

# Complete mitochondrial genomes from museum specimens clarify millipede evolution in the Eastern Arc Mountains

MARTIN NIELSEN<sup>1,2,\*</sup>, ASHOT MARGARYAN<sup>1</sup>, TEJS LIND NIELSEN<sup>3</sup>,  
HENRIK ENGHOFF<sup>4</sup> and MORTEN E. ALLENTOF<sup>3,5,\*</sup>

<sup>1</sup>Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

<sup>2</sup>Arctic Station, Faculty of Science, University of Copenhagen, Disko Island, Greenland

<sup>3</sup>Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

<sup>4</sup>Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

<sup>5</sup>Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Perth, WA, Australia

Received 7 December 2021; revised 30 May 2022; accepted for publication 9 June 2022

The Eastern Arc Mountains in Tanzania represent a hotspot for biological diversity of global importance. The level of endemism is high, and Eastern Arc biodiversity has been studied extensively in vertebrates and invertebrates, including millipedes. However, millipede evolution is vastly understudied at the molecular level. Therefore, we used next-generation ‘shotgun’ sequencing to obtain mitochondrial genome sequences of 26 museum specimens, representing six genera and 12 millipede species found across the Eastern Arc Mountains. Bayesian and maximum likelihood methods yielded consistent topologies with high node support, confirming a high level of congruence between molecular and morphological analyses. The only exception was a *Tropostreptus sigmatospinus* individual from Zanzibar, which was placed outside an otherwise monophyletic group consisting of mainland individuals of the same assumed species. For two species with a distribution across several mountain blocks (*Tropostreptus sigmatospinus* and *Tropostreptus hamatus*), each mountain population represents a distinct monophyletic lineage. In contrast, we also observe that distinct species exist sympatrically in the same montane forests, indicative of older speciation events that are not defined by current forest distribution. Our results are important for understanding speciation mechanisms in montane rain forests and highlight that ethanol-preserved invertebrates exhibit a tremendous potential for genomic analyses.

ADDITIONAL KEYWORDS: Diplopoda – mitogenomes – museomics – shotgun sequencing – Tanzania.

## INTRODUCTION

The Eastern Arc Mountains represent a globally important biodiversity hotspot recognized for its high concentration of endemic species (e.g. Burgess *et al.*, 2007). With > 400 endemic plant species and ≥ 96 endemic vertebrates distributed across ten mammals, 19 birds, 29 reptiles and 38 amphibian species, the area is highlighted in many major analyses and

recommendations for global biodiversity priorities (Burgess *et al.*, 2007). As early as the 1970s, the Eastern Arc Mountains were already identified as a centre of endemism (Hopkins, 1987), and they are listed as a Global 200 Ecoregion by the World Wildlife Fund (Olson & Dinerstein, 1998). They form part of a biodiversity hotspot ecoregion defined by Conservation International (Mittermeier *et al.*, 1998, 2004) and are recognized by BirdLife International as an Endemic Bird Area (Bibby *et al.*, 1992; Stattersfield *et al.*, 1998).

The Eastern Arc comprises 13 separate mountain formations with ~3300 km<sup>2</sup> of montane forest extending

\*Corresponding authors. E-mail: [mnapieceofnature@gmail.com](mailto:mnapieceofnature@gmail.com); [morten.allentoft@gmail.com](mailto:morten.allentoft@gmail.com)

southwards from the Taita Hills in Kenya to include a dozen isolated ranges in Tanzania. It is believed that a massive pan-African rain forest existed in the Eocene across tropical East Africa, but owing to climate change and aridification, this forest began to break up in the Oligocene–Early Miocene (Couvreur *et al.*, 2008). East African lowland forest became savannah, with isolated ‘islands’ of forests, possibly resulting in numerous vicariance events, driven by local adaptations or genetic drift in isolated populations (Couvreur *et al.*, 2008). Today, the Eastern Arc Mountains act as refugia for these ancient forests, presumably preserving reservoirs of old forest biodiversity. Moreover, East African rain forests might potentially have expanded and contracted multiple times, resulting in repeated cycles of species diversifications (Fjeldså & Bowie, 2008; Measey & Tolley, 2011; Papadopoulou & Knowles, 2015), and such a ‘species pump’ could explain why the Eastern Arc montane forests today have some of the highest densities of endemic plants and animals in the world. Evidence for vicariance in the Eastern Arc Mountains is plentiful, and many endemic species are found on a single mountain range only. The diversity of tree species in East Africa has been linked to vicariance events owing to the pan-African forest fragmentation since the Oligocene (Couvreur *et al.*, 2008), and a similar pattern has been proposed for some amphibians (Evans *et al.*, 2004). Bird diversity also appears to be consistent with repeated vicariance and dispersal events (Fjeldså & Bowie, 2008). A study on African violets (*Streptocarpus* Lindl. p.p., formerly *Saintpaulia* H. Wendl.; see Christenhusz, 2012; Nishii *et al.*, 2015) suggested a role for mountain refugia in driving speciation (Dimitrov *et al.*, 2012), and a study on species of laminate-toothed rats (*Otomys* F. Cuvier, 1824 spp.) supported the importance of palaeoclimate in modern speciation (Taylor *et al.*, 2009). Although vicariance (i.e. isolation) is widely accepted as a key driver for speciation in this region, conflicting data relate to the timing of the radiation in different lineages, hence their potential links to climatic and geological events. It is clear that the Eastern Arc is a perfect region of study for obtaining insights into fundamental evolutionary processes responsible for biodiversity on our planet.

Millipedes (Diplopoda) are terrestrial arthropods belonging to the guild of decomposers that are often abundant in temperate and tropical forest ecosystems. They are not vagile animals, and the distribution of millipede at species, genus or even family and order level is often very geographically restricted. Therefore, ancient biogeographical patterns, such as those shaped by vicariance in fragmenting rain forests, are expected to be preserved (Shelley & Golovatch, 2011; Enghoff, 2015). This seems to be true even at high taxonomic (family) levels in diplopods (e.g. Wesener & VandenSpiegel, 2009; Wesener, 2014).

The Eastern Arc region has a large diversity of millipedes that, to a great extent, is not yet scientifically documented. This is particularly true for the Udzungwa Mountains, which are the largest mountain range in the Eastern Arc. One family, Odontopygidae, has been subject to recent comprehensive morphological investigations (Enghoff, 2020). Of the 41 odontopygid species known from the Udzungwa Mountains, 36 are known from nowhere else and can be regarded as endemic (Enghoff, 2020). A recent morphological analysis of specimens from other Eastern Arc Mountain ranges has added another 25 new species, known from only one mountain range each (Enghoff, 2022). The Udzungwa species of the genus *Eviulisoma* Silvestri, 1910 (Paradoxosomatidae) were studied by Enghoff (2018), who found 22 species, all known from nowhere else, with 20 being new to science.

Spirostreptidae millipedes from the Eastern Arc have not been studied systematically, apart from *Tropostreptus* Enghoff, 2017, which is an example of a millipede genus known only from Tanzania and almost exclusive to the Eastern Arc Mountains, with several endemic species linked to specific mountain ranges. *Tropostreptus* species are forest dwellers with a medium-to-large body size (body length 5.5–14.0 cm; Fig. 1). The genus includes seven described species. *Tropostreptus austerus* (Attems, 1950) is endemic to the East Usambara Mountains, *Tropostreptus severus* Enghoff, 2017 to the Nguru Mountains, *Tropostreptus microcephalus* Enghoff, 2017 and *Tropostreptus droides* Enghoff, 2017 are both endemic to the Udzungwa Mountains and *Tropostreptus kipunji* Enghoff, 2017 to Mount Rungwe. Two species show wider distributions: *Tropostreptus hamatus* (Demange, 1970) is known from the East Usambara, Nguru, Rubeho, Uluguru and Udzungwa Mountains, and *Tropostreptus*



**Figure 1.** Typical *Tropostreptus* appearance exemplified by a *Tropostreptus hamatus* individual from Udzungwa Mountains, Tanzania (photograph credit: Nikolaj Scharff).

*sigmatospinus* Enghoff, 2017 is known from the Nguru, Uluguru and Udzungwa Mountains, the Rondo Plateau in south-east Tanzania and from the island of Zanzibar off the Tanzanian coast (Enghoff, 2017).

Although millipedes are among the most diverse groups of arthropods, encompassing thousands of species and occupying important niches in the ecosystem, little is known about their ecology and evolution. Taxonomic and phylogenetic analyses of millipedes are often based on morphological traits, hence our current knowledge at the molecular level remains limited. Few molecular studies have been published on millipede evolution (e.g. Pitz & Sierwald, 2010; Enghoff *et al.*, 2011), and before the present study only 13 full mitochondrial genomes of Diplopoda (divided across eight orders and ten families) have been deposited in GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>, accessed February 2020; Supporting Information, Table S1). As such, molecular data have largely been neglected in studies of millipedes, and there is an urgent need to adopt these methods for testing our current morphology-based phylogenies and for understanding millipede evolution in greater detail (Brewer *et al.*, 2012). Forest species, such as *Tropostreptus* in the Eastern Arc, have probably been restricted to these ancient montane forests for a long time, making them excellent model organisms for understanding millipede evolution in a biodiversity hotspot and for testing congruence between morphology-based and DNA-based millipede phylogenies.

In this study, we analyse 28 millipede museum specimens with next-generation sequencing technology. To provide an in-depth study of molecular millipede evolution in the Eastern Arc Mountains, we focus on the genus *Tropostreptus*. We present new molecular data from 12 different millipede species from Tanzania, including the seven known species of *Tropostreptus*, and especially the Eastern Arc Mountains. Based on ‘shotgun’ data, we assemble their full mitochondrial genomes, and by combining this dataset with 12 previously published millipede mitogenomes, we extend our phylogenetic insights beyond the Eastern Arc. Additionally, this study serves as a test case for assessing the molecular scientific potential and level of DNA degradation in old ethanol-preserved Diplopoda specimens, in addition to testing a previously published non-destructive DNA extraction approach (Nielsen *et al.*, 2019) on millipedes.

Lastly, we examine our genetic sequence data for the presence of the fungus *Rickia gigas* Santamaria, Enghoff & Reboleira, 2016. Several species of *Tropostreptus* (*T. austerus*, *T. droides*, *T. hamatus*, *T. severus* and *T. sigmatospinus*) are known to be host to this epizootic (possibly parasitic) fungus. A larger morph of what is regarded as the same species

of fungus has also been found on the giant East African millipede, *Archspirostreptus gigas* (Peters, 1855) (Santamaria *et al.*, 2016).

## MATERIAL AND METHODS

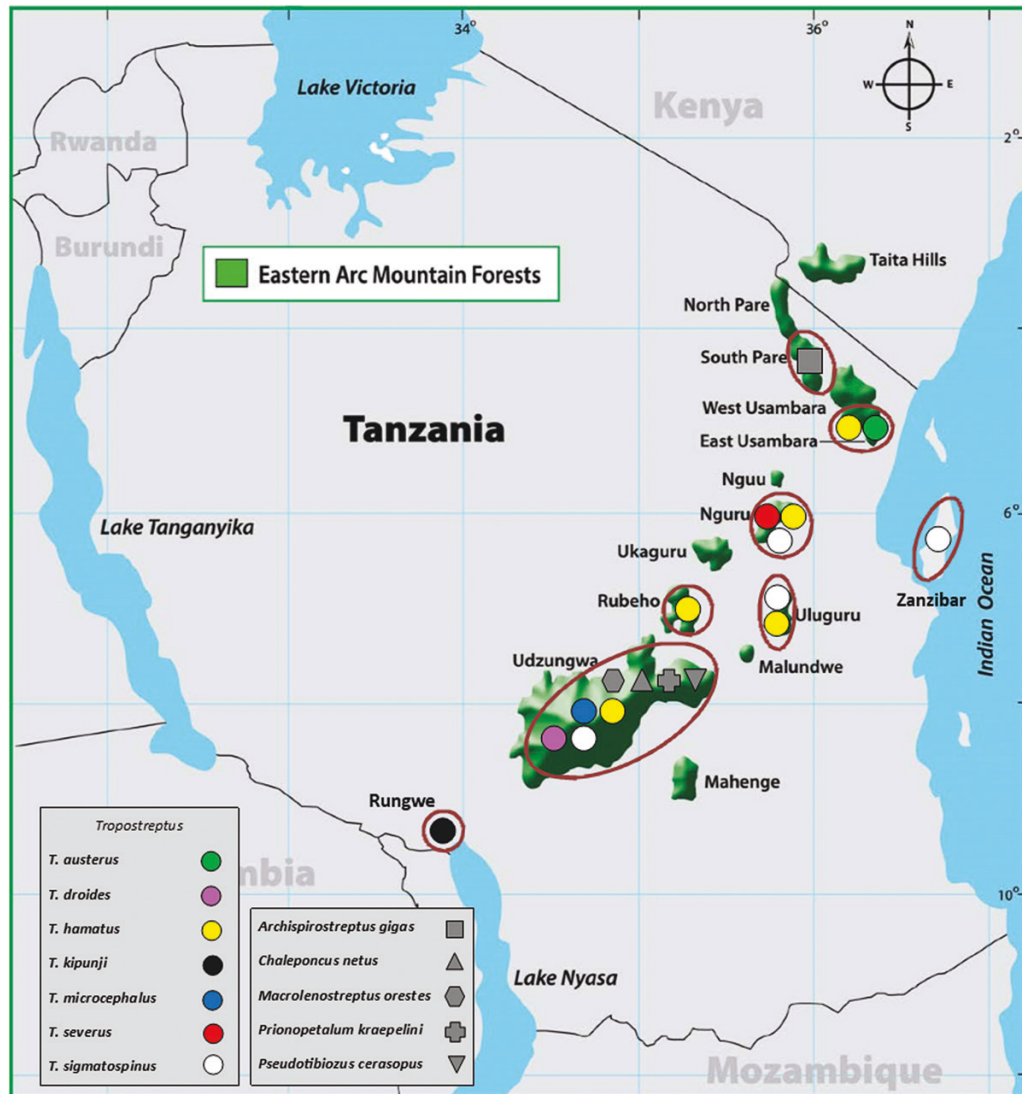
A total of 28 millipede museum specimens were sampled for this study. Twenty-three of these belonged to the genus *Tropostreptus* (Spirostreptidae) and included seven species. These had been collected in Tanzania between 1984 and 2016. An additional five specimens representing five species were included in our analyses, namely *Ar. gigas* (family Spirostreptidae), *Chaleponcus netus* Enghoff, 2014 (family Odontopygidae), *Macrolenostreptus orestes* Hoffman & Howell, 1996 (family Spirostreptidae), *Prionopetalum kraepelini* Attems, 1896 (family Odontopygidae) and *Pseudotibiozus cerasopus* Attems, 1914 (family Spirostreptidae). For details regarding sample origin and collecting dates, see Figure 2 and Table 1.

### DNA EXTRACTION AND SEQUENCING

All specimens were stored initially in 70% ethanol at the Natural History Museum of Denmark. The DNA extractions were carried out following a non-destructive protocol (Nielsen *et al.*, 2019) modified from the one described by Gilbert *et al.* (2007). The protocol aims at extracting DNA from the internal soft tissues of arthropods by submerging them in an incubation buffer, thereby leaving the exoskeleton intact. Entire specimens were used for DNA extraction, ensuring that the incubation buffer covered the specimen fully. The buffer consisted of 10 mmol/L Tris-HCl (pH 8.0), 10 mmol/L NaCl, 5 mmol/L CaCl<sub>2</sub>, 2.5 mmol/L EDTA, 1% sodium dodecyl sulphate, 1% proteinase K, 40 mmol/L dithiothreitol and molecular biology grade H<sub>2</sub>O. After the addition of buffer, the specimens were left at 56 °C overnight. They were then placed in 96% ethanol for a few hours to stop further digestion, before being transferred to 70% ethanol and returned to the museum collection.

The DNA was isolated from the incubation buffer using the QiaQuick PCR Purification Kit (Qiagen), following the manufacturer’s protocol apart from an extended 15 min incubation step at 37 °C before eluting in 50 µL EB buffer. Recovering *Rickia* DNA can be challenging (Haelewaters *et al.*, 2015; Sundberg *et al.*, 2018). Purified DNA extracts were therefore placed finally at –80 °C for ~30 min, representing a ‘deep-freeze’ step that has previously been found beneficial for extracting fungal DNA from ants (J. Agerbo Rasmussen, unpublished data). This is thought potentially to break the tough cell wall of the fungi and release the DNA,





**Figure 2.** Map showing the origin of the millipede specimens used in the study, with the accuracy of location restricted to mountain blocks. Coloured circles all represent *Tropostreptus* species, whereas grey symbols represent species from other millipede genera. Base map published by permission of the Eastern Arc Mountains Conservation Endowment Fund.

making it available for later amplification. The DNA extractions were carried out in a pre-PCR laboratory to minimize risk of contamination.

The majority of extracts (all *Tropostreptus* extracts and *C. netus*) were then prepared for sequencing using the single-tube library protocol (Carøe *et al.*, 2018). In short, in this protocol the DNA end-repair, the adapter ligation and the adapter fill-in reaction are all performed in the same tube. Products were then purified using bead purification (AMPure beads; Beckman Coulter, USA). A qPCR was initially carried out to estimate the number of PCR cycles required in the following index PCR, which was performed in reaction volumes of 100 µL consisting of 10× buffer, 2.5 mM MgCl<sub>2</sub>, 0.2 mM dNTP, 0.1 U/µL AmpliTaq Gold

(Applied Biosystems), 0.2 µM forward primer (BGI) and 0.2 µM index primer (BGI). The PCR amplifications were run at 95 °C for 12 min, followed by 9–13 cycles at 95 °C for 20 s, 60 °C for 30 s and 72 °C for 40 s, with a final extension step of 72 °C for 5 min.

The remaining four DNA extracts (representing *Ar. gigas*, *M. orestes*, *Pr. kraepelini* and *Ps. cerasopus*) were prepared for sequencing as follows. The DNA was fragmented to a peak size of 550 bp using a Covaris M220 (Covaris, Woburn, MA, USA), and blunt-ended, double-stranded libraries were constructed using the NEBNext DNA Sample Prep kit E6070 (New England Biolabs, Ipswich, MA, USA) with Illumina-specific adapters (Illumina, San Diego, CA, USA), following published guidelines (for details, see Allentoft *et al.*,

**Table 1.** Overview of samples, their origin, year of collection, mitochondrial genome size, GenBank accession number and museum voucher number. NHMD, Natural History Museum of Denmark

Species	Locality	Year of collection	Mitogenome size	GenBank accession number	Museum voucher (NHMD)
<i>Tropostreptus austerus</i>	East Usambara	2014	15 261	MT394523	621801
<i>Tropostreptus droides</i>	Udzungwa Scarp Forest Reserve	1984	15 172	MT394522	621802
<i>Tropostreptus droides</i> *	Udzungwa Scarp Forest Reserve	1984	–	–	621803
<i>Tropostreptus hamatus</i>	Udzungwa Mountains National Park	2016	15 151	MT394510	621804
<i>Tropostreptus hamatus</i>	Rubebo	1993	15 150	MT394521	621805
<i>Tropostreptus hamatus</i>	Uluguru	2000	15 156	MT394508	621806
<i>Tropostreptus hamatus</i>	Udzungwa, Ndundulu	1993	15 150	MT394520	621807
<i>Tropostreptus hamatus</i> *	Udzungwa Scarp Forest Reserve	1998	–	–	621808
<i>Tropostreptus hamatus</i>	Udzungwa Mountains National Park	2014	15 150	MT394514	621809
<i>Tropostreptus hamatus</i>	Nguru, Kanga forest	1984	15 152	MT394509	621810
<i>Tropostreptus hamatus</i>	East Usambara	1992	15 155	MT394502	621811
<i>Tropostreptus kipunji</i>	Rungwe	1996	15 170	MT394503	621812
<i>Tropostreptus kipunji</i>	Rungwe	1996	15 170	MT394511	621813
<i>Tropostreptus microcephalus</i>	Udzungwa, Ndundulu	2007	15 169	MT394516	621814
<i>Tropostreptus severus</i>	Nguru, Kanga forest	1984	15 209	MT394517	621815
<i>Tropostreptus sigmatospinus</i>	Zanzibar	1981	15 163	MT394519	621816
<i>Tropostreptus sigmatospinus</i>	Udzungwa, West Kilombero Scarp Forest Reserve	2000	15 172	MT394507	621817
<i>Tropostreptus sigmatospinus</i>	Uluguru, chamananyi	2000	15 174	MT394518	621818
<i>Tropostreptus sigmatospinus</i>	Udzungwa, Lower Khihansi	1997	15 002†	MT394505	621819
<i>Tropostreptus sigmatospinus</i>	Matundu	1996	15 172	MT394526	621820
<i>Tropostreptus sigmatospinus</i>	Udzungwa Mountains National Park	2014	15 174	MT394515	621821
<i>Tropostreptus sigmatospinus</i>	Udzungwa, Nyambanike	1994	15 171	MT394501	621822
<i>Tropostreptus sigmatospinus</i>	Nguru, Kanga forest	1989	15 176	MT394504	621823
<i>Archispirostreptus gigas</i>	South Pare Mountains, Mwala	2017	15 177	MT394525	621924
<i>Chaleponcus netus</i>	Udzungwa, West Kilombero Scarp Forest Reserve	2006	15 093	MT394513	621825
<i>Macrolenostreptus orestes</i>	Udzungwa Mountains National Park	2014	15 367	MT394512	621826
<i>Prionopetalum kraepelini</i>	Nguru, Kanga forest	2017	15 114	MT394524	621927
<i>Pseudotibiozus cerasopus</i>	Udzungwa Mountains National Park, Mizimu Camp	2012	15 121	MT394506	621828

\*Sequencing was unsuccessful.

†Partly complete.

2018) and run at 98 °C for 15 s, 65 °C for 30 s, 72 °C for 30 s, with a final extension step of 72 °C for 1 min. The PCR-amplified libraries were then purified using the Agencourt AMPure XP PCR purification kit (Beckman Coulter) and eluted in 50 µL elution buffer.

Sequencing of all *Tropostreptus* samples and *C. netus* was performed on a BGISEQ-500RS platform at BGI China, pooling ten libraries per lane with 150 Paired-end chemistry. The remaining samples were sequenced at the Danish National High-Throughput Sequencing Centre, using 150PE chemistry on an Illumina HiSeq4000 platform with Illumina indexes.

## BIOINFORMATICS

### Mitogenome assembly and annotation

We used ADAPTERREMOVAL v.2.2 (Schubert *et al.*, 2016) to trim the adapter sequences from the reads. Consecutive stretches of Ns, low-quality bases and sequences with a length of < 30 bp were removed. The mitogenomes were assembled using NOVOPLASTY v.2.6.3 (Dierckxsens *et al.*, 2017) with default parameters, which requires a reference sequence or a seed (i.e. a starting sequence for assembly initiation, usually a barcode region) for the mitochondrial DNA (mtDNA) assembly. Given that neither reference sequences nor barcodes were available for any of the 28 species, we mapped the raw reads to the barcode of *Bicoidens aridis* Mwabvu, 2009 (GenBank accession number KM982536), also from the family Spirostreptidae, with relaxed mapping parameters, using GENEIOUS v.9.1.8 (geneious.com). The resulting consensus sequences of each sample served as the seed for the assembly process by NOVOPLASTY. Annotations of the final assemblies were carried out using MITOZ v.2.3 (Meng *et al.*, 2019). GENEIOUS was also used to assess the quality of the assemblies manually and to visualize the annotations.

To estimate the levels of sequence coverage across the mitogenomes and to investigate DNA damage in these ethanol-preserved museum specimens, we mapped the adapter-trimmed sequence reads of each sample to its consensus mtDNA sequence using the BWA v.0.7.15 'aln' algorithm (Li & Durbin, 2009). Aligned sequences were filtered for mapping quality 30 and sorted using SAMTOOLS (Li *et al.*, 2009). Duplicate sequences were removed by PICARD MARKDUPLICATES (<http://picard.sourceforge.net>). Given that some of the samples were collected > 30 years ago, we estimated DNA damage parameters (such as average read length and the C→T transition rates at the first position of the 5' end of DNA reads), using MAPDAMAGE v.2.0 (Jónsson *et al.*, 2013). The depth of coverage for each sample was calculated using the 'genobox\_bam2avgdepth1.py'

script ([https://github.com/srcbs/GenoBox/blob/master/genobox\\_bam2avgdepth1.py](https://github.com/srcbs/GenoBox/blob/master/genobox_bam2avgdepth1.py)).

The same mapping approach was implemented to estimate the number of DNA reads originating from *R. gigas* fungus from each specimen by mapping the reads to the *R. gigas* fungus NSL1–NSL2 region (GenBank accession numbers MT416216 and MT416217).

### Phylogenetic analysis

Phylogenetic reconstruction was carried out based on 13 protein-coding genes of the mitochondrial genomes, which were aligned separately using the 'G-INS-i' algorithm in MAFFT plugin (v.1.3.6) in GENEIOUS v.9.1.8. The latter was also used to concatenate all 13 genes into longer contigs for phylogenetic analyses with RAXML-NG v.0.8.1 (Kozlov *et al.*, 2019) for a maximum likelihood (ML) phylogenetic tree analysis. We used default RAXML parameters with the GTR substitution model (estimated by RAXML-NG) and performed 100 bootstrap replicates to obtain node support values. An additional 12 millipede mitochondrial genomes were retrieved from GenBank (Supporting Information, Table S1; Fig. S1), but we focus specifically on the families Spirostreptidae and Odontopygidae (to which our samples belong). Hence, we include only two of these mitochondrial genomes from GenBank [*Thyropygus* Pocock, 1894, sp. and *Abacion magnum* (Loomis, 1943)] in our main analysis and figures. A full phylogeny, including all available millipede mitochondrial genomes, is presented in the Supporting Information (Fig. S1).

We used BEAST v.2.6.1 (Bouckaert *et al.*, 2019) to construct a Bayesian coalescent-based phylogeny to provide rough dates for the split times between various millipede lineages. Given that neither good fossil records nor reliable substitution rates are available for this group of invertebrates, we used the general 'standard' arthropod mitochondrial clock substitution rate of  $1.15 \times 10^{-8}$  per site per year (equivalent to 2.3% divergence per Myr) (Brower, 1994) as a proxy for a general millipede mtDNA substitution rate, with relaxed log-normal distribution priors. However, given the potentially wide range of substitution rates in arthropods (Papadopoulos *et al.*, 2010; Allio *et al.*, 2017), the exact split times of various millipede lineages should be interpreted with caution. In addition, considering findings that various partitioning schemes of the mtDNA do not significantly affect the estimates of divergence times in Bayesian phylogenetic dating (Jin & Brown, 2018), we decided not to partition the mtDNA sequences for this analysis. The tree was constructed using a GTR+I+G nucleotide substitution model as determined by JMODELTEST (Darriba *et al.*, 2012) run with the

Akaike information criterion and a Bayesian skyline prior, which was previously shown to be a reliable prior for inter- and intraspecies-level phylogenetic trees (Ritchie *et al.*, 2017). We ran the Markov chain Monte Carlo chains for  $10^8$  generations with sampling every  $10^4$  generations and designated the first  $10^6$  states as burn-in. We checked the output for convergence to a stationary distribution and sufficient effective sample size estimates ( $\geq 200$ ) using TRACER v.1.7.1 (Rambaut *et al.*, 2018). TREEANNOTATOR from the BEAST v.2.6.1 package and FIGTREE v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) were used to visualize the resulting phylogeny from the BEAST analysis.

In an attempt to obtain genome-wide phylogenetic information in the absence of reference genomes, we reconstructed a *k*-mer-based phylogenetic tree using our shotgun sequencing data. We estimated the proportion of shared *k*-mers ( $k = 35$ ) between the samples to create a distance matrix for the phylogenetic reconstruction with the AAF (alignment- and assembly-free) software package (Fan *et al.*, 2015). We used duplicate-free (PCR and optical) fastq files as input for the AAF. The duplicates were removed by 'clumpify' from the BBMAP package (BBMAP; Bushnell B.; [sourceforge.net/projects/bbmap/](https://sourceforge.net/projects/bbmap/)) using the 'dedup=t' parameter.

## RESULTS

A total of 1.46 billion paired-end DNA reads were sequenced for this project, with an average of 56 million read pairs per sample. Of these, the average non-collapsed read length per sample ranged between 94 and 129 bp (average of 97 bp).

From the 28 specimens, we were able to obtain 25 complete and one partial mitochondrial genome representing all the sampled species (Table 1). The two remaining specimens (one *Tropostreptus hamatus* and one *Tropostreptus droides*) did not yield sufficient sequences for a successful mitogenome assembly. Successfully sequenced specimens were collected between 4 and 40 years ago. For the assembled mitochondrial genomes, the average depth of coverage was high, varying between  $114\times$  and  $4050\times$ , with an average of  $1111\times$ . Furthermore, we found no signs of significant DNA deamination damage, estimated as the C→T transition rates at the first position of the 5' end of reads, in any of the samples (Supporting Information, Fig. S2; Table S2), indicating high levels of DNA preservation for all specimens.

The mitochondrial genome sizes range from 15 093 to 15 367 bp, which is similar to that reported from other Diplopoda species (14 747–15 791 bp; Supporting Information, Table S1). Likewise, our 26 new sequenced mitochondrial genomes comprise 13 protein-coding

regions, two ribosomal subunits and 22 transfer RNAs (tRNAs), as also seen in all other published Diplopoda mitochondrial genomes.

The *Tropostreptus* samples showed an average pairwise similarity of 90.2% for the *COX1* gene, with the highest intraspecific similarity found for the two *Tropostreptus kipunji* samples (100%) and the lowest observed among the eight *Tropostreptus sigmatospinus* samples (95.4%). A similar pattern (although with slightly lower values) was observed when estimating across all the coding regions (CDS) in the mitogenome (Supporting Information, Table S3).

## MITOCHONDRIAL GENOME ORGANIZATION

All our sequenced millipede mitochondrial genomes showed the same order and orientation of genes. The gene arrangement is unusual, in that genes transcribed in opposite directions are located in two different parts of the genome that are separated by two non-coding regions (Fig. 3). Specifically, genes for the first 8000 bp are transcribed in one direction, and genes for the remaining ~8000 bp genes are transcribed in the opposite direction. This applies for all genes except tRNA-Cys. This arrangement has also been reported for *Thyropygus* sp. and *Narceus annularis* (Rafinesque, 1820) (Lavrov *et al.*, 2002). For the remaining ten Diplopoda species with full mitochondrial genomes available from GenBank, the order and orientation of the genes differ in various ways (e.g. Dong *et al.*, 2016).

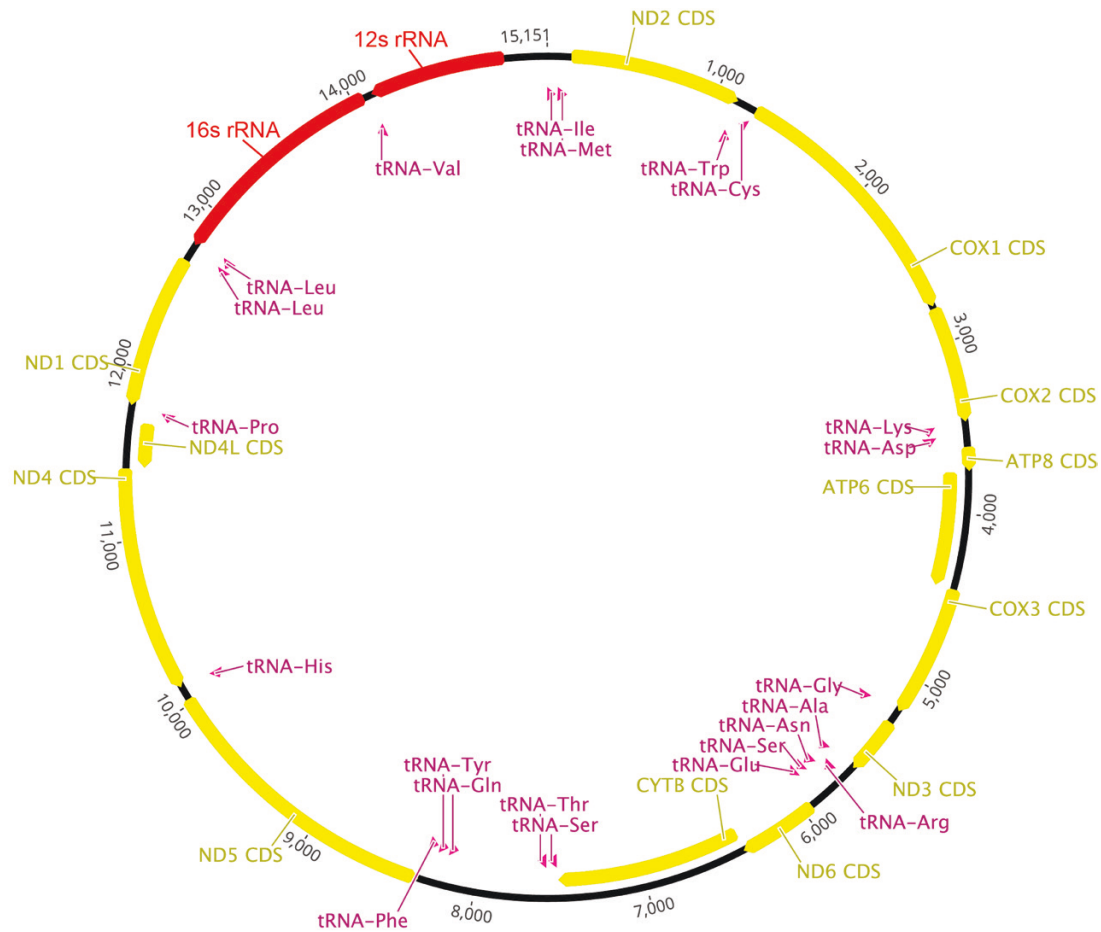
## RICKIA FUNGUS

The DNA sequenced from the associated *Rickia* Cavara, 1899 fungus was identified in a single sample; a *Tropostreptus hamatus* individual from Udzungwa (NHMD 621809). We identified 111 sequences mapping with > 99% sequence identity to the small subunit ribosomal RNA gene of *R. gigas* (GenBank accession number MT416216.1). This allowed us to reconstruct a consensus sequence of 559 bp with >  $3\times$  coverage. Our consensus sequence was identical to the *R. gigas* reference sequence, apart from a single G→A substitution. Our new *Rickia* sequence has been uploaded to GenBank with the accession number MT704348.

## PHYLOGENETIC ANALYSES

The maximum likelihood-based phylogenetic tree for all analysed samples is shown in Figure 4. A corresponding phylogenetic tree including all 12 previously published millipede mitochondrial genomes is shown in the Supporting Information (Fig. S2). Overall, the node support values are high. All *Tropostreptus* species come





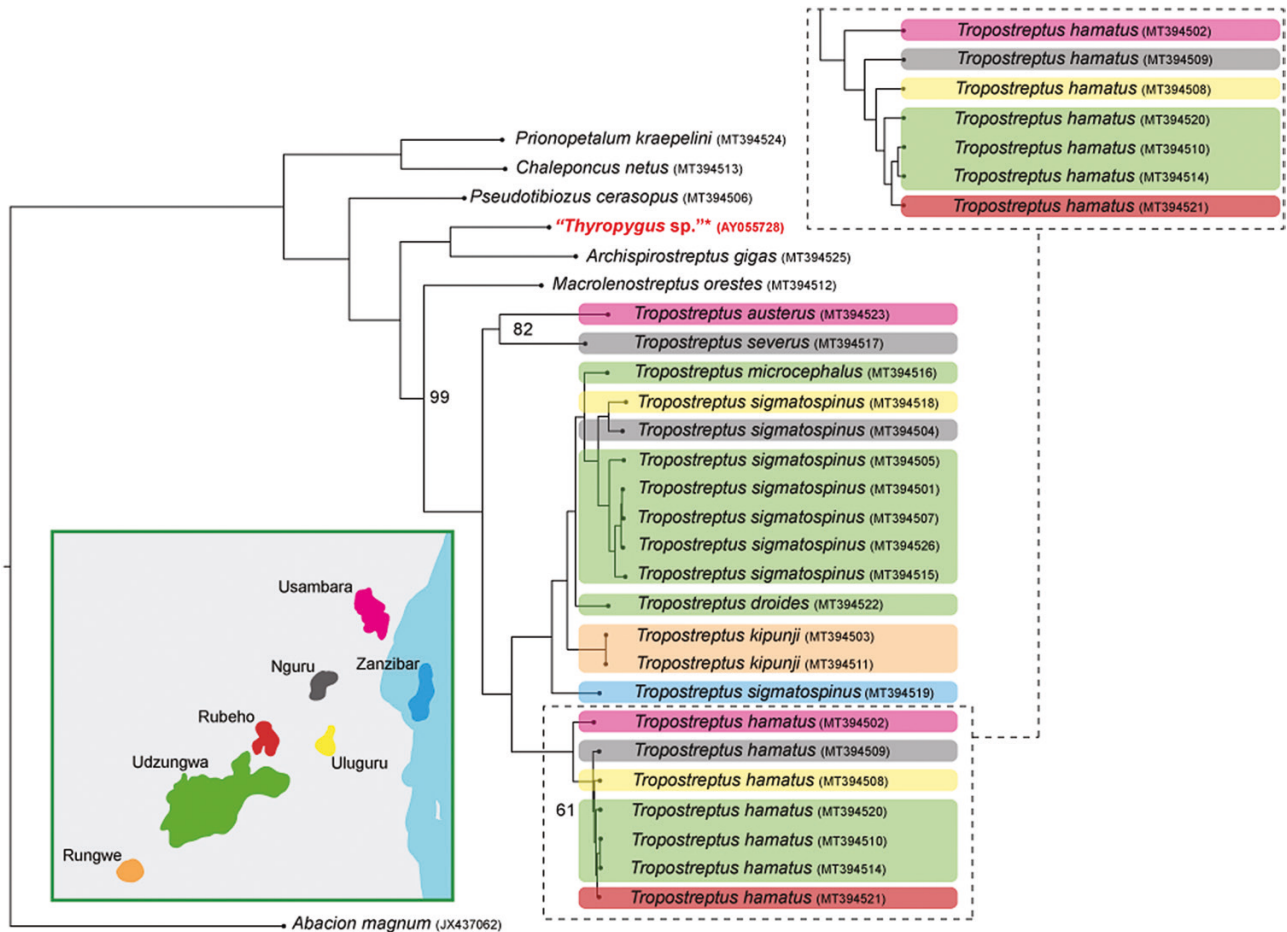
**Figure 3.** The gene order of mitochondrial coding sequences shared among all analysed millipede species in this study, which include all known species of *Tropostreptus* (*T. droides*, *T. hamatus*, *T. kipunji*, *T. microcephalus*, *T. severus* and *T. sigmatospinus*), in addition to *Archispirostreptus gigas*, *Chaleponcus netus*, *Macrolenostreptus orestes*, *Prionopetalum kraepelini* and *Pseudotibiozus cerasopus*. Colour key: red, ribosomal RNA (rRNA); pink, transfer RNA (tRNA); yellow, protein-coding sequences (CDS). Arrows indicate gene transcription orientation.

out as distinct monophyletic groups, except for a single *Tropostreptus sigmatospinus* specimen from Zanzibar. This individual falls basal to a purely Eastern Arc clade consisting of *Tropostreptus kipunji*, *Tropostreptus droides*, *Tropostreptus microcephalus* and the remaining *Tropostreptus sigmatospinus* individuals. The basal split in *Tropostreptus* separates a clade consisting of *Tropostreptus austerus* + *Tropostreptus severus* from the rest. The widespread *Tropostreptus hamatus*, represented by seven individuals, forms a monophyletic cluster, sister to the remaining species. Apart from the Zanzibar individual, another cluster is formed by seven of the eight individuals of the equally widespread *Tropostreptus sigmatospinus*. The last three species form a 'paraphyletic ladder' leading to the main *Tropostreptus sigmatospinus* clade and with a similar morphology *Tropostreptus droides* and *Tropostreptus kipunji* on the two lower rungs. Two of the species, *Tropostreptus hamatus*

and *Tropostreptus sigmatospinus*, have a distribution across several mountain blocks, and this intraspecific geographical division is also recovered in the topology. For *Tropostreptus hamatus*, covering five mountain blocks, the first split separates Usambara, followed by Nguru, Uluguru and then Rubeho and Udzungwa. Our analyses suggest that *Tropostreptus sigmatospinus* is polyphyletic, with the Zanzibar lineage being a more distant relative. Disregarding this Zanzibar outlier, the first phylogenetic split appears to have separated Udzungwa from Nguru + Uluguru.

A Bayesian coalescent-based phylogeny produced in BEAST recovered the same topology as the maximum likelihood method and again with strong node support (Fig. 5). Age estimation for lineage divergence was based on a general arthropod mtDNA substitution rate and should therefore be considered with caution. We find the oldest divergence within *Tropostreptus* occurring  $\sim 22 \pm 5$  Mya and the youngest species split in this





**Figure 4.** Maximum likelihood-based phylogeny, with 100 bootstrap replicates and using all the 26 mitochondrial genomes generated in this study. The dataset was supplemented with *Thyropygus* sp. and *Abacion magnum* as outgroups, with sequences derived from GenBank. GenBank accession numbers are given in parentheses. Colours represent *Tropostreptus* sample origins. The upper right inset shows the topology of the *Tropostreptus hamatus* lineage, enlarged to clarify the branching order. Only support values < 100 are shown. \**Thyropygus* sp. (red font) is very likely to be a species misidentification; for more information, see Discussion text.

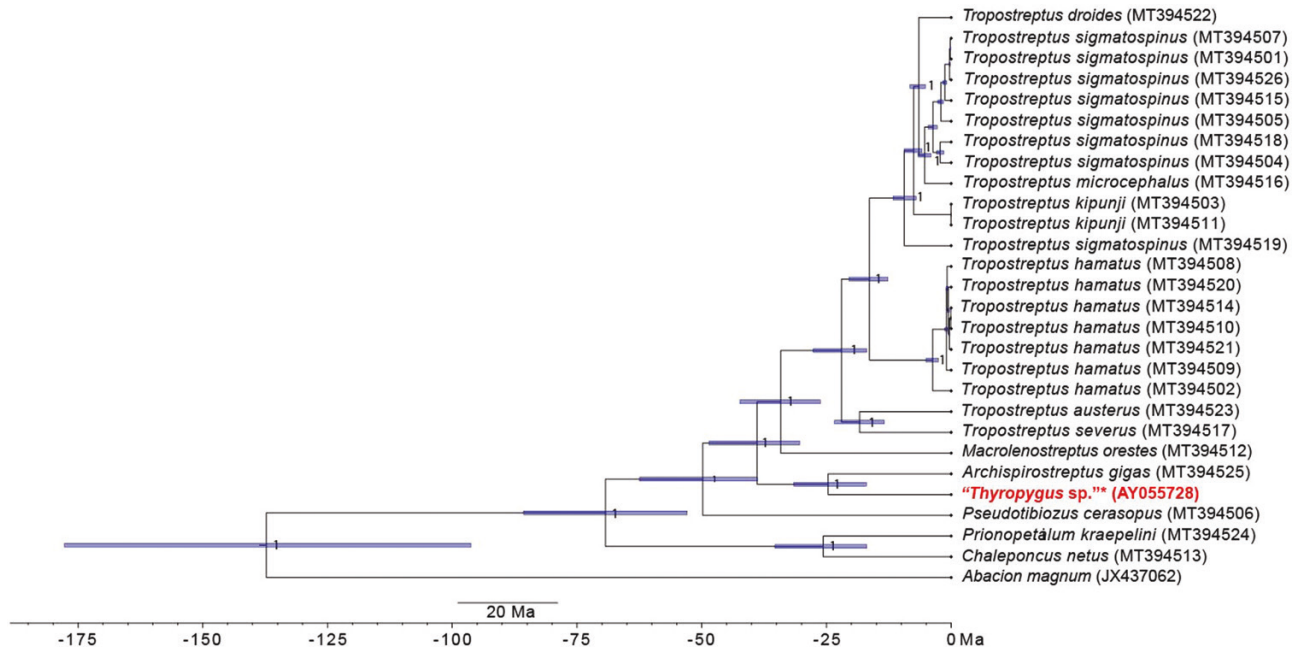
genus (the divergence of *Tropostreptus microcephalus* and *Tropostreptus sigmatospinus*) occurring ~5 Mya. We also observe that the majority of divergence events between *Tropostreptus* species have occurred in the last 10 Myr. Two of the species, *Tropostreptus hamatus* and *Tropostreptus sigmatospinus*, inhabit several mountains, enabling analyses of intraspecific divergence. The oldest *Tropostreptus hamatus* divergence occurred ~5 Mya, separating the Usambara population from the rest (Nguru, Uluguru and Rubeho), which then split up within the last 1–2 Myr. The oldest *Tropostreptus sigmatospinus* divergence occurred ~10 Mya when considering the Zanzibar lineage, but a mere 5 Mya if this lineage is not considered, when the Udzungwa population separated from Nguru and Uluguru.

The phylogenetic reconstruction, based on the genome-wide *k*-mer data (Supporting Information,

Fig. S3) is similar overall to the maximum likelihood and Bayesian phylogenies that were based on the mitochondrial genomes. Two notable differences are three *Tropostreptus sigmatospinus* samples (the one from Zanzibar and its two geographically closest mountains: Uluguru and Nguru) clustering together as a sister group to the remaining *Tropostreptus sigmatospinus* samples, and *Tropostreptus droides* and *Tropostreptus microcephalus* being grouped together.

## DISCUSSION

Few phylogenetic studies have been conducted on Spirostreptidae or on the order Spirostreptida in general (but see Iniesta *et al.*, 2020), and even fewer studies have been performed using molecular data of millipedes (e.g. Mwabvu *et al.*, 2013; VandenSpiegel *et al.*, 2021).



**Figure 5.** Bayesian phylogeny, with species divergence age estimates reconstructed with BEAST using all the 26 mitochondrial genomes generated in this study. The dataset was supplemented with *Thyropygus* sp. and *Abacion magnum* as outgroups, derived from GenBank. GenBank accession numbers are provided in parentheses. Blue bars indicate the 95% highest probability density intervals for node ages. Age estimation for lineage divergence was based on a general arthropod mitochondrial DNA substitution rate and should be considered with caution. \**Thyropygus* sp. (red font) is very likely to be a misidentification; for more information, see the Discussion.

With only 12 published millipede mitogenomes to date (Supporting Information, Table S1), our effort triples the number of publicly available millipede mitogenomes. Two samples failed to provide enough sequences for mitochondrial genome assembly and were therefore not included in the analyses (Table 1): one *Tropostreptus droides* specimen (collected in 1984 in the Udzungwa Scarp Forest Reserve) and one *Tropostreptus hamatus* specimen (collected in 1998 in Udzungwa Scarp Forest Reserve). Although these specimens were stored in the museum, preserved in ethanol, the field collecting procedure is unknown and might have been unsuitable for later molecular analysis (e.g. if they were stored in formalin initially). We found no signs of significant DNA deamination (C→T) damage in any of the analysed samples, although some were collected 40 years ago. Successful assembly of full mitochondrial genomes from museum samples (> 10 years since collection) has been reported for other arthropods (e.g. Gilbert *et al.*, 2007; Heintzman *et al.*, 2014; Timmermans *et al.*, 2016), and our study underlines the molecular research potential in old ethanol-preserved arthropods.

#### GENOME SYNTENY AND FEATURES

All our sequenced mitochondrial genomes (*Tropostreptus* spp. and *Ar. gigas*, *C. netus*, *M. orestes*,

*Pr. kraepelini* and *Ps. cerasopus*) are highly similar in length and compositions. They share the same unusual gene arrangement, whereby genes transcribed in the same direction are grouped together (except for tRNA-Cys) in the same half of the genome. The two gene groups are separated by two non-coding regions (Fig. 3). This arrangement is shared with previously sequenced millipedes, such as *N. annularis*, which belong to the order Spirobolida of the same superorder, Juliformia, as our analysed specimens (Lavrov *et al.*, 2002). To our knowledge, it has not been observed in other arthropod mitochondrial genomes (Lavrov *et al.*, 2002) and might therefore be unique to Juliformia, but more mitochondrial genomes from both within Juliformia (but not belonging to the orders Spirostreptida and Spirobolida) and outside Juliformia are needed to confirm this. Lavrov *et al.* (2002) proposed that this gene arrangement might have been caused by events of duplication and non-random loss of genes (Lavrov *et al.*, 2002; Brewer *et al.*, 2013).

#### TROPOSTREPTUS PHYLOGENY AND EVOLUTION

Twenty-one of our included mitochondrial genomes belong to *Tropostreptus*, allowing for a thorough investigation of the evolution of this genus in the Eastern Arc. We observe a clear genetic structure in

*Tropostreptus*, with distinct lineages (both inter- and intraspecific) being defined by the mountain blocks (Fig. 4). This is consistent with previous genetic results of Eastern Arc gene pools (e.g. cat snakes: Gravlund, 2002; chameleons: Tolley *et al.*, 2011; African violets: Dimitrov *et al.*, 2012), where forest-adapted species inhabit the montane forests and are absent from the adjacent savannah lowlands. Today, the mountains capture the oceanic winds from the Indian Ocean, which maintains sufficient humidity for dense rain forest to grow, resulting in the forest 'sky islands' (Lovett, 1993a, b; Burgess *et al.*, 2007). Until 30 Mya, the Eastern Arc region is thought to have been covered by rain forest (Rodgers, 1998; Couvreur *et al.*, 2008), and an uplifting of the Eastern Arc Mountains is believed to have occurred within the last 7 Myr (although this is debated), changing the whole topography of East Africa (Griffiths, 1993; Ring, 2014; Macgregor, 2015). Climatic and geological fluctuations through time have thus repeatedly affected the forest cover and, presumably, resulted in a multitude of vicariance events when species were isolated in patchy forest remnants (Lovett, 1993a; Sepulchre *et al.*, 2006; Couvreur *et al.*, 2008). For these reasons, the splitting order we observe in the *Tropostreptus* phylogeny might well reflect forest fragmentation in ancient times.

We observe a general trend, whereby northern lineages appear to split off first. The earliest split in *Tropostreptus* separates the *Tropostreptus austerus* + *Tropostreptus severus* lineage from the rest, and the second split separates these two species, today occupying Nguru and Usambara Mountains in the north. A similar intraspecific pattern is evident in more recent splits in *Tropostreptus hamatus* and *Tropostreptus sigmatospinus* (Fig. 4), suggesting a repeated pattern of vicariance events occurring first in the north. A separation of species between northern and southern mountains has also been observed in several other Eastern Arc taxa, including amphibians (Blackburn & Measey, 2009), gastropods (Tattersfield *et al.*, 1998) and reptiles (Gravlund, 2002; Tolley *et al.*, 2011), but also in well-dispersing taxa, such as birds (Fjelds   & Bowie, 2008).

This indicates a forest retraction southwards during dry periods, resulting in vicariance events, followed by forest expansion and thus northward recolonization of species during periods with higher humidity. Northward migration is also observed in other Eastern Arc species, such as chameleons (Tolley *et al.*, 2011; Ceccarelli *et al.*, 2014). A recent cycle of forest expansion/retraction can explain why *Tropostreptus hamatus* and *Tropostreptus sigmatospinus* exist across several of the mountains without having evolved into distinct species yet. Other events have isolated *Tropostreptus kipunji* in the forest on Mount Rungwe, the most south-westerly occurring species in

the Eastern Arc region, in addition to *Tropostreptus sigmatospinus* in Zanzibar and, potentially, also the Rondo Plateau, from where *Tropostreptus* has been observed but for which molecular data are still lacking (Enghoff, 2017).

Regarding the timing of the species splits (Fig. 5), several major events might have played a role. Around 30 Mya the Antarctic ice sheet started to form (Couvreur *et al.*, 2008), along with rifting that started to occur in northern East Africa (Ring, 2014), possibly initiating the fragmentation of the pan-African forest. Through millions of years, the rifting would continue southwards (Ring, 2014), affecting the topology and possibly related to the forest fragmentation responsible for the divergence of *Tropostreptus austerus* and *Tropostreptus severus* observed ~22 Mya. The observed divergence of *Tropostreptus hamatus* and the split between *Tropostreptus austerus* and *Tropostreptus severus* correspond well to the closing of the Tethys Sea (17 Mya), which would have altered ocean currents and, probably, the climate of the area (Couvreur *et al.*, 2008). Likewise, the isolation of the *Tropostreptus kipunji* lineage corresponds with the uplifting of Mount Rungwe from ~8 Mya (Ring, 2014). Finally, between 5 Mya and today, we observe a radiation in *Tropostreptus hamatus* and *Tropostreptus sigmatospinus* (Fig. 5). A reasonable explanation for this is the uplift of the Eastern Arc Mountains, shifting the precipitation from the lowlands to the mountains (Lovett, 1993a, b), in combination with the Antarctic ice sheet forming, thus decreasing global humidity (Polyak *et al.*, 2010). This would lead to the emergence of savannah in the lowlands between the mountains (Sepulchre *et al.*, 2006; S  galen *et al.*, 2007; Couvreur *et al.*, 2008), isolating the montane forest and limiting migration between populations of forest-restricted species.

We emphasize that we have neither good fossil records nor mtDNA mutation rates estimated specifically for millipedes, which is why the split times of our millipede phylogenetic tree should be interpreted with caution. Moreover, comparable studies with dated phylogenies of Eastern Arc species are sparse, hence it is difficult to compare the split times we have estimated with those of other species in the region. Examining two separate studies of chameleons (*Kinyongia* Tilbury, Tolley & Branch, 2006 and *Trioceros* Swainson, 1839) with dated phylogenies based on both mitochondrial and nuclear markers did show some correspondence with our dated splits (Tolley *et al.*, 2011; Ceccarelli *et al.*, 2014). Tolley *et al.* (2011) dated the earliest split between the northern and southern Eastern Arc species to ~28 Mya, and both studies show several radiation events between 5 and 20 Mya, corresponding to the same overall time frame that we are discussing for the millipedes. In contrast, the chameleons display fewer speciation events during



the last 5 Myr than the millipedes, perhaps suggesting that the latter have been more susceptible to vicariance during more recent climatic events.

#### CONGRUENCE BETWEEN MORPHOLOGY AND MOLECULAR DATA

We highlight that one of the millipede mitochondrial genomes, obtained from GenBank and identified as *Thyropygus* sp. (AY055728), is likely to have been misidentified. It was purchased as an 'African giant millipede' (Dennis V. Lavrov, pers. comm.), but *Thyropygus* is an exclusively Asian genus belonging to the family Harpagophoridae (Pimvichai *et al.*, 2009). Several large African species of Spirostreptidae are available in the pet trade, including *Ar. gigas*, with which '*Thyropygus* sp.' groups in our trees. '*Thyropygus* sp.' is therefore likely to be a close relative of *Ar. gigas* and not a real *Thyropygus* sp. Assuming that '*Thyropygus* sp.' is indeed a misidentified spirostreptid (*Archispirostreptus* sp.?), the branching order of our samples (Fig. 4) is consistent with current morphology-based classification, indicating monophyly of the families Odontopygidae (represented by *Pr. kraepelini* and *C. netus*) and Spirostreptidae (remaining species apart from the far outgroup *Ab. magnum*).

With the exception of the *Tropostreptus sigmatospinus* specimen from Zanzibar, all species represented by more than one individual are retrieved as monophyletic in our molecular analysis (Fig. 4). Enghoff (2017) did not include a molecular-based phylogenetic analysis, but divided *Tropostreptus* into three morphological distinct groups, viz. the *hamatus* group (*hamatus*, *sigmatospinus* and *microcephalus*), the *droides* group (*droides* and *kipunji*) and the *austerus* group (*austerus* and *severus*). In the molecular phylogeny, the morphologically defined *austerus* group comes out monophyletic and as sister to the remaining species, but the two other morphologically defined groups are not retrieved as monophyletic. Instead, the *droides* group is retrieved as paraphyletic and nested in the *hamatus* group, but the two species of the *droides* group are morphologically similar despite being widely separated geographically.

As mentioned, the *Tropostreptus sigmatospinus* specimen from Zanzibar does not group with the other Eastern Arc *Tropostreptus sigmatospinus* samples (Fig. 4; but see Supporting Information, Fig. S3). Instead, it comes out at the base of a clade consisting of, in ascending sequence, *Tropostreptus kipunji*, *Tropostreptus droides*, *Tropostreptus microcephalus* and Eastern Arc *Tropostreptus sigmatospinus* (monophyletic). Pairwise similarity for the *Tropostreptus sigmatospinus* samples also increases markedly when excluding the Zanzibar specimen (Supporting Information, Table S3). Thus, we speculate

that this Zanzibar specimen (the only non-Eastern Arc specimen in our study) most probably represents an unrecognized species. To elucidate this further, studies of more non-Eastern Arc *Tropostreptus* specimens are needed, given that *Tropostreptus sigmatospinus* has also been recorded from the Rondo Plateau in south-east Tanzania (Enghoff, 2017). We note that the genome-wide *k*-mer-based phylogenetic reconstruction displays notable differences; for instance, *droides* and *microcephalus* group together, and the *Tropostreptus sigmatospinus* specimen from Zanzibar groups together with its geographically closest *Tropostreptus sigmatospinus* samples from Nguru and Uluguru, but given the lack of node value support with this distance-based method, these deviating observations should be interpreted with caution and warrant further investigation (Supporting Information, Fig. S3).

#### RICKIA FUNGUS

We were able to detect the presence of a millipede-associated *R. gigas* fungus in a single sample of *Tropostreptus hamatus* from Udzungwa. *Rickia gigas* is a species of Laboulbeniales (an order of fungi known to parasitize arthropods) that has been observed visually in several *Tropostreptus* species and in *Ar. gigas* (Santamaria *et al.*, 2016; Enghoff, 2017). The DNA extraction and molecular detection of Laboulbeniales have proved difficult, and new methods had to be developed specifically for this (Haelewaters *et al.*, 2015; Sundberg *et al.*, 2018). It is therefore interesting that one of our samples revealed the presence of *R. gigas*, although we applied no special treatment for its detection, except for a short storage of DNA extract at  $-80^{\circ}\text{C}$ . Future studies aiming to investigate *Rickia* fungi from millipedes are therefore highly feasible with next-generation 'shotgun' sequencing data, and it should also be investigated systematically whether this short deep-freeze storage and/or the DNA extraction procedure enhances the potential for detection of *Rickia* DNA.

#### CONCLUSION

We have shown how it is possible to derive complex genetic information from old millipede museum specimens and how this can be used for detailed phylogenetic analysis even decades after collecting the samples. The genetic information retrieved in this study has markedly increased the number of publicly available mitogenomes from millipedes. Furthermore, it brings new knowledge on the evolutionary processes in the Eastern Arc Mountains, a biodiversity hotspot of great scientific and conservation interest. Our results highlight the unique and vital role that museum collections play by offering an enormous genetic research potential. New studies of both vertebrates and

invertebrates could therefore consider the usability of existing material stored in museum collections before venturing out to collect new and potentially endangered species.

## ACKNOWLEDGEMENTS

Thanks to Francis B. N. Sabuni (Eastern Arc Mountains Conservation Endowment Fund) for permission to use the base map for [Figure 2](#) and to Danny Haelewaters (Purdue University) for providing us with *Rickia gigas* DNA sequences. Thanks to Tom Gilbert (GLOBE Institute, University of Copenhagen) for his support, and thanks to Sofie Holdflod Nielsen and Luisa dos Santos Bay Nielsen (both at the GLOBE Institute, University of Copenhagen) for technical assistance. M.E.A. is funded by the Independent Research Fund Denmark (Sapere Aude, grant no. 7027-00147B). The study also received funding from Dr Bøje Benzons Støttefond.

## CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## AUTHOR CONTRIBUTIONS

M.N., A.M., H.E. and M.E.A. designed the study. M.N. and T.L.N. carried out laboratory work, and A.M. carried out bioinformatics and analysis. All authors contributed to the writing and approved the final version.

## DATA ACCESSIBILITY

Sequence data are available at <https://doi.org/10.17894/ucph.64a6ffaa-b6b2-48c3-a05b-70bb6b316006>.

## REFERENCES

- Allentoft ME, Rasmussen AR, Kristensen HV. 2018. Centuries-old DNA from an extinct population of aesculapian snake (*Zamenis longissimus*) offers new phylogeographic insight. *Diversity* **10**: 14.
- Allio R, Donega S, Galtier N, Nabholz B. 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. *Molecular Biology and Evolution* **34**: 2762–2772.
- Bibby CJ, Collar NJ, Crosby MJ, Heath MF, Imboden C, Johnson TH, Long AJ, Stattersfield AJ, Thirgood SJ. 1992. Putting biodiversity on the map: priority areas for global conservation. International council for bird preservation. Cambridge, United Kingdom.
- Blackburn DC, Measey GJ. 2009. Dispersal to or from an African biodiversity hotspot? *Molecular Ecology* **18**: 1904–1915.
- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **15**: e1006650.
- Brewer MS, Sierwald P, Bond JE. 2012. Millipede taxonomy after 250 years: classification and taxonomic practices in a mega-diverse yet understudied arthropod group. *PLoS One* **7**: e37240.
- Brewer MS, Swofford L, Spruill CL, Bond JE. 2013. Arthropod phylogenetics in light of three novel millipede (Myriapoda: Diplopoda) mitochondrial genomes with comments on the appropriateness of mitochondrial genome sequence data for inferring deep level relationships. *PLoS One* **8**: e68005.
- Brower AV. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 6491–6495.
- Burgess N, Butynski TM, Cordeiro NJ, Doggart NH, Fjeldsø J, Howell KM, Kilahama FB, Loader SP, Lovett JC, Mbilinyi B, Menegon M. 2007. The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. *Biological Conservation* **134**: 209–231.
- Carøe C, Gopalakrishnan S, Vinner L, Mak SS, Sinding MH, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MT. 2018. Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution* **9**: 410–419.
- Ceccarelli FS, Menegon M, Tolley KA, Tilbury CR, Gower DJ, Laserna MH, Kasahun R, Rodriguez-Prieto A, Hagmann R, Loader SP. 2014. Evolutionary relationships, species delimitation and biogeography of Eastern Afrotropical horned chameleons (Chamaeleonidae: Trioceros). *Molecular Phylogenetics and Evolution* **80**: 125–136.
- Christenhusz MJM. 2012. On African violets and Cape primroses—towards a monophyletic *Streptocarpus* (Gesneriaceae). *Phytotaxa* **46**: 3–9.
- Couvreur TLP, Chatrou LW, Sosef MSM, Richardson JE. 2008. Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biology* **6**: 54.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* **45**: e18.

- Dimitrov D, Nogués-Bravo D, Scharff N. 2012.** Why do tropical mountains support exceptionally high biodiversity? The Eastern Arc mountains and the drivers of *Saintpaulia* diversity. *PLoS One* **7**: e48908.
- Dong Y, Zhu L, Bai Y, Ou Y, Wang C. 2016.** Complete mitochondrial genomes of two flat-backed millipedes by next-generation sequencing (*Diplopoda*, *Polydesmida*). *ZooKeys* **637**: 1–20.
- Engelhoff H. 2015.** Chapter 13. Diplopoda — geographical distribution. In: A. Minelli, ed. *Treatise on zoology - anatomy, taxonomy, biology. The Myriapoda*, Vol. 2. Leiden: Brill, 329–336.
- Engelhoff H. 2017.** A new East African genus of spirostreptid millipedes (Diplopoda, Spirostreptida, Spirostreptidae), with notes on their fungal ectoparasite *Rickia gigas*. *Zootaxa* **4273**: 501–530.
- Engelhoff H. 2018.** A mountain of millipedes VII: the genus *Eviulisma* Silvestri, 1910, in the Udzungwa Mountains, Tanzania, and related species from other Eastern Arc Mountains. With notes on *Eoseviulisma* Brolemann, 1920, and *Suohelisma* Hoffman, 1963 (Diplopoda, Polydesmida, Paradoxosomatidae). *European Journal of Taxonomy* **445**: 1–90.
- Engelhoff H. 2020.** A mountain of millipedes VIII. The genus *Aquattuor* Frederiksen, 2013 revisited – a new species from the Udzungwa Mts, Tanzania, another from the Nguru Mts, and introduction of the first pair of male legs as a source of taxonomic characters (Diplopoda, Spirostreptida, Odontopygidae). *European Journal of Taxonomy* **626**: 1–32.
- Engelhoff H. 2022.** Mountains of millipedes. The family Odontopygidae in the Eastern Arc Mountains of Tanzania (Diplopoda, Spirostreptida). *European Journal of Taxonomy* **803**: 1–136.
- Engelhoff H, Petersen G, Seberg O. 2011.** Phylogenetic relationships in the millipede family Julidae. *Cladistics* **27**: 606–616.
- Evans BJ, Kelley DB, Tinsley RC, Melnick DJ, Cannatella DC. 2004.** A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. *Molecular Phylogenetics and Evolution* **33**: 197–213.
- Fan H, Ives AR, Surget-Groba Y, Cannon CH. 2015.** An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. *BMC Genomics* **16**: 522.
- Fjeldså J, Bowie RCK. 2008.** New perspectives on the origin and diversification of Africa's forest avifauna. *African Journal of Ecology* **46**: 235–247.
- Gilbert MTP, Moore W, Melchior L, Worobey M. 2007.** DNA extraction from dry museum beetles without conferring external morphological damage. *PLoS One* **2**: e272.
- Gravlund P. 2002.** Molecular phylogeny of Tornier's cat snake (*Crotaphopeltis tornieri*), endemic to East African mountain forests: biogeography, vicariance events and problematic species boundaries. *Journal of Zoological Systematics and Evolutionary Research* **40**: 46–56.
- Griffiths CJ. 1993.** The geological evolution of East Africa. In: JC Lovett, SK Wasser, eds. *Biogeography and Ecology of the Rain Forests of Eastern Africa*, Cambridge: Cambridge University Press, 9–22.
- Haelewaters D, Gorczak M, Pfliegler WP, Tartally A, Tischler M, Wrzosek M, Pfister DH. 2015.** Bringing *Laboulbeniales* into the 21st century: enhanced techniques for extraction and PCR amplification of DNA from minute ectoparasitic fungi. *IMA Fungus* **6**: 363–372.
- Heintzman PD, Elias SA, Moore K, Paszkiewicz K, Barnes I. 2014.** Characterizing DNA preservation in degraded specimens of *Amara alpina* (Carabidae: Coleoptera). *Molecular Ecology Resources* **14**: 606–615.
- Hopkins B. 1987.** Vegetation map of Africa. The vegetation of Africa: a descriptive memoir to accompany the Unesco/AETFAT/UNSO Vegetation map of Africa by F. White. *Journal of Ecology* **75**: 1214–1216.
- Iniesta LFM, Engelhoff H, Brescovit A, Bouzan RS. 2020.** Phylogenetic placement of the monotypic genus *Holopodostreptus* Carl, 1913 and notes on the systematics of Pseudonannolenidae (Spirostreptida: Cambalidea). *Invertebrate Systematics* **34**: 661–677.
- Jin Y, Brown RP. 2018.** Partition number, rate priors and unreliable divergence times in Bayesian phylogenetic dating. *Cladistics* **34**: 568–573.
- Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013.** mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**: 1682–1684.
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019.** RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**: 4453–4455.
- Lavrov DV, Boore JL, Brown WM. 2002.** Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. *Molecular Biology and Evolution* **19**: 163–169.
- Li H, Durbin R. 2009.** Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009.** The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Lovett JC. 1993a.** Climatic history and forest distribution in eastern Africa. In: JC Lovett, SK Wasser, eds. *Biogeography and ecology of the rain forests of Eastern Africa*. Cambridge: Cambridge University Press, 23–29.
- Lovett JC. 1993b.** Eastern Arc moist forest flora. In: JC Lovett, SK Wasser, eds. *Biogeography and ecology of the rain forests of Eastern Africa*. Cambridge: Cambridge University Press, 33–55.
- Macgregor D. 2015.** History of the development of the East African Rift System: a series of interpreted maps through time. *Journal of African Earth Sciences* **101**: 232–252.
- Measey GJ, Tolley KA. 2011.** Sequential fragmentation of Pleistocene forests in an East Africa biodiversity hotspot: chameleons as a model to track forest history. *PLoS One* **6**: e26606.



- Meng G, Li Y, Yang C, Liu S. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research* **47**: e63.
- Mittermeier RA, Gil PR, Hoffmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreux J, Da Fonseca GAB. 2004. *Hotspots revisited. Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Mexico City, Cemex.
- Mittermeier RA, Myers N, Thomsen JB, da Fonseca GAB, Olivieri S. 1998. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conservation Biology* **12**: 516–520.
- Mwabvu T, Lamb J, Slotow R, Hamer M, Barraclough D. 2013. Is millipede taxonomy based on gonopod morphology too inclusive? Observations on genetic variation and cryptic speciation in *Bicoidens flavicollis* (Diplopoda: Spirostreptida: Spirostreptidae). *African Invertebrates* **54**: 349–356.
- Nielsen M, Gilbert MTP, Pape T, Bohmann K. 2019. A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental DNA* **1**: 144–154.
- Nishii K, Hughes M, Briggs M, Haston E, Christie F, DeVilliers MJ, Hanekom T, Roos WG, Bellstedt DU, Möller M. 2015. *Streptocarpus* redefined to include all Afro-Malagasy Gesneriaceae: molecular phylogenies prove congruent with geographical distribution and basic chromosome numbers and uncover remarkable morphological homoplasies. *Taxon* **64**: 1243–1274.
- Olson DM, Dinerstein E. 1998. The global 200: a representation approach to conserving the Earth's most biologically valuable ecoregions. *Conservation Biology* **12**: 502–515.
- Papadopoulou A, Anastasiou I, Vogler AP. 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Molecular Biology and Evolution* **27**: 1659–1672.
- Papadopoulou A, Knowles LL. 2015. Genomic tests of the species-pump hypothesis: recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. *Evolution* **69**: 1501–1517.
- Pimvichai P, Enghoff H, Panha S. 2009. A revision of the *Thyropygus alleatus* group. Part 1: the *T. opinatus* subgroup (Diplopoda: Spirostreptida: Harpagophoridae). *Zootaxa* **2016**: 17–50.
- Pitz KM, Sierwald, P. 2010. Phylogeny of the millipede order Spirobolida (Arthropoda: Diplopoda: Helminthomorpha). *Cladistics* **26**: 497–525.
- Polyak L, Alley RB, Andrews JT, Brigham-Grette J, Cronin TM, Darby DA, Dyke AS, Fitzpatrick JJ, Funder S, Holland M, Jennings AE, Wolff E. 2010. History of sea ice in the Arctic. *Quaternary Science Reviews* **29**: 1757–1778.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901–904.
- Ring U. 2014. The East African rift system. *Austrian Journal of Earth Sciences* **107**: 132–146.
- Ritchie AM, Lo N, Ho SYW. 2017. The impact of the tree prior on molecular dating of data sets containing a mixture of inter- and intraspecies sampling. *Systematic Biology* **66**: 413–425.
- Rodgers WA. 1998. An introduction to the conservation of the Eastern Arc Mountains. *Journal of East African Natural History* **87**: 7–18.
- Santamaria S, Enghoff H, Reboleira AS. 2016. Hidden biodiversity revealed by collections-based research—Laboulbeniales in millipedes: genus *Rickia*. *Phytotaxa* **243**: 101–127.
- Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes* **9**: 88.
- Ségalen L, Lee-Thorp JA, Cerling T. 2007. Timing of C<sub>4</sub> grass expansion across sub-Saharan Africa. *Journal of Human Evolution* **53**: 549–559.
- Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin J-J, Brunet M. 2006. Tectonic uplift and Eastern Africa aridification. *Science* **313**: 1419–1423.
- Shelley RM, Golovatch SI. 2011. Atlas of myriapod biogeography. I. Indigenous ordinal and supra-ordinal distributions in the Diplopoda: perspectives on taxon origins and ages, and a hypothesis on the origin and early evolution of the class. *Insecta Mundi* **158**: 1–134.
- Stattersfield AJ, Crosby MJ, Long AJ, Wege DC. 1998. *Endemic Bird Areas of the World: Priorities for Biodiversity Conservation*. Cambridge: BirdLife International.
- Sundberg H, Ekman S, Kruys A. 2018. A crush on small fungi: an efficient and quick method for obtaining DNA from minute ascomycetes. *Methods in Ecology and Evolution* **9**: 148–158.
- Tattersfield P, Seddon MB, Meena C. 1998. Ecology and conservation of the land-snails of the Eastern Arc Mountains. *Journal of East African Natural History* **87**: 119–138.
- Taylor PJ, Maree S, Van Sandwyk J, Kerbis Peterhans JC, Stanley WT, Verheyen E, Kaliba P, Verheyen W, Kaleme P, Bennett NC. 2009. Speciation mirrors geomorphology and palaeoclimatic history in African laminate-toothed rats (Muridae: Otomyini) of the *Otomys denti* and *Otomys lacustris* species-complexes in the 'Montane Circle' of East Africa. *Biological Journal of the Linnean Society* **96**: 913–941.
- Timmermans MJTN, Viberg C, Martin G, Hopkins K, Vogler AP. 2016. Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. *Biological Journal of the Linnean Society* **117**: 83–95.
- Tolley KA, Tilbury CR, Measey GJ, Menegon M, Branch WR, Matthee CA. 2011. Ancient forest fragmentation or recent radiation? Testing refugial speciation models in chameleons within an African biodiversity hotspot. *Journal of Biogeography* **38**: 1748–1760.
- VandenSpiegel D, Henrard A, Mathys A. 2021. *Sechelleptus arborivagus* sp. nov., a new arboreal spirostreptid millipede (Diplopoda, Spirostreptidae) endemic to Mayotte Island (Comoros Archipelago), Indian Ocean. *European Journal of Taxonomy* **755**: 1–21.

**Wesener T. 2014.** A new phylogenetic analysis of the Sphaerotheriida (Diplopoda) with a revision of the Australian giant pill-millipedes. *Invertebrate Systematics* **28**: 196–213.

**Wesener T, VandenSpiegel D. 2009.** A first phylogenetic analysis of giant pill-millipedes (Diplopoda: Sphaerotheriida), a new model Gondwanan taxon, with special emphasis on island gigantism. *Cladistics* **25**: 545–573.

### SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Figure S1.** Phylogenetic tree based on all published millipede mitogenomes using maximum likelihood (ML). GenBank accession numbers are given in parentheses. \*Samples from the present study.

**Figure S2.** DNA damage plots for the oldest sample, *Tropostreptus sigmatospinus*, collected in 1981 in Zanzibar. All other samples showed equal or lower levels of DNA damage.

**Figure S3.** Distance-based phylogenetic tree using the whole shotgun sequencing data and an estimated proportion of shared  $k$ -mers. We used all the shotgun sequencing data to reconstruct a phylogenetic tree for all analysed millipedes. The proportion of shared  $k$ -mers ( $k = 35$ ) was estimated to create a distance matrix for phylogenetic reconstruction with the AAF (alignment- and assembly-free) software package. GenBank accession numbers are given directly after the species name, starting MT.

**Table S1.** All available mitochondrial genomes of millipedes found in GenBank (accessed April 2020).

**Table S2.** Analysis of DNA damage for all analysed samples.

**Table S3.** Pairwise identity for the *COX1* gene and for all the coding DNA sequences (CDS) for all analysed samples and for species with more than one sample.