carl Zeiss AxioVert micromanipulator. The extracted DNA was amplified using DOP-PCR (degenerate oligonucleotide-primed polymerase chain reaction) described in (Liehr 2009, 2017), using the following family of primers: 5′-CCG ACT CGA GNN NNN NAT GTG G-3′. Letter "N" denotes an equal proportion of the A, T, G, and C nucleotides in a given position. After the DOP-PCR procedure, 126-bp paired-end reads were obtained with Illumina HiSeq. 2500.

4.3 | Read preprocessing

We checked the sequencing quality of the obtained paired-end reads using FastQC v0.11.9 (Andrews, 2010). Then, we trimmed and filtered the trimmed reads by length using Trimmomatic v0.39 (Bolger et al., 2014) with an option -phred33 and the following parameters: HEADCROP: 25, ILLUMINACLIP:primer.fa:8:30:10 (where primer.fa contains the sequence of the DOP-PCR primer), ILLUMINACLIP:-TruSeq. 2-PE.fa:2:30:10, SLIDINGWINDOW: 4:20, MINLEN: 20.

4.4 | Contamination assessment

Raw reads were examined for eukaryotic, bacterial, viral and viroid contamination with NCBI Magic-BLAST v1.6.0 (Boratyn et al., 2019) using the following NCBI reference databases: ref_euk_rep_genomes, 16S_ribosomal_RNA, ref_prok_rep_genomes, ref_viruses_rep_genomes, ref_viroids_rep_genomes (https://ftp.ncbi.nlm.nih.gov/blast/db/).

4.5 | Contig assembly and extraction of tandem repeat monomers

Contigs were assembled from the trimmed reads using the seqclust tool from TAREAN v0.3.8 (TAndem REpeat ANalyzer) (Novák et al., 2017) with the following options: -tax METAZOA3.0 -opt ILLUMINA. The assembly quality was evaluated with QUAST v5.0.2 (Gurevich et al., 2013). Tandem repeat monomers (TRMs) were extracted from the obtained contigs with TRF (Tandem Repeats Finder) v4.09 (Benson, 1999), as TAREAN did not find tandem repeats in the assembly. The GC-content of the contigs and TRMs was calculated using the BioPython package (Cock et al., 2009).

4.6 | Prediction of species-specific FISH probes

TRMs were sorted in the descending order of the coverage of the corresponding contigs, where the coverage estimations were taken

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from the TRF output. In this way, we prioritized the most covered repeat monomers, assuming that the coverage correlates with the genomic representation of the repeats and hence can be used to guide the search for highly represented fluorescent FISH probes. To find species-specific probes, we selected the top 10% of the highest-covered repeat monomers (henceforth these selected monomers are termed "the top repeat monomers") in each species and for each of these monomers we calculated its Levenshtein distances (Levenshtein, 1965) to all repeat monomers of the other species. Next, for each top repeat monomer of each species, we calculated its distance to the whole set of monomers of the other species by finding its minimal Levenshtein distance. An approximate Levenshtein distance was calculated with the adist function from the R package utils v3.6.3 (R Core Team, 2023). Then, the top monomers of both species were arranged in the descending order of their distance to the whole set of monomers of the other species, and for each species several top-ranking monomers from these lists were selected to manually extract candidate species-specific 30-bp FISH probes. In our analysis, we used R packages tidyr v1.2.0 (Wickham, Vaughan, et al., 2023), dplyr v1.0.9 (Wickham, François, et al., 2023), and stringr v1.4.0 (Wickham, 2022).

4.7 | Computational filtering of candidate probes

To exclude possible nonspecific probes before doing FISH experiments, we mapped the candidate probes predicted for one species on the contigs of the other species and removed any mapped probes. For this mapping, we used Bowtie2 v2.4.4 (Langmead & Salzberg, 2012). Also, we used the (Thermo Fisher Scientific, Multiple Primer Analyzer) web service to exclude self-hybridizing and cross-hybridizing probes.

4.8 | FISH experiments

We prepared metaphase plates using the standard protocol by (Ford & Hamerton, 1956) with minor modifications described in (Spangenberg et al., 2022). Oligonucleotides corresponding to the candidate species-specific FISH probes for the parental species *D. raddei nairensis* and *D. valentini* and labeled with green and red fluorescent 5′-tags were synthesized by Sintol (Moscow, Russia). We used the TAMRA (5-Carboxytetramethylrhodamine) dye of a red color to mark chromosomes inherited from *D. raddei nairensis* and the FAM (6-Carboxyfluorescein) dye of a green color to mark chromosomes inherited from *D. valentini*. Tandem DNA repeats allowed the use of oligonucleotides with terminal labeling and did not require a signal amplification system. We used the FISH protocol described in Stepakov et al. (2015). Slides were washed in three changes of 2 × standard saline citrate at 30°C, dehydrated in 70%–96% ethanol series, air-dried, and mounted in Vectashield antifade mounting medium with DAPI, 4′,6-diamidino-2-phenylindole (Vector Laboratories). Slides were analyzed using a Carl Zeiss Axio Imager D1 fluorescence light microscope with FS01, FS38HE, and FS43HE filter sets and a monochrome camera Axiocam HRm (Carl Zeiss), using the Genetic Polymorphisms Core Facility of the VIGG RAS (State Contract No 0092-2022-0002). A 36-bp test DNA FISH probe CLsat36 was a kind gift from Vitaly Korchagin (Institute of Gene Biology of the Russian Academy of Sciences). We used it to reveal the chromosomal localization of the repeats of the CLsat family.

4.9 | Schematics and data visualization

In the graphical abstract, we used the silhouette of *Podarcis erhardii* (Bedriaga, 1882) by Alex Slavenko (in public domain) from the PhyloPic database, as there were no *Darevskia* silhouettes in the database at the time of the manuscript preparation: https://www.phylopic.org/images/bf7d9c5f-83c0-435a-b09f-dc6111ece257/podarcis-erhardii. In Figure 1a, we used graphical elements from the nf-core Graphic design guidelines by James A. Fellows Yates and the nf-core team (nf-core Graphic design guidelines). In Figure 1b,c and Supporting Information S1: Figure S3, we used the R package ggplot2 v3.3.6 (Wickham, 2016).

5 | DATA AND CODE AVAILABILITY

Raw sequencing reads were deposited in the NCBI Short Read Archive (SRR25825523 for *D. raddei nairensis*; SRR25825522 for *D. valentini*). The descriptions of the sequenced samples are available in the NCBI BioSample database (SAMN37212597 for *D. raddei nairensis*; SAMN37212598 for *D. valentini*). The NCBI BioProject record associated with this study is PRJNA1011242. The Nextflow pipeline reproducing the computational results described in this paper is available at https://github.com/nikitin-p/darevskia-pericentromere-analysis under the MIT License.

AUTHOR CONTRIBUTIONS

Pavel Nikitin, Sviatoslav Sidorov, and Victor Spangenberg designed experiments. Ahmed Al-Rikabi performed microdissection experiments. Victor Spangenberg performed FISH and microscopy experiments. Marine Arakelyan, Oxana Kolomiets, Thomas Liehr, and Ksenia Klimina provided resources. Pavel Nikitin, Sviatoslav Sidorov, and Vitaly Korchagin analyzed data. Pavel Nikitin, Sviatoslav Sidorov, and Victor Spangenberg wrote the manuscript.

ACKNOWLEDGMENTS

The authors are grateful to A. N. Bogomazova for collaboration. The authors thank the Genetic Polymorphisms Core Facility of the Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow. Sequencing was performed using the core facilities of the Lopukhin FRCC PCM "Genomics, proteomics, metabolomics" (http://rcpcm.org/nauchnye-issledovanija/centr-kollektivnogo-polzovaniia). The authors acknowledge the use of the Makarich HPC cluster at the A.N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University (Moscow, Russia) and the Jex HPC cluster

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Young & Coktas, 2023) at the MRC Laboratory of Medical Sciences (LMS), as well as the resources and support provided by the IT and Bioinformatics Facilities at the MRC LMS. This research was funded by the State Contract of VIGG RAS 0092-2022-0002 (cytogenetic experiments), the Science Committee of Republic of Armenia within the research project No 23-IRF-1F06 (microscopy), and the Russian National Foundation under Grant Agreement No 22-14-00227 (expeditions and keeping the animals).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in NCBI Short Read Archive (SRR25825523 for *D. raddei nairensis*; SRR25825522 for *D. valentini*). The Nextflow pipeline reproducing the computational results is available on GitHub (https://github.com/nikitin-p/darevskia-pericentromere-analysis) under the MIT License. https://github.com.

ETHICS STATEMENT

The manipulations with the animals followed the international rules of the Manual on Humane Use of Animals in Biomedical Research. All experiments were approved by the Ministry of Environment Republic of Armenia, permission No. 3/29.7/1043 in accordance with the Regulations for Laboratory Practice.

ORCID

Pavel Nikitin http://orcid.org/0000-0003-4251-0837
Sviatoslav Sidorov http://orcid.org/0000-0002-1730-265X
Thomas Liehr http://orcid.org/0000-0003-1672-3054
Ksenia Klimina http://orcid.org/0000-0002-5563-644X
Ahmed Al-Rikabi http://orcid.org/0000-0003-4474-4493
Vitaly Korchagin http://orcid.org/0000-0003-2264-6294
Oxana Kolomiets http://orcid.org/0000-0002-1915-0039
Marine Arakelyan http://orcid.org/0000-0002-6334-5714
Victor Spangenberg http://orcid.org/0000-0002-6623-9124

REFERENCES

Ahmad, S. F., Singchat, W., Jehangir, M., Suntronpong, A., Panthum, T., Malaivijitnond, S., & Srikulnath, K. (2020). Dark matter of primate genomes: Satellite DNA repeats and their evolutionary dynamics. *Cells*, *9*, 2714. https://doi.org/10.3390/cells9122714

Altemose, N., Logsdon, G. A., Bzikadze, A. V., Sidhwani, P., Langley, S. A., Caldas, G. V., Hoyt, S. J., Uralsky, L., Ryabov, F. D., Shew, C. J., Sauria, M. E. G., Borchers, M., Gershman, A., Mikheenko, A., Shepelev, V. A., Dvorkina, T., Kunyavskaya, O., Vollger, M. R., Rhie, A., … Miga, K. H. (2022). Complete genomic and epigenetic maps of human centromeres. *Science*, *376*, eabl4178. https://doi.org/10.1126/science.abl4178

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data [Online]. Accessed February 12, 2024. Available from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Barley, A. J., Reeder, T. W., Nieto-Montes De Oca, A., Cole, C. J., & Thomson, R. C. (2021). A new diploid parthenogenetic whiptail lizard from sonora, Mexico, is the "missing link" in the evolutionary transition to polyploidy. *The American Naturalist*, *198*, 295–309. https://doi.org/10.1086/715056

Benson, G. (1999). Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, *27*, 573–580. https://doi.org/10.1093/nar/27.2.573

Biscotti, M. A., Olmo, E., & Heslop-Harrison, J. S. (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Research*, *23*, 415–420. https://doi.org/10.1007/s10577-015-9499-z

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boratyn, G. M., Thierry-Mieg, J., Thierry-Mieg, D., Busby, B., & Madden, T. L. (2019). Magic-BLAST, an accurate RNA-seq aligner for long and short reads. *BMC Bioinformatics*, *20*, 405. https://doi.org/10.1186/s12859-019-2996-x

Ciobanu, D., Grechko, V. V., Darevsky, I. S., & Kramerov, D. A. (2004). New satellite DNA in Lacerta s. str. lizards (Sauria: Lacertidae): Evolutionary pathways and phylogenetic impact. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *302*(6), 505–516. https://doi.org/10.1002/jez.b.21014

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & De Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, *25*, 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

Dedukh, D., Altmanová, M., Klíma, J., & Kratochvíl, L. (2022). Premeiotic endoreplication is essential for obligate parthenogenesis in geckos. *Development*, *149*, dev200345. https://doi.org/10.1242/dev.200345

Dedukh, D., Majtánová, Z., Marta, A., Pšenička, M., Kotusz, J., Klíma, J., Juchno, D., Boron, A., & Janko, K. (2020). Parthenogenesis as a solution to hybrid sterility: The mechanistic basis of meiotic distortions in clonal and sterile hybrids. *Genetics*, *215*, 975–987. https://doi.org/10.1534/genetics.119.302988

Ford, C. E., & Hamerton, J. L. (1956). The chromosomes of man. *Nature*, *178*, 1020–1023. https://doi.org/10.1038/1781020a0

Freitas, S., Rocha, S., Campos, J., Ahmadzadeh, F., Corti, C., Sillero, N., Ilgaz, Ç., Kumlutaş, Y., Arakelyan, M., Harris, D. J., & Carretero, M. A. (2016). Parthenogenesis through the ice ages: A biogeographic analysis of Caucasian rock lizards (genus Darevskia. *Molecular Phylogenetics and Evolution*, *102*, 117–127. https://doi.org/10.1016/j.ympev.2016.05.035

Fu, J., Murphy, R. W., & Darevsky, I. S. (2000). Divergence of the cytochrome *b* gene in the *Lacerta raddei* complex and its parthenogenetic daughter species: Evidence for recent multiple origins. *Copeia*, *2000*, 432–440. https://doi.org/10.1643/0045-8511(2000)000[0432:DOTCBG]2.0.CO;2

Giovannotti, M., Nisi Cerioni, P., Caputo, V., & Olmo, E. (2009). Characterisation of a GC-rich telomeric satellite DNA in *Eumeces schneideri* Daudin (Reptilia, Scincidae). *Cytogenetic and Genome Research*, *125*, 272–278. https://doi.org/10.1159/000235933

Giovannotti, M., Nisi Cerioni, P., Rojo, V., Olmo, E., Slimani, T., Splendiani, A., & Caputo Barucchi, V. (2018). Characterization of a satellite DNA in the genera *Lacerta* and *Timon* (Reptilia, Lacertidae) and its role in the differentiation of the W chromosome. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *330*, 83–95. https://doi.org/10.1002/jez.b.22790

Giovannotti, M., Nisi Cerioni, P., Splendiani, A., Ruggeri, P., Olmo, E., & Caputo Barucchi, V. (2013). Slow evolving satellite DNAs: The case of a centromeric satellite in Chalcides ocellatus (Forskål, 1775) (Reptilia, Scincidae). *Amphibia-Reptilia*, *34*, 401–411. https://doi.org/10.1163/15685381-00002905

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Grechko, V. V., Ciobanu, D. G., Darevsky, I. S., Kosushkin, S. A., & Kramerov, D. A. (2006). Molecular evolution of satellite DNA repeats and speciation of lizards of the genus Darevskia (Sauria: Lacertidae). *Genome*, *49*, 1297–1307. https://doi.org/10.1139/g06-089

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, *29*, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hasson, D., Alonso, A., Cheung, F., Tepperberg, J. H., Papenhausen, P. R., Engelen, J. J. M., & Warburton, P. E. (2011). Formation of novel CENP-A domains on tandem repetitive DNA and across chromosome breakpoints on human chromosome 8q21 neocentromeres. *Chromosoma*, *120*, 621–632. https://doi.org/10.1007/s00412-011-0337-6

Henikoff, S., Ahmad, K., & Malik, H. S. (2001). The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science*, *293*, 1098–1102. https://doi.org/10.1126/science.1062939

Komissarov, A. S., Gavrilova, E. V., Demin, S. J., Ishov, A. M., & Podgornaya, O. I. (2011). Tandemly repeated DNA families in the mouse genome. *BMC Genomics*, *12*, 531. https://doi.org/10.1186/1471-2164-12-531

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

Levenshtein, V. I. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Doklady Akademii Nauk SSSR*, *163*, 845–848. http://mi.mathnet.ru/dan31411

Liehr, T. (Ed.). (2009). *Fluorescence in situ hybridization (FISH) —Application guide*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70581-9

Liehr, T. (Ed.). (2017). *Fluorescence in situ hybridization (FISH): Application guide*, *Springer protocols handbooks*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-52959-1

Liehr, T. (2022). *Cytogenetics and molecular cytogenetics* (1st ed.). CRC Press. https://doi.org/10.1201/9781003223658

Lisachov, A., Rumyantsev, A., Prokopov, D., Ferguson-Smith, M., & Trifonov, V. (2023). Conservation of major satellite DNAs in snake heterochromatin. *Animals: An Open Access Journal from MDPI*, *13*, 334. https://doi.org/10.3390/ani13030334

Matsubara, K., Uno, Y., Srikulnath, K., Seki, R., Nishida, C., & Matsuda, Y. (2015). Molecular cloning and characterization of satellite DNA sequences from constitutive heterochromatin of the habu snake (Protobothrops flavoviridis, Viperidae) and the Burmese python (Python bivittatus, Pythonidae). *Chromosoma*, *124*, 529–539. https://doi.org/10.1007/s00412-015-0529-6

Murphy, R. W., Fu, J., Macculloch, R. D., Darevsky, I. S., & Kupriyanova, L. A. (2000). A fine line between sex and unisexuality: The phylogenetic constraints on parthenogenesis in lacertid lizards. *Zoological Journal of the Linnean Society*, *130*, 527–549. https://doi.org/10.1111/j.1096-3642.2000.tb02200.x

Murtskhvaladze, M., Tarkhnishvili, D., Anderson, C. L., & Kotorashvili, A. (2020). Phylogeny of caucasian rock lizards (Darevskia) and other true lizards based on mitogenome analysis: Optimisation of the algorithms and gene selection. *PLoS One*, *15*, e0233680. https://doi.org/10.1371/journal.pone.0233680

Newton, A. A., Schnittker, R. R., Yu, Z., Munday, S. S., Baumann, D. P., Neaves, W. B., & Baumann, P. (2016). Widespread failure to complete meiosis does not impair fecundity in parthenogenetic whiptail lizards. *Development*, *143*(23), 4486–4494. https://doi.org/10.1242/dev.141283

nf-core Graphic design guidelines. https://nf-co.re/docs/contributing/design_guidelines

Novák, P., Ávila Robledillo, L., Koblížková, A., Vrbová, I., Neumann, P., & Macas, J. (2017). TAREAN: A computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Research*, *45*, e111. https://doi.org/10.1093/nar/gkx257

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., … Phillippy, A. M. (2022). The complete sequence of a human genome. *Science*, *376*, 44–53. https://doi.org/10.1126/science.abj6987

Ochkalova, S., Korchagin, V., Vergun, A., Urin, A., Zilov, D., Ryakhovsky, S., Girnyk, A., Martirosyan, I., Zhernakova, D. V., Arakelyan, M., Danielyan, F., Kliver, S., Brukhin, V., Komissarov, A., & Ryskov, A. (2022). First genome of rock lizard Darevskia valentini involved in formation of several parthenogenetic species. *Genes*, *13*, 1569. https://doi.org/10.3390/genes13091569

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rhie, A., Nurk, S., Cechova, M., Hoyt, S. J., Taylor, D. J., Altemose, N., Hook, P. W., Koren, S., Rautiainen, M., Alexandrov, I. A., Allen, J., Asri, M., Bzikadze, A. V., Chen, N.-C., Chin, C.-S., Diekhans, M., Flicek, P., Formenti, G., Fungtammasan, A., … Phillippy, A. M. (2023). The complete sequence of a human Y chromosome. *Nature*, *621*, 344–354. https://doi.org/10.1038/s41586-023-06457-y

Rojo, V., Martínez-Lage, A., Giovannotti, M., González-Tizón, A. M., Cerioni, P. N., Barucchi, V. C., Galán, P., Olmo, E., & Naveira, H. (2015). Evolutionary dynamics of two satellite DNA families in rock lizards of the genus Iberolacerta (Squamata, Lacertidae): Different histories but common traits. *Chromosome Research*, *23*, 441–461. https://doi.org/10.1007/s10577-015-9489-1

Roudykh, I. A., Grechko, V. V., Ciobanu, D. G., Kramerov, D. A., & Darevsky, I. S. (2002). Variability of restriction sites in satellite DNA as a molecular basis of taxonprint method: Evidence from the study of Caucasian rock lizards. *Russian Journal of Genetics*, *38*, 937–941. https://doi.org/10.1023/A:1016839929007

Spangenberg, V., Arakelyan, M., Cioffi, M. D. B., Liehr, T., Al-Rikabi, A., Martynova, E., Danielyan, F., Stepanyan, I., Galoyan, E., & Kolomiets, O. (2020). Cytogenetic mechanisms of unisexuality in rock lizards. *Scientific Reports*, *10*, 8697. https://doi.org/10.1038/s41598-020-65686-7

Spangenberg, V., Arakelyan, M., Galoyan, E., Martirosyan, I., Bogomazova, A., Martynova, E., Bello Cioffi, M., Liehr, T., Al-Rikabi, A., Osipov, F., Petrosyan, V., & Kolomiets, O. (2021). Meiotic synapsis of homeologous chromosomes and mismatch repair protein detection in the parthenogenetic rock lizard *Darevskia unisexualis*. *Molecular Reproduction and Development*, *88*, 119–127. https://doi.org/10.1002/mrd.23450

Spangenberg, V., Kolomiets, O., Stepanyan, I., Galoyan, E., De Bello Cioffi, M., Martynova, E., Martirosyan, I., Grishaeva, T., Danielyan, F., Al-Rikabi, A., Liehr, T., & Arakelyan, M. (2020). Evolution of the parthenogenetic rock lizard hybrid karyotype: Robertsonian translocation between two maternal chromosomes in *Darevskia rostombekowi*. *Chromosoma*, *129*, 275–283. https://doi.org/10.1007/s00412-020-00744-7

Spangenberg, V., Losev, M., Volkhin, I., Smirnova, S., Nikitin, P., & Kolomiets, O. (2021). DNA environment of centromeres and non-homologous chromosomes interactions in mouse. *Cells*, *10*, 3375. https://doi.org/10.3390/cells10123375

Spangenberg, V., Redekop, I., Simanovsky, S. A., & Kolomiets, O. (2022). Cytogenetic analysis of the bimodal karyotype of the common european adder, *Vipera berus* (Viperidae). *Animals: An Open Access Journal from MDPI*, *12*, 3563. https://doi.org/10.3390/ani12243563

Stepakov, A., Galkina, S., Bogomaz, D., Gaginskaya, E., & Saifitdinova, A. (2015). Modified synthesis of 6-carboxyfluorescein (6-FAM): Application to probe labeling for conventional cytogenetics. *British Journal of Applied Science & Technology*, *7*(4), 423–428.

Suwala, G., Altmanová, M., Mazzoleni, S., Karameta, E., Pafilis, P., Kratochvíl, L., & Rovatsos, M. (2020). Evolutionary variability of

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

W-linked repetitive content in lacertid lizards. *Genes*, *11*, 531. https://doi.org/10.3390/genes11050531

Thermo Fisher Scientific, Multiple Primer Analyzer. www.thermofisher.com/ru/ru/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html

Tokarskaya, O., Kan, N., Petrosyan, V., Martirosyan, I., Grechko, V., Danielyan, F., Darevsky, I., & Ryskov, A. (2001). Genetic variation in parthenogenetic Caucasian rock lizards of the genus Lacerta (L. dahli, L. armeniaca, L. unisexualis) analyzed by DNA fingerprinting. *Molecular Genetics and Genomics*, *265*, 812–819. https://doi.org/10.1007/s004380100475

Uzzell, T., & Darevsky, I. S. (1975). Biochemical evidence for the hybrid origin of the parthenogenetic species of the Lacerta saxicola complex (Sauria: Lacertidae), with a discussion of some ecological and evolutionary implications. *Copeia*, *1975*, 204. https://doi.org/10.2307/1442879

Wickham, H. (2016). *Ggplot2*. Springer Science+Business Media, LLC [Online]. Accessed February 12, 2024. Available from https://ggplot2.tidyverse.org

Wickham, H. (2022). stringr: Simple, consistent wrappers for common string operations [Online]. Accessed February 12, 2024. Available from https://stringr.tidyverse.org

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A grammar of data manipulation [Online]. Accessed February 12, 2024. Available from https://dplyr.tidyverse.org

Wickham, H., Vaughan, D., & Girlich, M. (2023). tidyr: Tidy messy data [Online]. Available from https://tidyr.tidyverse.org

Young, G., & Coktas, A. (2023). Jex: High performance computing at the MRC laboratory of medical sciences. *Zenodo*. https://doi.org/10.5281/zenodo.10245364

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Nikitin, P., Sidorov, S., Liehr, T., Klimina, K., Al-Rikabi, A., Korchagin, V., Kolomiets, O., Arakelyan, M., & Spangenberg, V. (2024). Variants of a major DNA satellite discriminate parental subgenomes in a hybrid parthenogenetic lizard Darevskia unisexualis (Darevsky, 1966). *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 1–12. https://doi.org/10.1002/jez.b.23244

15525015, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jez.b.23244 by IMPERIAL COLLEGE OF LONDON, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License