Evolutionary Persistence and Speciation under Ancient Asexuality

Inauguraldissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

M. Sc.

Hüsna Öztoprak

aus Köln

Köln, 2023

Berichterstatter:

Prof. Dr. Michael Bonkowski, Terrestrische Ökologie, Universität zu Köln

Prof. Dr. Stefan Scheu, Tierökologie, JF Blumenbach Institut

Prof. Dr. Karine van Doninck, Université libre de Bruxelles

Vorsitz der Prüfung:

Prof. Dr. Tim Mansfeldt Bodengeographie/Bodenkunde, Universität zu Köln

Beisitzer:

Dr. Jens Bast, Universität zu Köln

Tag der mündlichen Prüfung: 11.01.2024

"*The cure for ignorance is to question*."

– Prophet Muhammad (s.a.w.) reported by Abu Daawood

"*Acquire knowledge and teach it to people. Learn with it dignity and tranquility and humility for those who teach you and humility for those whom you teach*."

– Umar ibn Al-Khattab (RA)

"*There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved*"

– Darwin (1859) *First edition of Origin of Species*

Acknowledgements

I would like to express my heartfelt gratitude to those who have been instrumental in my journey throughout my PhD. First and foremost, I extend my sincere thanks to Dr. Jens Bast for contributing to my growth as a researcher, teaching me to persist just like *Platynothrus peltifer* and for trusting me to be his 'official' first PhD Student. I'll fondly remember our sampling trips, culinary excursions and conversations about our not-so-niche-anymore music taste. This endeavor would not have been possible without Prof. Dr. Michael Bonkowski, who encouraged me generously since the very beginning of my academic career. I am grateful to Prof. Dr. Stefan Scheu, Prof. Dr. Karine van Doninck and Prof. Dr. Thomas Mansfeldt, for being part of my thesis committee and taking the time to read my thesis. I'd like to thank Mr. Fritz Schüller, my biology teacher, for fostering my love for biology and being a great teacher.

In addition, I am grateful to my co-workers, particularly Dr. Nadège Guiglielmoni, and students Svenja Wulsch, Viktoria Bednarski, and Katharina Burak, for standing up for me whenever I could not and for making the daily grind in the academic world bearable. The camaraderie and moments of enjoyment in their company eased the challenges of the office environment.
Thanks should also go to my friends who provided unwavering emotional support during the toughest times of my PhD: Barbara Paffendorf, Yasemin Birbir, Fernando Kreuz, Leonie Eißing, Marcel D. Solbach, Maria Sachs and Jule Freudenthal. Their willingness to listen, offer advice, and be a source of comfort was invaluable in maintaining my morale and mental health.
Further, I am also thankful to colleagues I met, who impacted and inspired me: Dr. Alexander Brandt, for checking in on me from time to time and encouraging me in a brotherly way. Dr. Fernando Villagomez for teaching me about oribatid mites and happily identifying all the mites I found interesting and cute. And of course, Dr. Kenneth Dumack, who was there since day one, always interested in what I had to share, gave me advice, believed in me, and kept on hoping that I might return to work on protists someday.

And finally, I am deeply indebted to my family. My parents Alpaslan and Serpil Öztoprak instilled in me a love for reading, and learning, and an unshakable belief in the power of education. Thank you for fostering an environment where I could be curious, explore the world, and be brave. And to my siblings Esma and Yunus, thank you for believing in me even during times when I doubted myself and reminding me that you are always there, fighting in my corner.

Table of Content

Acknowledgements **5**
Summary **7**
General Introduction **9**
Unraveling the Genomic Mysteries of Ancient Asexuality 9
Ancient Asexual Scandals 10
Unraveling the Genomic Mysteries of Asexual Speciation 11
Platynothrus peltifer as a Model Organism to Study Ancient Asexuality and Speciation 12
Aims and Outline of this Study 14
Contributions to the Chapters of this Thesis **15**
Chapter 1 Challenges in Genomic Studies of Oribatid Mites and the Need for High-Quality Assemblies of Single-individuals **16**
Modified salting out method for high-molecular-weight gDNA extraction (oribatid mites) 19
Chapter 2 Haplotypic Independence Contributes to Evolvability in the Long-term Absence of Sex in a Mite **28**
Chapter 3 Ancient Asexual Speciation in the Oribatid Mite *Platynothrus peltifer* **38**
Chapter 4 A Female Heterogametic ZW Sex-determination System in Acariformes **56**
Chapter 5 Science Outreach and Empowerment: Bridging the Gap for Diverse Audiences **70**
Having babies in soil: is sex really necessary? 72
General Discussion **80**
Importance of Single Individual Genomics from Natural Populations 80
Unraveling Ancient Asexuality 80
Unraveling Asexual Speciation 81
Discoveries in Sex Determination 82
Bridging Science and Education 82
Conclusion 83
General Outlook **84**
References **85**
General References 85
References Chapter 1 91
References Chapter 2 93
References Chapter 3 98
References Chapter 4 102
References Chapter 5 107
Appendix **108**
Supplementary Materials Chapter 2 108
Supplementary Materials Chapter 3 141
Supplementary Materials Chapter 4 147
Eidesstaatliche Erklärung **151**
Lebenslauf **152**

Summary

The enigma of why sexual reproduction persists in nature despite its many costs continues to challenge the field of evolutionary biology. Compared to sexual reproduction, asexuality is more straightforward. However, it is rare in animals and often regarded as an evolutionary dead-end. 'Ancient asexuals' challenge this view, as they successfully persisted over evolutionary time. Understanding how these evolutionary scandals thrive in the absence of sex will, in turn, provide insights into the selective advantage of sexual reproduction.

This thesis explored the consequences of ancient asexuality and emphasized its effects on genome evolution in natural populations. The oribatid mite *Platynothrus peltifer* was introduced as an ideal subject for studying ancient asexuality and speciation. Employing a comprehensive approach, I generated genome, transcriptome, and population data. Within this data, we found evidence of intra-individual haplotype divergence, conserved genomic synteny among populations, and notable inter-individual population variation. Thus, we proposed haplotypic independence to contribute to evolvability as a way to adapt and escape extinction in long-term asexuals.

P. peltifer is known to exhibit multiple independent evolutionary lineages. Asexual speciation may be more prevalent in certain taxa, like oribatid mites, due to their distinctive ecological niches and long-term evolutionary persistence. Understanding how asexuals diversify into new species is crucial as it challenges traditional evolutionary paradigms, reveals hidden biodiversity, and provides insights into the dynamics of speciation, evolution, taxonomy, and conservation. Thus, this study delved into the complexities of species delimitation, while emphasizing the need for a holistic perspective in comprehending asexual speciation. I suggested speciation of the Japanese population and identified potential cryptic lineages in the Canadian population of *P. peltifer*. European populations lacked distinct phylogenetic boundaries and genetic differences. However, haplotypic divergence in European populations implied cohesion mechanisms on a global scale among geographical lineages. Furthermore, this work laid the foundation for future sex-asex comparisons in oribatid mites as it provided high-quality genomes of sexual and asexual oribatid mites. This will determine whether the observed effects are lineage-specific or general consequences of ancient asexuality.

This study established a new protocol allowing for haplotype-specific genomics of tiny non-model-organisms. It discussed how ancient asexuals might have evolved and benefitted from their genome dynamics and how this facilitates asexual speciation. An in-depth analysis of *Platynothrus peltifer* provided evidence of its long-term asexuality and speciation. It introduced haplotypic independence as a way to allow for 'evolution' thus contributing to the persistence of ancient asexuals.

General Introduction

Unraveling the Genomic Mysteries of Ancient Asexuality

Despite the possibility of straightforward asexual reproduction, the majority of metazoans reproduce via costly sex (Lehtonen et al. 2012; Bell 1982). Even after decades of research, it remains largely unanswered why sex is maintained in natural populations and why it is needed every generation. This paradox is known as the “queen of problems" in evolutionary biology. The persistence of sexual reproduction in the face of competition from asexually reproducing individuals requires that the disadvantages of sex are fully compensated by its benefits (Neiman et al. 2014). Such disadvantages include smaller population size, possibly transmitted sexual diseases, fewer offspring, and the two-fold ‘cost of males’ experienced by sexual females who invest equally in sons and daughters, but only daughters contribute to the next generation. Short-term benefits of sex act on the level of genes and individuals and require a rapidly fluctuating direction of selection. The consensus is that through the breakage of linked alleles with contrasting fitness effects, more variation is generated by effective selection (Felsenstein 1974). Thus, the fixation probability is reduced when selection acts on a locus in linkage with another beneficial mutation (Hill and Robertson 1966). Hypotheses on the selective advantage of sex inherently predict that asexuality also comes with specific disadvantages that manifest in the genome over evolutionary times. Hence, most of the asexual lineages are observed at the tips of phylogenies, deeming asexuality an evolutionary dead-end (Judson and Normark 1996; Engelstädter 2008; Schwander and Crespi 2009).

Long-term consequences of asexuality act on the lineage level, manifesting over multiple generations. They can be divided into consequences driven by selection and/or mechanistic effects. Selection-driven consequences include less effective selection in asexuals due to strong physical linkage (Muller 1964; Hill and Robertson 1966; Felsenstein 1974; Keightley and Otto 2006; Fisher 1999). This is proposed to lead to decreased rates of adaptation, increased accumulation of deleterious mutations, and repetitive elements fueling the Hill-Robertson effect and Muller’s ratchet, ultimately leading to the extinction of asexual lineages. The accumulation of deleterious mutations in asexual genomes was shown in computational simulations (Muller 1964; Hill and Robertson 1966; Keightley and Otto 2006) as well as experimental evolutionary studies (Morgan et al. 2014). Genome-based studies show less effective selection in asexuals in *Timema* spp. (Bast et al. 2019), and no difference in Aphididae spp. (aphids; Ollivier et al. 2012). However, data on natural populations are scarce. Natural asexual populations of oribatid mites were shown to maintain effective selection and purge deleterious mutations even more successfully than sexual populations (Brandt et al. 2017). Mechanistic consequences of asexuality include independent divergence of homologous chromosomes known as the ‘Meselson-effect’. The phylogenetic history of sexual lineages shows initial separation into populations and then into separate

haplotypes because recombination maintains the similarity between haplotypes. In contrast, in asexual haplotype phylogenies, the haplotypes are expected to separate before population separation (Birky 1996). Therefore interallelic divergence is expected to be higher in asexuals. The lack of recombination and chromosomal rearrangements thus results in the independent evolution of homologous chromosomes. Within a single individual, this leads to high heterozygosity between two alleles of the same gene. Multiple studies revealing elevated heterozygosity levels between alleles were shown to be driven by the effects of a hybrid origin of asexuality, rather than prolonged asexuality (Jaron et al. 2021). Support for such within-individual divergence is limited. To date, the only documented case of haplotype divergence in parthenogenetic animals on the genome scale stems from *Oppiella nova* (oribatid mites; Brandt et al. 2021). The genomic peculiarities of oribatid mites that allow them to escape the predicted detrimental consequences of asexuality and persist over evolutionary time are still unclear.

Genomes of other asexual animal species have frequently been analyzed on some peculiarities that were suggested to be generally linked to asexuality, such as horizontal gene transfer, genomic rearrangements, gene family expansions, gene losses, and gene conversion (Flot et al. 2013; Danchin et al. 2010; Faddeeva-Vakhrusheva et al. 2017). Yet, parthenogenesis is regarded as a lineage-level trait, and up until now, it was not possible to detect genomic features as general consequences of parthenogenesis (Jaron et al. 2021).

Ancient Asexual Scandals

Obligate asexuality is assumed to be rare in animals, occurring only in approximately 0.1 % of species (White 1977; Bell 1982). This observation led to the assumption that asexual lineages are bound for early extinction (see potential reasons above). Still, there are some asexuals, who escape the detrimental fate of asexuality and persist over time. They are called ancient asexual scandals (Judson and Normark 1996). Understanding the peculiarities of how they are able to persist in the absence of sex will in turn contribute to identifying the advantage of sexual reproduction. Some of the most established representatives of ancient asexuals include bdelloid rotifers, darwinulid ostracods, and oribatid mites (Judson and Normark 1996; Schön et al. 2009; Neiman et al. 2009). Bdelloid rotifers were shown to engage in non-canonical forms of sex (Gandolfi et al. 2003; Signorovitch et al. 2015; Schwander 2016; Debortoli et al. 2016; Vakhrusheva et al. 2020). Like the bdelloid rotifers, Darwinulid ostracods arose from a single transition to asexuality, limiting comparative analyses (Tran Van et al. 2021; Wallace and Snell 1991; Ricci 1987). Thus, the identified effects might be lineage-specific. In contrast, oribatid mites provide evolutionary replicates as they have lost sex multiple times independently through time (Norton and Palmer 1991; von Saltzwedel et al. 2014). This enables a comparative analysis of sexual and asexual species, helping to disentangle lineage-specific effects from the true consequences of ancient asexuality.

Oribatid mites are thus a suitable non-model system to reveal the prerequisites for persistence without sex.

Unraveling the Genomic Mysteries of Asexual Speciation

Speciation is the formation of new and distinct species in the course of evolution. The origin of species was once referred to as the "Mystery of Mysteries" by Charles Darwin (Darwin 1909). Since then, speciation research has become one of the most active areas of evolutionary biology. Understanding speciation requires knowledge of whether, when and why conditions favoring divergence arise and how the genetic response of organisms drives reproductive isolation and species separation. Speciation is a dynamic process. Two irrevocable processes lead to speciation: first, gene pool division in the form of reduced or absent gene flow. This can happen through geographical isolation (allopatry), differentiation of a new niche (sympatry), or uneven gene flow (parapatry). But these barriers are not to be absolute, as long as divergence is more promoted than gene flow and cohesion (Hernández-Hernández et al. 2021; Barraclough 2019). Ecological divergence seems to play the driving role and reproductive isolation the facilitating role in independent evolution (Barraclough 2019). However, neither seems to be sufficient alone for creating the pattern of species (Sobel et al. 2010). Secondly, there is a genetic response to the new conditions, which leads to discrete entities. Explaining phenotypic differentiation in a population and diversity in nature. Shedding light on the steps leading to reproductive isolation and subsequent species diversification in sexual organisms has been a fascinating and challenging topic in research for a very long time.

These ideas of mechanisms were developed with sexual organisms in mind. Therefore recombination and reproductive isolation have played crucial roles in theories of speciation. Even though sexual reproduction may result in stronger clustering patterns compared to their asexual counterparts, sex per se is not a prerequisite for speciation. Despite the vast efforts to understand the genomic basis of speciation in sexual organisms, the genomic basis of speciation in asexual organisms is completely unknown. Mechanisms leading to speciation differ under asexuality, as asexual lineages are reproductively isolated per definition. However, reproductive barriers might still affect asexuals as they not only impose premating but also postmating isolation. This includes prezygotic and postzygotic effects, which entail divergent selection, and genetic incompatibilities dependent on cytology (Seehausen et al. 2014). Similar effects might also influence parthenogenetic embryos. The understanding of asexual speciation and its underlying molecular mechanisms and the impact of the cytological mechanism of asexuality are still in their infancy. Compared to sexuals there are only a few known cases of asexual speciation. Bdelloid rotifers are an ancient asexual group that includes more than 400 species defined by phenotype. The taxonomic challenges presented by asexually reproducing organisms have led researchers to explore novel species concepts and criteria tailored to their unique

reproductive modes (Barraclough et al. 2003; Birky et al. 2010). Hence, the phylogenetic and evolutionary species concepts can be applied to asexually reproducing organisms.

An understanding of what asexual species are and how to define them is essential for testing theories about the evolutionary advantage of sex and the consequences of losing sex. Criteria of species concepts are i) initial separation ii) cohesion iii) monophyly and iv) distinguishability.
There are some predictions regarding possible differences between the processes of sexual and asexual adaptation and speciation. Firstly, asexuals should diversify more readily than sexuals. Having no reproductive barriers is a requirement for sexual reproduction like the production of gametes and pairing of homologous chromosomes during meiosis (Barraclough 2019). Reproductive isolation is generally regarded to be a limiting factor for speciation in sexuals (Coyne et al. 2004). In asexuals, this mechanism is non-existent and therefore secures reproductive isolation between diverging populations. Secondly, it is predicted that sexual populations might diverge more rapidly than asexuals. Recombination enables the spread of beneficial gene combinations. This may allow sexual adaptation at a higher pace and promote greater clustering than in asexuals (McDonald et al. 2016; Barraclough 2019).
Genomes are highly organized on multiple levels in the cell. The folding and unfolding of genomic DNA are known to influence various basic biological processes (Cavalli and Misteli 2013). The DNA compaction machinery reorganizes chromosomes to allow homologous chromosome pairing. In contrast to chromosome dynamics in sexuals, asexuality releases constraints from chromosome pairing. Therefore, genome structure can be more plastic in asexuals than in sexuals (McElroy et al. 2021). Variations in gene copy number, insertions, and deletions between homologous chromosomes lead to a potential for evolutionary novelty and diversification that is not as easily given in sexuals. Especially with speciation, it is known that structural rearrangements influence differentiation, which is crucial for sexual organisms to overcome the effect of recombination. Therefore, an in-depth assessment of ploidy and haplotype structure is essential. Moreover, the correlation between the diversity of expression patterns found in asexuals, which result from gene family expansions and traces of positive selection, will give interesting clues to the success of ancient asexuals. Overall, characterizing the genomic features that are contributing to the differentiation and speciation in ancient asexuals will help to understand how asexuals can adapt and therefore persist over time.

Platynothrus peltifer as a Model Organism to Study Ancient Asexuality and Speciation

An organism that combines the possibility to study ancient asexual persistence and speciation is the oribatid mite *Platynothrus peltifer.* It has a single generation per year, with a life cycle of 170 days from hatching to adult emergence (Jalil 1972). It likely reproduces parthenogenetically at least for tens of millions of years, probably even longer (Heethoff et al. 2007). Reproduction was suggested to be by

automictic thelytoky with terminal fusion (Taberly 1987), with one to seven (unfertilized) eggs (Jalil 1972) between March and September (Grandjean 1950). Inverted meiosis was proposed to be the driving force for fixed heterozygosity in natural populations (Heethoff et al. 2007). Within this parthenogenetic species, substantial divergence occurred leading to its radiation despite the absence of sex (Figure 1). The distribution of extant populations reflects continental drift (Hammer and Wallwork 1979) and has been shown to exhibit mitochondrial diversification and morphological flexibility (Heethoff et al. 2007), highlighting *P. peltifer* uniqueness in investigating asexual speciation.

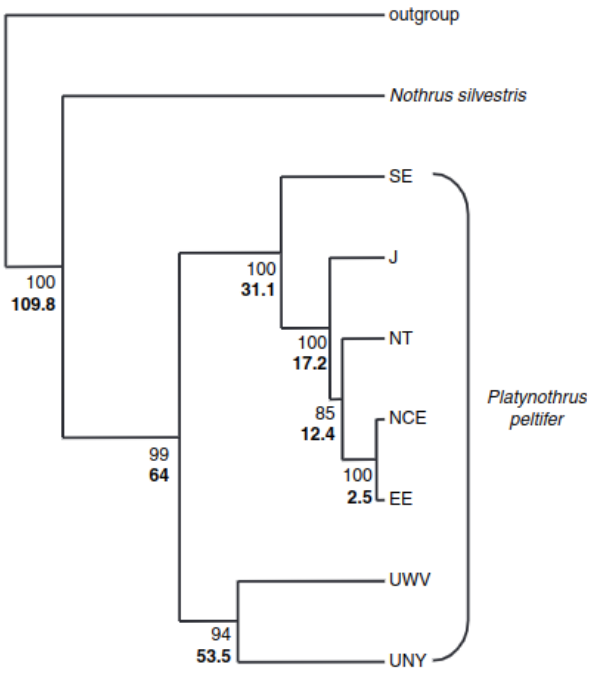


Figure 1: Excerpt from Heethoff et al., 2007 showing the seven different evolutionary lineages of *Platynothrus peltifer*, with molecular clock assumption. Numbers indicate bootstrap values. Bold numbers indicate the age (Myr). Southern Europe (SE), Japan (J), Nothern Tyrolia (NT), Nothern/Central Europe (NCE), Eastern Europe (EE), USA West Virginia (UWV), USA New York (UNY).

Aims and Outline of this Study

In this thesis, I unraveled novel ways of evolvability of ancient asexuals and shed light on asexual speciation. To address this I propose three hypotheses that underpin our exploration of the genomic features associated with ancient asexuality:

1) *Platynothrus peltifer* reproduces in the absence of sex for millions of years which is reflected by unique evolutionary patterns of independent haplotype evolution (i.e. the 'Meselson effect').
2) Such independent haplotype evolution could potentially contribute to evolvability.
3) Asexual adaptation and diversification into novel species proceed differently from sexual species and involve structural variation.

To test these hypotheses, I first established a protocol for single-individual high-molecular-weight *g*DNA extraction from minuscule non-model organisms (Chapter 1). This formed the basis for in-depth, genome-wide analyses for which I generated genome, transcriptome, and population data from five individuals from five distinct populations each. Second, I generated a haplotype-resolved chromosome-scale genome assembly of the obligate ancient asexual oribatid mite *Platynothrus peltifer*. Further, I utilized the newly generated population data to reveal increased intra-individual haplotype divergence, suggesting evolution without sex for 20 my. This haplotypic independence might contribute to evolvability in these ancient asexuals (Chapter 2). In-depth population analyses additionally show conserved genome synteny and heterozygosity patterns among worldwide populations (Chapter 3). Moreover, assembling the genomes of the sexual outgroup *Hermannia gibba* and asexual outgroup *Nothrus palustris,* establish these species as part of a broader comparative approach in oribatid mites to reveal common features in ancient asexuals and their independent transitions to asexuality (Chapter 3-4). Additionally, identifying the sex-determination mechanisms in *H. gibba* provided an ancestral example in sex-chromosome evolution (Chapter 4). Lastly, I emphasize the importance of scientific outreach and present research addressing children, introducing diverse reproductive modes in soil, highlighting the consequences of asexuality, and milestones of oribatid mite research (Chapter 5).

Contributions to the Chapters of this Thesis

To describe specific contributions by the author of this thesis CRediT (Contributor Roles Taxonomy) was used. Details can be found here: https://credit.niso.org/

Chapter 1

Challenges in Genomic Studies of Oribatid Mites and the Need for High-Quality Assemblies of Single-individuals

Protocols.io (2023) https://doi.org/10.17504/protocols.io.yxmvm3yybl3p/v1

Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing

Chapter 2

Haplotype Independence Contributes to Evolvability in the Long-term Absence of Sex in a Mite

bioRxiv (2023) https://doi.org/10.1101/2023.09.07.556471 *Submitted to Science*

Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing - original draft, Writing - review & editing

Chapter 3

Ancient Asexual Speciation in the Oribatid Mite *Platynothrus peltifer*

in preparation

Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft

Chapter 4

A Female Heterogametic ZW Sex-determination System in Acariformes

bioRxiv (2023) https://doi.org/10.1101/2023.10.24.563255

Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing - original draft

Chapter 5

Science Outreach and Empowerment: Bridging the Gap for Diverse Audiences

Front. Young Minds. (2021) https://doi.org/10.3389/frym.2021.611659

Conceptualization, Project administration, Visualization, Writing - original draft, Writing - review & editing

Chapter 1

Challenges in Genomic Studies of Oribatid Mites and the Need for High-Quality Assemblies of Single-individuals

Abstract

Answering questions of evolutionary biology necessitates the use of population genetic studies, for which single-individual approaches are often needed. Recent studies have made progress, allowing research on previously unattainable organisms. Yet, the genome quality of such organisms lacks continuity, which remains a bottleneck for in-depth analysis. This study uses oribatid mites, a group of minute soil organisms, to tailor a *g*DNA extraction protocol, which now yields >20 ng of high-molecular-weight *g*DNA. Thus, overcoming limitations associated with small invertebrate organisms. Results showcase Femto Pulse Chromatograms and HiFi read metrics, demonstrating excellent read quality and continuity. Overall, the established protocol demonstrates its ability to generate ample DNA for *de novo* genome assembly, empowering researchers to delve into the genomics of small invertebrate organisms.

Introduction

The sequencing of DNA has revolutionized the field of molecular biology. The development of several waves of sequencing technologies impacted our understanding of genome evolution. From the early days of Sanger sequencing to the emergence of second-and third-generation sequencing, each new technology has enabled new possibilities. Ultimately, we entered an era where we can study *any* organism and untangle its genomic diversity and evolutionary history. Single-individual genomics is a powerful tool, that addresses evolutionary theories in natural environments. One such theory deals with the divergence of alleles over evolutionary time in the absence of sex, a phenomenon observed in ancient asexual diploid organisms. To explore the independent evolution of haplotypes, phased continuous genomes are needed. However, this presents a challenge when studying ancient asexuals like oribatid mites, which are tiny soil-living invertebrates. The difficulty of obtaining high-molecular-weight genomic DNA (HMW *g*DNA) for these organisms, resulted in discontinuous genomes. Compared to Illumina sequencing, which can manage poor DNA quality and quantity, long-read procedures have higher requirements: technologies like Pacific Biosciences (PacBio) Single Molecule

Real-Time (SMRT) sequencing requires 3 µg/ 1 Gb genome for standard HiFi Sequencing or 5 ng with >20 kb HMW *g*DNA for the Ultra-Low DNA Input Sequencing (Application Note — PacBio 2021). Moreover, there is a lack of pure HMW *g*DNA extraction protocols accounting for difficulties in extracting DNA of invertebrates, i.e. chitinous exoskeleton, small size, and high contamination (internal (gut) and external). Commercially available kits promise easy, quick, and efficient HMW *g*DNA extractions (e.g. Wizard HMW DNA Extraction Kit (Promega, Madison, USA), NEB Monarch® High Molecular Weight DNA Extraction Kit (New England Biolabs, Massachusetts USA), Quick-DNA HMW MagBead Kit (Zymo Research, California USA) and MagAttract HMW DNA Kit (Qiagen, Netherlands)). However, they are not optimized to work with invertebrate organisms, which usually are tiny and cannot be reared in the lab. Nevertheless, invertebrates harbor more diversity than vertebrates (Eisenhauer & Hines, 2021). Thus to uncover such diversity, a particular HMW extraction protocol is needed.

Case study: *Platynothrus peltifer*, Camisiidae, Acariformes, is a small (<1 mm) invertebrate of interest in evolutionary biology due to its successful persistence of life without sex (Judson & Normark, 1996; Maraun et al., 2003). With a life cycle of 170 days at 25°C (Jalil, 1972), *P. peltifer* cannot be cultured in the lab and must be obtained from a natural population in the wild. For wild-caught organisms in general, the likelihood of including cryptic species (and or different genetic lineages) is imminent and greatly impairs assembly quality. As *P. peltifer* is proposed to reproduce asexually for tens of millions of years (Heethoff et al., 2007), high levels of heterozygosity are expected. Here, I provide an HMW *g*DNA extraction protocol for small arthropods, tested on oribatid mites, which generates >20 ng of HMW *g*DNA of a single individual (~700 µm). The quality of *g*DNA allows for long-read sequencing techniques like PacBio HiFi sequencing, which is used as a basis for the generation of phased *de novo* assemblies.

Material and Methods

Sample Collection

Details of sample collection and sample preparation are given in Chapter 2. In short: five different populations of *Platynothrus peltifer* were collected. The Russian, Japanese, and Canadian populations were sent by collaborators from Moscow (Central Russia), Yamanashi (Chūbu), and Moncton (New Brunswick) respectively. The German population of *P. peltifer* and its sexual outgroup *Hermannia gibba* were sampled in Dahlem (North Rhine-Westphalia). The Italian population was sampled in Montan (Trentino). Additionally, specimens of *Nothrus palustris,* which is a closely related asexual outgroup of *P. peltifer,* were sent from Göttingen (Lower Saxony, Germany).

High-molecular-weight (HMW) *g*DNA extraction protocol

Specimens were cleansed prior to DNA extraction. Cleansing followed a newly established workflow, including cleansing specimens with detergent, ethanol, and bleach. Optimization of the salting-out method for DNA extraction for oribatid mites included modifications on the lysis of cells and sample purification. To identify the most efficient lysis of a single individual tests whether the disruption of tissue was more efficient if conducted with a razor blade or pestle and whether sonification of tissue increases DNA yield were conducted. Additionally tested was if an increase in time, temperature, and amount of proteinase K increases digestion. Furthermore, the influence of various volumes of buffer was tested.

HMW DNA Quality Assessment

The quality of HMW *g*DNA was assessed using a pulsed-field power supply like the Femto Pulse System (Agilent), which is more sensitive than other gel-based instruments, especially when low concentrations of DNA are expected. Samples were chosen if size distribution depicted HMW DNA smears and fragments ≥165 kb.

Sequencing

For details on sequencing see Chapter 2. In short: for the *de novo* assembly individuals were long-read sequenced with PacBio HiFi (SMRTbell® Libraries from Ultra-Low DNA Input). Plots on read length and quality of HiFi reads were generated using NanoPlot v1.41.0 (De Coster & Rademakers, 2023).

Assessing genome complexity of species

27-mers in the HiFi reads were analyzed using KAT v2.4.2 (Mapleson et al., 2017) with the modules kat hist and kat gcp (default parameters). Ploidy was further investigated using kmc v3.2.1 with parameters -k 27 -ci 1 -cs 10000 and Smudgeplot v0.2.5 (Rhyker Ranallo-Benavidez et al., 2020) with default parameters. To obtain a frequency count of *k*-mers Jellyfish v2.2.8 (Marçais & Kingsford 2011) was used as input to estimate genome heterozygosity and repeat content with GenomeScope (http://qb.cshl.edu/genomescope/; last access 17.10.23; (Rhyker Ranallo-Benavidez et al., 2020).

Results

The high-molecular-weight genomic DNA (HMW *g*DNA) extraction protocol was tested on more than 360 individual oribatid mites, mainly on *Platynothrus peltifer*. Additionally tested specimens include *Hermannia gibba*, and multiple species of genera such as *Nothrus*, *Camisia*, *Heminothrus*, *Pergalumna*, *Carabodes*, *Hydrozetes*, and *Steganaccarus*.

Establishment of single individual low-input high-molecular-weight *g*DNA extraction

Lysis of oribatid tissue and exoskeleton was most efficient with at least 3x freeze and thaw cycles followed by disruption of the whole individual with a pestle. Sonification did not yield higher results and disrupting tissue with a razor blade even led to decreased DNA yield. Incubation time and temperature of DNA digestion with proteinase K did not show verifiable effects on DNA yield. It is advised to follow the manufacturer's instructions for optimal proteinase K digestion. Furthermore, ethanol volume during purification was most efficient with ~2 volumes of sample.

The following protocol is the now-established high-molecular-weight (HMW) *g*DNA extraction method in the working group. Further, it was shared with colleagues worldwide and successfully used on Collembolans and even Nematodes (personal communication Dr. Kamil Jaron, June 2023). It enables the generation of HMW *g*DNA of ~20 ng, even up to 200 ng of a single oribatid mite, depending on the genus.

Modified salting out method for high-molecular-weight *g*DNA extraction (oribatid mites)

Hüsna Öztoprak* & Jens Bast

PROTOCOL integer ID: 88360 - Protocols.io (2023)

*Corresponding author

Abstract

This protocol describes a low-cost, high-molecular-weight genomic DNA extraction method for a single minuscule specimen (modified from Miller et al., 1988). DNA extractions from oribatid mites are typically challenging, because of the small body size of 150-1400 µm. Their chitinous exoskeleton does not dissolve during DNA extraction, impedes DNA purification, and leads to additional loss of DNA. Therefore, high-molecular DNA from oribatid mites has been thus far unattainable, especially from single individuals. We established a high-molecular-weight *g*DNA extraction protocol for mites that enables the generation of high-quality phased genomes for small non-model organisms. There are three options to utilize this protocol: i) for high-molecular *g*DNA extraction ii) for high-molecular *g*DNA extraction, while preserving the exoskeleton for morphological analysis, and iii) DNA extraction with Chitinase to yield more *g*DNA.

As specimens are collected from natural populations and are not cultured in the lab. They are cleansed prior to DNA extraction to minimize external contamination. Cleansing includes brushing the specimen in distilled water and in distilled water with detergent (fit GmbH, Zittau, Germany). Once the external residue is not visible anymore, the specimen is incubated in NaCIO 0.05% (DonKlorix; CP GABA GmbH, Hamburg, Germany) and ethanol 70% for 30 seconds each and rinsed in distilled water again.

Keywords: Sample preparation, DNA extraction, Low-cost, low-input, HMW DNA

Material and Regents

1. Proteinase K (Qiagen, catalog number: 19131)
2. Yeast tRNA (Invitrogen, catalog number: AM7119)
3. RNase Cocktail (Invitrogen, catalog number: AM2286)

Solutions

1. TNES buffer (see Recipes)

Recipes

1. Final concentration of TNES buffer freshly made before each extraction, ddH_20 used to dilute.: 400 mM NaCl, 20 mM EDTA, 50 mM Tris pH 8.0, 0.5 % SDS

Procedure

Version i) High-molecular *g*DNA extraction

DNA Extraction

- Submerge cleansed specimen in 195 µl TNES buffer and flash freeze by holding tube in liquid nitrogen.
- Using a sterile pestle, homogenize by applying pressure to grind the specimen between pestle head and the walls of the tube.

If low DNA yield is expected. Leave pestle in tube to ensure maximum digestion of material.

Consider including multiple (3x) freeze and thaw cycles and vortexing to disrupt tissue fully.

- Add 5 µl proteinase K.
- Vortex for 3 sec. Centrifuge briefly.
- Incubate at 55°C for ~60 min (completely dissolve the specimen).

If there is indigestible debris left over centrifuge (18000 rcf for 5 min) and transfer the supernatant to a fresh tube.

- Add 1.5 µl yeast tRNA, flick to mix briefly then spin down.

- Add 65 µl 5 M NaCl and 290 µl 96 % EtOH, mix by inversion.
- This should clarify the solution. Store at -20°C for 1h.

Solution can be stored overnight at this step.

DNA Purification

- *Optional*: add 1 µl Pellet Paint Co-Precipitant *(for colorful pellet)*
- Spin down at 18000 rcf for 15 min at 4°C.

Know the expected position of the DNA pellet, as it can be difficult to see.

- Ghostly pellet should be visible.
- Remove supernatant.
- Add 0.5 ml chilled 70 % EtOH (make fresh).
- Spin at 18000 rcf for 5 min.
- Repeat ethanol rinse.
- Carefully remove supernatant.
- Leave tube open to air dry. Pellet should have a glassy appearance.
- Using a wide-bore pipette tip, add 21 µl TE Buffer (elution buffer) and gently resuspend the DNA pellet with pipette mixing.
- Let DNA resuspend at 4°C overnight.
- Add 2 µl RNase Cocktail. Incubate at 37°C for approx. 60 min.

Note: For femto pulse systems elution in TE Buffer, more specifically EDTA is not recommended. Alternatively use 0.1 mM EDTA or EB, Tris-HCl (pH 8-8.5)

Version ii) High-molecular gDNA extraction preserving exoskeleton

DNA Extraction

- To observe the specimen under a microscope, cleansed specimens are placed on a sterile slide and submerged with TNES buffer until fully covered.
- Remove one genital plate cautiously with a sharp needle and stir tissue without destroying the exoskeleton.
- Transfer specimen in 195 µl TNES buffer.
- Add 5 µl proteinase K.
- Incubate at 37°C overnight.
- Transfer specimen with sterile needle to a tube containing 70 % EtOH and store it for morphological analysis.

- Add 1.5 µl yeast tRNA, flick to mix briefly then spin down.
- Add 65 µl 5 M NaCl and 290 µl 96 % EtOH, mix by inversion.
- This should clarify the solution. Store at -20°C for 1h.

DNA Purification follows the same procedure as above.

Version iii) *g*DNA extraction with chitinase digestion

DNA Extraction

- Submerge specimen in 195 µl TNES buffer and flash freeze by holding tube in liquid N.
- Using a sterile pestle, homogenize by applying pressure to grind the specimen between pestle head and the walls of the tube.
- Add 2 µl chitinase (1 U/ml).
- Vortex for 5 sec. Centrifuge briefly.
- Incubate at 55°C for ~60 min.
- Add 5 µl proteinase K.
- Vortex for 5 sec. Centrifuge briefly.
- Incubate at 55°C for ~60 min (completely dissolve the specimen).

If there is indigestible debris left over centrifuge (18000 rcf for 5 min) and transfer the supernatant to a fresh tube.

- Add 1.5 µl yeast tRNA, flick to mix briefly then spin down.
- Add 65 µl 5 M NaCl and 290 µl 96% EtOH, mix by inversion.
- This should clarify the solution. Store at -20°C for 1h.

DNA Purification follows the same procedure as above.

HMW DNA Quality Assessment

Henceforth, shown samples of *Nothrus palustris* (Nps), *Hermannia gibba* (Hga), and *Platynothrus peltifer* (Ppr) were chosen for HiFi sequencing and later used for *de novo* generation of assemblies. The integrity and purity of *g*DNA were assessed using the Femto pulse System (Agilent). The Femto Pulse gel images and traces of the seven samples show fragments with the majority of DNA >20 kb, fulfilling the minimum requirement for generating PacBio HiFi reads for *de novo* assembly (Figure 1). All the samples of *N. palustris*, *H. gibba* and *P. peltifer* from Germany (DE), Italy (IT), Russia (RU), and Japan (JP) show a narrow peak at a fragment size of around 165,500 bp, except the Canadian Ppr (CA). The majority of Canadian Ppr *g*DNA has 4010 - 15400 bp. All samples show an initial peak of less than 200 bp, most likely representing RNA contamination in the sample, which was separated in follow-up procedures like size selection.

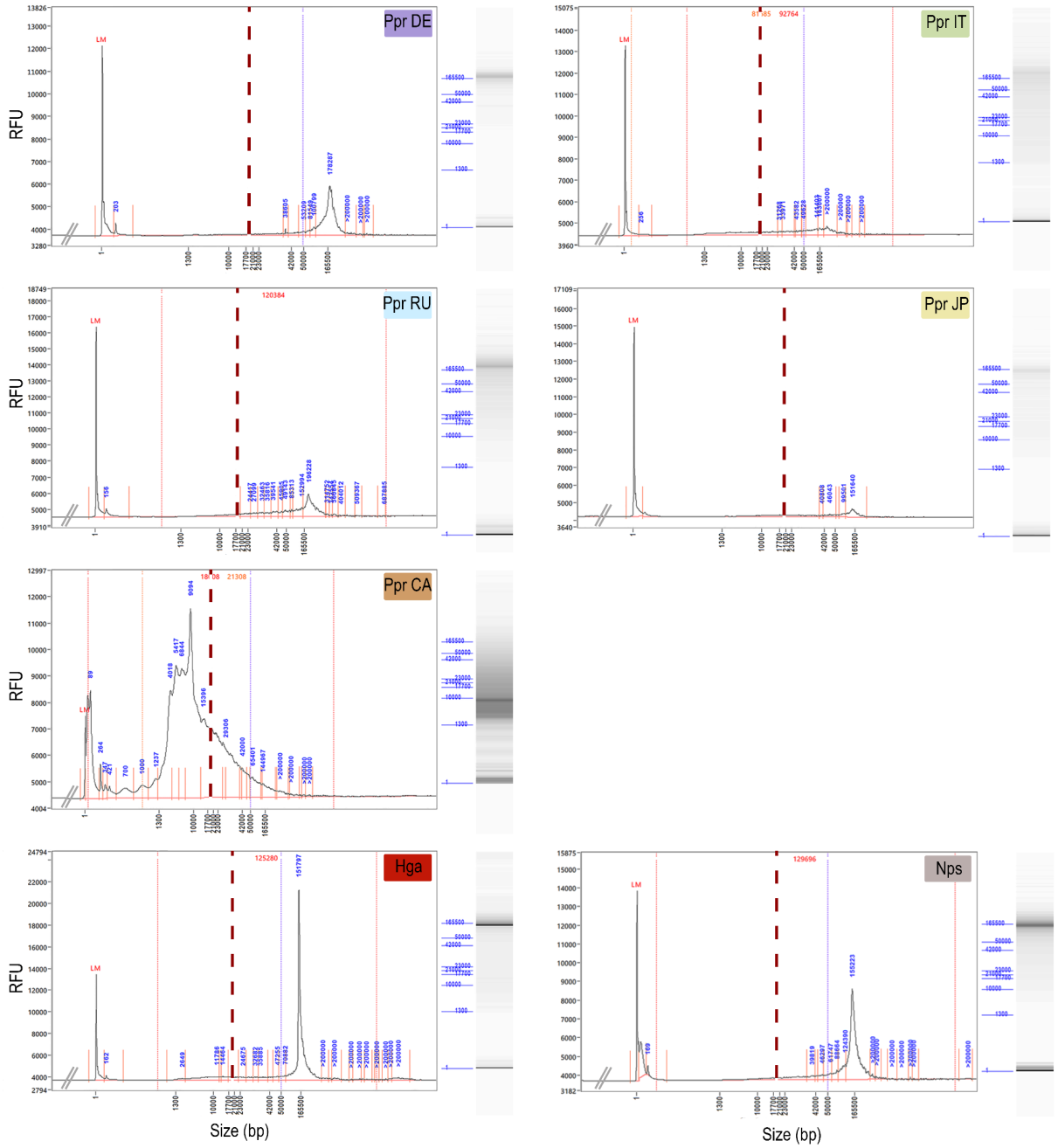


Figure 1: Modified Femto pulse chromatograms of gDNA of *Platynothrus peltifer* (Ppr) sampled from Germany (DE), Italy (IT), Russia (RU), Japan (JP), and Canada (CA) chosen for long-read sequencing prior to size selection. Additionally, one sample of *Hermannia gibba* (Hga) and *Nothrus palustris* (Nps) were chosen. The red dotted line indicates a fragment size of 20 kb, which is the minimum requirement for PacBio HiFi SMRTbell® Libraries.

Long-read sequencing

The minimum amount of 8 ng of HMW *g*DNA of a single specimen was used for PacBio HiFi sequencing. It resulted in at least 2.3 M reads for each individual, which equals approximately 100X coverage (Figure 2a). For the Canadian *P. peltifer* and *H. gibba* more than 3 M reads were sequenced. The mean read length across all samples was 12,132 bp (Figure 2b), ranging from 6,153 bp (Ppr CA) to 14,709 (*N. palustris*). The read length for the Canadian *P. peltifer* was half as high compared to the other samples, with an N_{50} of 6,539 (Figure 2c). The German *P. peltifer* had an N_{50} of 15,100 bp, *N. palustris* of 15,429 bp and *H. gibba* of 15,041 bp. The median base call quality score was in all samples above 30, ranging from 31.2 (Ppr DE) to 42.3 (Ppr CA; Figure 2d). This resulted in a probability of incorrect base call of 1 in 1000, meaning 99.9% inferred base call accuracy.

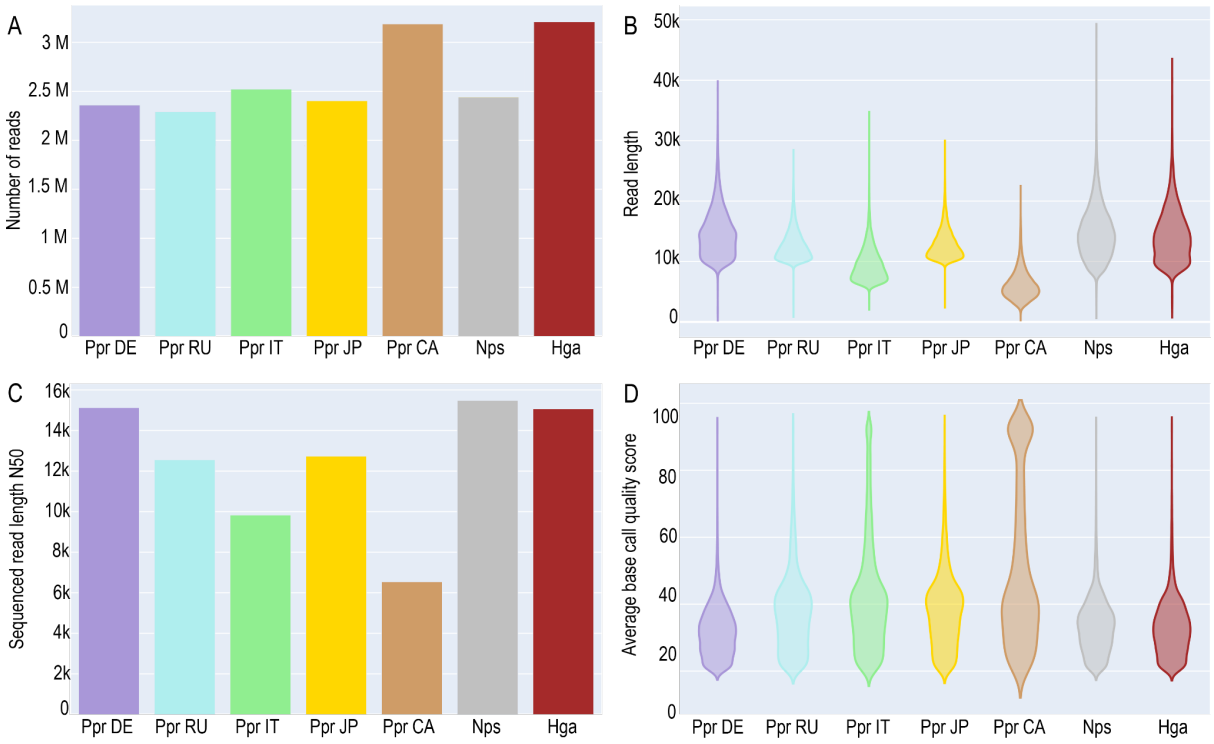


Figure 2: Comparison of HiFi reads generated with the established protocol for high-molecular-weight gDNA extraction for oribatid mites. Plots visualize **a**) number of reads, **b**) read length, **c**) N_{50}-read-length, and **d**) the base call quality score. Color code: red for *Hermannia gibba* (Hga), grey for *Nothrus palustris* (Nps), purple, cyan, green, yellow and brown for *Platynothrus peltifer* (Ppr) populations from Germany (DE), Russia (RU), Italy (IT) Japan (JP) and Canada (CA) respectively.

To judge sequencing complexity such as ploidy, heterozygosity and genome duplications of each species we conducted *k*-mer analyses on the generated HiFi reads of German *P. peltifer*, *N. palustris*, and *H. gibba* using a *k*-mer size of 27.

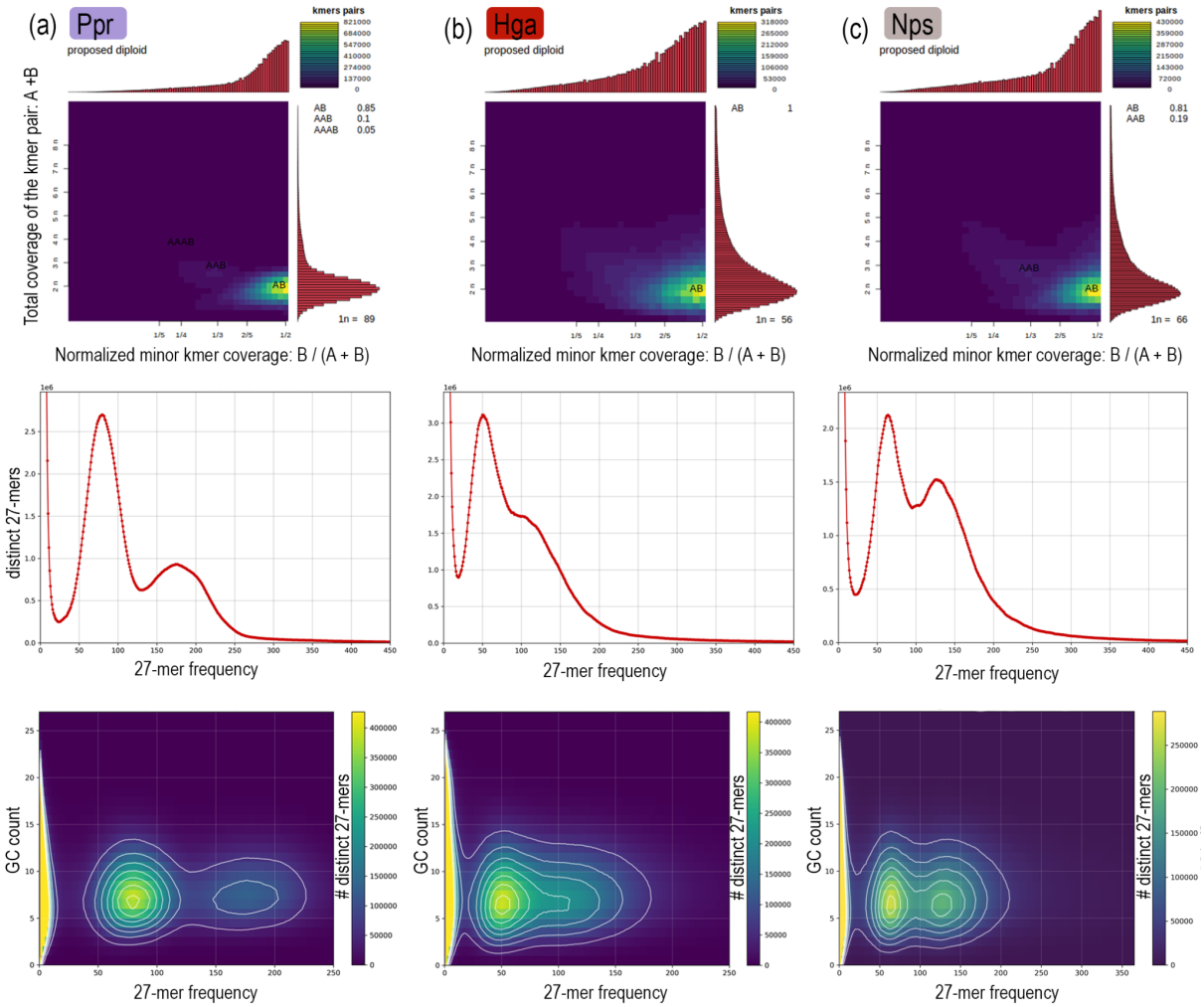


Figure 3: *k*-mer analysis of the PacBio HiFi reads (*k*=27) of **a**) *Platynothrus peltifer* (Ppr), **b**) *Hermannia gibba* (Hga), *and* **c**) *Nothrus palustris* (Nps). Distinct smudges of *k*-mers with an A/B configuration identify the genomes as diploid without signs of genome duplication. Heterozygous and homozygous content is represented in two peaks respectively in the histograms. Plots show GC content of these two peaks, with no residue of DNA contamination.

The smudgeplots for each species showed a single smudge with an A/B configuration, identifying each species as diploid (Figure 3). Their *k*-mer spectra histograms showed two peaks corresponding to the heterozygous *k*-mers and the homozygous *k*-mers respectively, suggesting diploidy in all species. The asexual *N. palustris* exhibited an elevated homozygous peak (Figure 3c). The homozygous peak in sexual *H. gibba* was not prominent (Figure 3b). The *k*-mer coverage vs GC count plot showed a peak

at 50-100X for *P. peltifer*, 50-75X for *N. palustris*, and 50X coverage for *H. gibba* with a GC spread from 2 to 15 in all species. There was no additional peak of higher GC detectable, indicating no significant contamination in any reads.

GenomeScope 3.0 estimated genome size, heterozygosity and repetitiveness based on *k*-mer frequencies for *P. peltifer*, *H. gibba* and *N. palustris* (Table 1).

Table 1: Genomescope statistics on HiFi reads from *Platynothrus peltifer* (Ppr), *Hermannia gibba* (Hga) and *Nothrus palustris* (Nps).

	Ppr	Hga	Nps
expected genome length [bp]	171,269,487	313,701,730	231,232,057
expected heterozygosity [%]	2.35	1.77	1.33
coverage	91	62	71
duplication/repetitiveness [%]	4.76	7.29	6.44

With a genome length of ~171 Mb for *P. peltifer*, 231 Mb for *N. palustris* and 314 Mb for *H. gibba*, genome size is, compared to other arthropods (from 98 Mb to 18.6 Gb; Kelley et al., 2014; Gregory, 2023) small in these oribatid mites. The heterozygosity of *P. peltifer* was highest at 2.35 % and lowest in *N. palustris* at 1.33 %. Following this, the increased heterozygosity was also visible in a higher first peak and a lower second peak of distinct 27-mers in the histograms (Figure 3). Duplication was estimated to be highest in *H. gibba* at 7.29 % and lowest in *P. peltifer* at 4.76 %, which can also be seen in the higher coverage peaks of *H. gibba* and *N. palustris* in the *k*-mer spectrum representing increasingly high copy repetitive sequences in the genomes.

Discussion

Understanding evolution in natural populations requires high-quality and contiguous genomes. Specifically, in oribatid mites previous studies lacked high-quality assemblies (Bast et al., 2016; Brandt et al., 2021; Brückner et al., 2022). Therefore, establishing a distinct extraction protocol of high-molecular-weight genomic DNA (HMW *g*DNA) based on single individuals was essential for genomic studies on these ancient asexual scandals. During the development of the protocol, it became clear that the adjusted salting out method was the only known method that could yield the quality and quantity needed for single individual genome sequencing in oribatid mites. The extracted HMW *g*DNA quality satisfied the strict requirements for long-read sequencing. It generated high-quality reads, thus underlying the capability of the protocol to extract DNA useful to generate *de novo* assemblies of

invertebrate organisms. Even the initially more fragmented DNA from the Canadian *Platynothrus peltifer* sample generated HiFi reads with comparably high base call quality to all the other samples. Although, oribatid mite DNA is known to be associated with fungal taxa like Basidiomycota, Ascomycota and Zygomycota (Renker et al., 2005) and harbor a substantial amount of body-surface contamination (Remén et al., 2010), no excessive contamination was detected in our samples. We conclude that starving the specimen prior to DNA extraction and body-surface cleansing with detergent, bleach and ethanol is an essential step in minimizing contamination in sample preparation.

We were able to confirm *P. peltifer*, *Hermannia gibba*, and *Nothrus palustris* as diploid organisms. Distinct levels of heterozygosity were detected among *P. peltifer*, *H. gibba* and *N. palustris*, supporting the applicability of the extraction protocol for studying heterozygous organisms and specifically genetic diversity in oribatid mites. Concluding that the established protocol for HMW *g*DNA extraction provides ample DNA for HiFi sequencing useful to generate high-quality genome assemblies of small invertebrate organisms.

Chapter 2

Haplotypic independence contributes to evolvability in the long-term absence of sex in a mite

Öztoprak, H.[†*], Gao, S.[†], Guiglielmoni, N., Brandt, A., Zheng, Y., Becker, C., Becker, K., Bednarski V., Borgschulte, L., Burak, K.A., Dion-Côté, A-M., Leonov, V., Opherden, L., Shimano, S. & Bast, J.[*]

bioRxiv (2023)
Submitted to Science (2023)

† Shared first authorship
*Corresponding authors

Abstract

Some unique asexual species persist over time and contradict the consensus that sex is a prerequisite for long-term evolutionary survival. How they escape the dead-end fate remains enigmatic. Here, we generated a haplotype-resolved genome assembly based on a single individual and collected genomic data from worldwide populations of the parthenogenetic diploid oribatid mite *Platynothrus peltifer* to identify signatures of persistence without sex. We found that haplotypes diverge independently since the transition to asexuality at least 20 my ago. Multiple lines of evidence indicate disparate evolutionary trajectories between haplotypic blocks. Our findings imply that such haplotypic independence can lead to non-canonical routes of evolvability, helping some species to adapt, diversify and persist for millions of years in the absence of sex.

One-Sentence Summary

Functionally different chromosome sets in an asexual mite species showcase a survival strategy spanning millions of years.

Some oribatid mite species are rare evolutionary anomalies, as they maintain effective purifying selection and persist and diversify over millions of years in the absence of sex (*1–4*). Oribatid mites are diverse (>10,000 species), small (150-1400 µm), mainly soil-living decomposers that were among the first arthropods to colonize land during the Devonian (*5*, *6*). Notably, a high number of species in this animal order (10 %) reproduce via parthenogenesis, of which many radiated and form diverse phylogenetic clades (*7*, *8*). Such evolutionary exceptions are invaluable because understanding the peculiarities for success without sex will help to identify the adaptive value and constraints of sex vice-versa (*9*). To date, however, the population-genomic signatures of persistence without sex, i.e. how peculiar genome dynamics of ancient asexuals might contribute to maintaining effective selection, adaptation and diversification remain largely unknown (*1*, *3*, *10*).

A phased genome assembly to identify signatures of ancient asexuality

One of the most iconic old asexual species is the diploid oribatid mite *Platynothrus peltifer* (C.L. Koch, 1893; **Fig. 1A**). Previous work suggested a transition to asexuality tens of millions of years ago, likely predating the separation of Europe and North America (*2*). The cytological mechanism for asexuality is a form of automictic thelytoky during which recombination is restricted to chromosome ends, maintaining heterozygosity and resulting in 'effective clonality' (*11–13*). *Platynothrus peltifer* has a generation time of one year and produces only female offspring in lab rearings, and extremely rare spanandric males are infertile (*14*, *15*).

To reveal the genomic substrate for persistence without sex, we first generated a chromosome-scale reference genome from a natural population and additionally resolved haplotypes of *P. peltifer* from a single individual, sampled in Germany (**Fig. 1**). In short, PacBio HiFi long reads and TELL-Seq linked reads stemming from one diploid individual were assembled into haplotype-collapsed scaffolds that were ordered to chromosome-scale with chromosome conformation capture (Hi-C) data from pooled individuals of the same population. The same long-read and linked-read data were used to generate a haplotype-resolved assembly; following, phased haplotypic blocks were anchored to each other and the chromosome-scale assembly (**Fig. 1B, fig. S1)**. The chromosome-level genome assembly is of high completeness and contiguity and spans 219 Mb comprising nine chromosomes with 24,932 annotated genes (**Fig. 1C, fig. S2, fig. S3, table S1**). The haplotypes comprise 63 blocks B anchored to 44 blocks A, representing 92.7 % (203 Mb) of the genome being phased (**Fig. 1C, table S2**). The genome size estimate via flow cytometry closely matches the assembly size (6 % difference: 232 Mb vs 219 Mb) and together with *k*-mer spectra ploidy analyses suggest no ancient whole genome duplication (**fig. S4, supplementary text**). Additionally, we assembled the complete mitochondrial genome from the same individual (**fig. S5**).

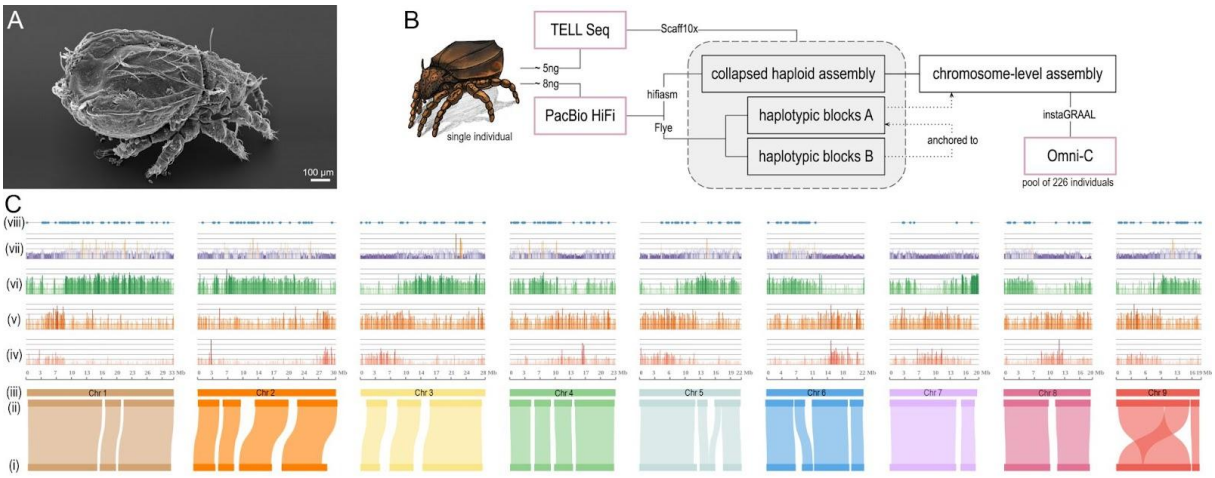


Fig. 1. Assembly strategy and characteristics of the *Platynothrus peltifer* genome. (**A**) *P. peltifer* electron microscopy picture. Scale bar indicates 100 µm. (**B**) Schematic workflow of single-individual sequencing and assembly. (**C**) Genomic properties of the aligned haplotypic blocks (i) A and (ii) B, anchored to the (iii) nine pseudochromosomes, for which densities of 100 kb blocks of (iv) coding sequences, (v) gene number, (vi) transposable elements, (vii) heterozygosity percentage, and (viii) horizontally transferred gene locations are shown along the genome.

Haplotype dynamics and divergence across worldwide populations

After the transition from sexuality to asexuality, spontaneous mutations are predicted to occur independently on each haplotype in a diploid, 'effectively clonal' asexual. Over time, the absence of sex, recombination and gene flow should thus manifest into increasing intra-individual heterozygosity, which drives the divergence and independent evolution of haplotypes. Consequently, if the transition to asexuality occurred a considerable amount of time ago, such that different populations separated after, haplotypes should be more diverged within individuals than between individuals from different populations, generating haplotype trees mirroring each other's topology. Such independent evolution of haplotypes is the strongest evidence for obligate asexual reproduction over long time periods (aka, the 'Meselson effect') (*16*). However, despite this straightforward prediction, empirical evidence remains equivocal and the only strong evidence for independent haplotype evolution in animals stems from an oribatid mite species (*1*).

To investigate haplotypic independence associated with ancient asexuality in *P. peltifer*, we sequenced five individuals per population from German (Dahlem), Italian (Montan), Russian (Moscow Oblast), Japanese (Yamanashi) and Canadian (Moncton) populations (**Fig. 2A**). Genetic divergence clusters individuals by their geographical locations (**Fig. 2B**). Next, to identify haplotypic differences among individuals and populations, we analyzed genetic diversity patterns over the whole nine chromosomes.

Overall dynamics of mean heterozygosity between haplotypes within individuals are similar among populations, with shared regions of high and low heterozygosity (**Fig. 2C**). This suggests differences in purifying selection or mutation rate heterogeneity among the various chromosomal regions, shared by all populations. German, Italian and Russian (European) populations feature similar levels of mean individual heterozygosity of 1.6 % to 1.8 % along chromosomes. As expected under asexuality, individuals within these populations also share a large proportion of heterozygous sites, with Italian individuals sharing most (**fig. S6**). The mean heterozygosity of individuals from Canadian (1.4 %) and Japanese (1.0 %) populations is consistently lower compared to the other populations, but shows strikingly similar heterozygosity dynamics along chromosomes (**Fig. 2C**). These differences are thus not the consequence of large distinct stretches of loss of heterozygosity (LOH) that can occur under some non-clonal forms of asexuality (*17*). Individuals within these populations moreover share a lower proportion of heterozygous sites, especially Canadian individuals (**fig. S6**). These findings suggest considerable differences in processes that can affect the overall divergence of haplotypes among populations, such as the rate of spontaneous mutations increasing heterozygosity, the rate of gene conversion removing heterozygosity and/or suggest independent and more recent transitions to asexuality for the Japanese and Canadian populations.

Haplotypes evolve independently under asexuality since 20 million years ago

We generated haplotype-specific trees using phased data to identify if the transition to asexuality occurred considerable amounts of time before the separation into different populations (**Fig. 2D**, **fig. S7, table S3**). As expected under ancient asexuality, a perfect split of the two haplotypes displaying mirror topologies of populations could be identified, including all individuals of European populations, representing the 'Meselson effect' (**Fig. 2D**). Contrastingly, haplotype trees of Japanese and Canadian populations lack such clear separation. While some individuals display a shared separation of haplotypes as expected under asexuality, this is confined to each respective population (**Fig. 2D**). Taken together, heterozygosity patterns are corroborated by haplotype topologies (**Fig. 2C, D, fig. S7**) and suggest an ancient transition to asexuality for the ancestor lineage of the European populations and a more recent independent loss of sex for both the Japanese and Canadian populations. Heterozygosity dynamics along chromosomes are very similar in all populations and imply conserved synteny in the ancestral lineage for all transitions (**Fig. 2C**). Thus, while likely not sharing the same transition event, the ancestor of all *P. peltifer* populations was a closely related lineage, indicating comparable ages of asexuality or very conserved genome evolution. Although the mechanism of transition to asexuality is unknown, hybridization is unlikely as it can not generate the shared haplotype differences among the European populations. Moreover, hybridization would entail substantial genomic changes and thus can

not explain the similar heterozygosity patterns for all populations (*18*). As remnants of *Wolbachia* endosymbiont can be detected in the *P. peltifer* genome, the transition to asexuality might have happened via reproductive manipulation of this endosymbiont (**fig. S8**). This type of transition often occurs via gamete duplication and results in fully homozygous lineages (*25*). Taken together, and given that haplotypic divergence of the Canadian and Japanese populations are considerably lower, transitions to asexuality in *P. peltifer* likely occurred via such a mechanism that substantially removes heterozygosity.

Following, we estimated the age of asexuality (**Fig. 2E**). Haplotypic divergence (i.e., heterozygosity) under asexuality over a substantial amount of time can be used to infer the age of asexuality. Heterozygosity gain over time involves the combined effects of accumulating differences between haplotypes via novel mutations (μ), and decreases via gene conversion events. Consequently, we first estimated the spontaneous mutation rate of *P. peltifer* by sequencing mothers and their eggs from the German population and measured $\mu = 2.05\times10^{-9}$ (**supplementary text**). Second, we estimated gene conversion to occur with a minimum track length of 500 bp in the German *P. peltifer* (**fig. S9**). Simulations of these combined effects show that contrary to common assumptions, heterozygosity reaches an equilibrium value that is independent of gene conversion track lengths over time (**Fig. 2E, fig. S10**). Using these biological parameters of *P. peltifer*, about 20×10^{6} generations are necessary to attain the mean 1.5 % heterozygosity equilibrium for *P. peltifer* (**Fig. 2E**). Given a generation time of one generation per year for *P. peltifer*, and assuming that the pre-asexuality heterozygosity for *P. peltifer*'s ancestor was substantially lower and that heterozygosity currently is at equilibrium, the European *P. peltifer* lineage has reproduced asexually since at least 20 million years. Moreover, these findings again suggest that the Japanese and Canadian populations are younger as they have not yet reached equilibrium.

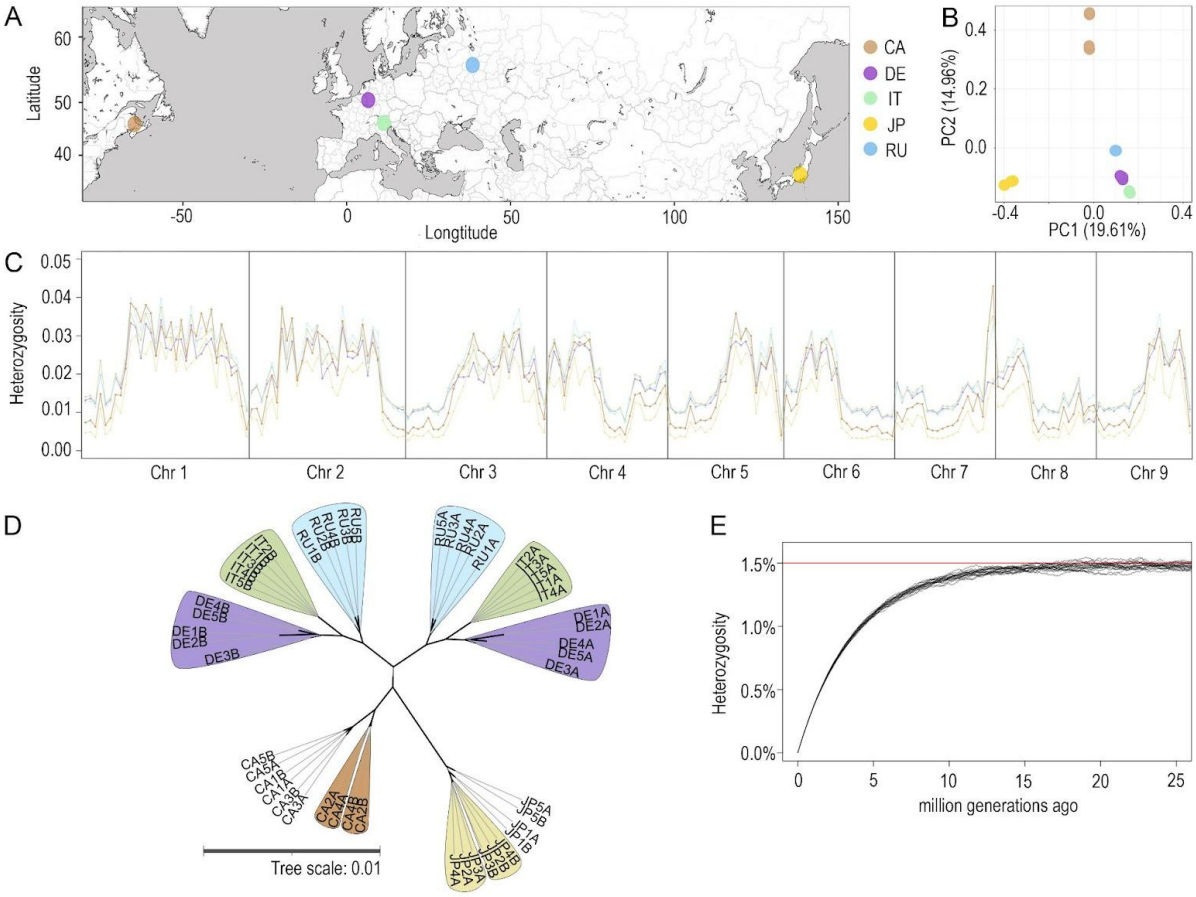


Fig. 2. Haplotype dynamics of worldwide *Platynothrus peltifer* populations suggest independent haplotype evolution for at least 20 my under asexuality of the DE-IT-RU lineages. (**A**) Sampling spots of the *P. peltifer* populations worldwide. Background gray shading indicates water bodies. Abbreviations: CA: Canada, DE: Germany, IT: Italy, JP: Japan, RU: Russia, color code kept throughout the figure. (**B**) Principal component analysis (PCA) of genomic population data visualizing 34.6 % of the total variability. (**C**) Mean heterozygosity between haplotypes within individuals for each population along the nine chromosomes, separated into 1 Mb genomic blocks. (**D**) Unrooted consensus haplotype tree showing the Meselson effect, i.e., complete separation of haplotypes A and B with mirror topologies, among *P. peltifer* populations from Germany, Russia and Italy. Bold branches indicate consensus supports > 96 %. (**E**) Simulations of increased divergence between haplotypes over time under gene conversion track length of 1000 bp, reaching a plateau after 20 million generations, using biological parameters of *P. peltifer*.

Adaptation and evolutionary innovation in the long-term absence of sex

Asexual organisms are typically regarded as evolutionary dead-ends because linkage of loci should result in decreased efficiency of purifying selection and reduced adaptive potential (*19*). Contrary to

evidence for these predictions in some younger asexual arthropods (*20*, *21*), asexual species of oribatid mites can maintain effective purifying selection (*3*). However, the genomic substrate for their unique ability to evolve and adapt in the long-term absence of sex remains unknown. Hence, we identified different evolutionary dynamics associated with the exceptional independent evolution of haplotypes that could contribute to adaptation and evolutionary novelty. This is exemplified in the haplotype-resolved genome of German *P. peltifer* (**Fig. 3**) encompassing allelic diversification and functionalization, the impact of horizontally transferred genes, and transposable element activity.

The decoupling and independent evolution of haplotypes after the transition to asexuality in *P. peltifer* (**Fig. 2**) mimics duplications of the whole genome with genes representing quasi-homeologs. Gene duplications are a major source of evolutionary innovation and allow access to novel phenotypes, as they can provide a novel function while preserving the old function of one duplicate and systematically increase mutational robustness (*22*, *23*). To test this prediction, we first analyzed differences in genetic variation between haplotypes that show a mean heterozygosity of 2.4 % (**Fig. 1C**). Among 11,029 allelic 1:1 orthologs, the vast majority differ from each other by non-synonymous (88.0 %) and/or synonymous (96.8 %) variants. Approximately 10 % gained or lost a stop codon and about 11 % showed frameshift, and nearly 15 % non-frameshift insertions and/or deletions (**fig. S11**). Non-synonymous to synonymous divergence (K_a/K_s) analysis suggests allele-specific direction of selection with 4.6 % (470/10,238) of alleles showing signs of strong positive selection ($K_a/K_s > 1$), potentially indicating rapid evolution (**Fig. 3A**). Over time, diverging alleles (similar to duplicated genes) might acquire novel functionalities and expression patterns. Hence, we also tested whether differentially expressed alleles (DEAs) contribute to functional diversification. About 9.1 % (936) of alleles are differentially expressed and show elevated K_a/K_s values compared to equivalently expressed alleles (EEAs) (**Fig. 3B, C**). Specifically, upstream and exonic regions of DEAs showed elevated variant densities compared to EEAs, unlike introns and 5' UTRs, suggesting functional and adaptive diversifications of alleles as well as associated transcription factors (**Fig. 3D**). Up- or down-regulation of DEAs is not haplotype-specific and can be different for each allele. These DEAs are enriched for basic cell functions, such as ribosomes, translation, and protein production (**fig. S12**).

Another mechanism providing novel traits to organisms is horizontal transfer of pre-existing genes (*24*). Horizontal gene transfers (HGTs) represent 2.0 % (504) of *P. peltifer* genes, which is within the typical range of asexual animals (**Fig. 1C**, **Fig. 3B**) (Jaron et al. 2021). These HGT stem mainly from bacteria, but also fungi, plants and metazoans. Of these HGTs, 92.9 % (468) are expressed and 61.5 % (319) contain intronic regions, indicating adjustments to functional integration in the host genome. From the 62 allelic 1:1 HGT copies that show signs of diversification, 25.2 % are differentially expressed, again suggesting selection of diverging HGT alleles (**Fig. 3B, E, F**). These HGTs might have arrived before the transition to asexuality, but divergent haplotypes likely contributed to novel (sub-) functions,

similarly to overall alleles. This is why we next identified HGTs that were incorporated after the transition to asexuality and before potential gene conversion events, i.e. HGTs that reside only on one haplotype (**fig. S13**). Of the 33 'orphan HGTs', 19 (57.6 %) are expressed and 16 (48.5 %) contain at least an intron, which is less compared to allelic HGTs, indeed indicating a more recent arrival with less time to adjust to the host background. Notably, HGT functional annotations, including differentially expressed and orphan HGTs, suggest the involvement in processes to digest plant cell walls (e.g. Glycosyl hydrolases). Further, HGT genes of the UGT family may contribute to pesticide resistance, indicating a contribution to the mite's ecology as soil-living decomposers (**fig. S14**, **S15, S16, supplementary data**). Overall, the molecular underpinnings driving diverging haplotypes under asexuality, specifically allelic differentiation of DEAs and HGTs, involve in addition to single nucleotide variants, (non-) frameshift insertions and deletions and stop gains/losses, most pronounced in HGT alleles (**Fig. 3F**, **fig. S11**).

Another possible agent of evolvability are transposable elements (TEs). They proliferate throughout genomes independently of the host cell cycle and their activity is often deleterious (*26*). While TEs likely evolve to be more benign in asexual genomes compared to sexuals (*4*, *26*), TE activity can occasionally be beneficial by accelerating evolution and by rewiring regulatory networks (*27*, *28*). We identified TEs and their haplotype-specific activity in the *P. peltifer* reference genome. Transposable elements comprise 27 % of the genome assembly and show signs of recent and past activity as emphasized by Kimura substitution levels (**fig. S17, table S4**). Moreover, TE density distributions suggest effective selection against TE insertion within genes, in concordance with previous results (**fig. S18**) (*4*). We detect noticeable differences in historical TE activity between the largest phased haplotypic blocks of chromosomes (**Fig. 3G**, **fig. S19**). Interestingly, pronounced differences occur from 6 %-12 % TE copy divergence, suggesting an increase in TE activity for one haplotype and a decrease in the other over 29 to 59 my ago. This correlates with the transition to asexuality of the *P. peltifer* lineage over 20 my ago. Very recent activity is largely restricted to one haplotype and, similar to haplotypic alleles, suggests divergence of one haplotype and more conservation of the other in TE activity and content.

Overall, our analyses suggest conservation of one allelic copy (or haplotype) and relaxed selection and/or functional adaptive divergence in the other, similar to innovation via gene duplication (*22*). This suggests that evolution of genes and regulatory regions via haplotypic divergence, modification of pre-existing genes (HGTs) and haplotype-specific activity of transposable elements can lead to increased evolvability via evolutionary innovation and robustness.

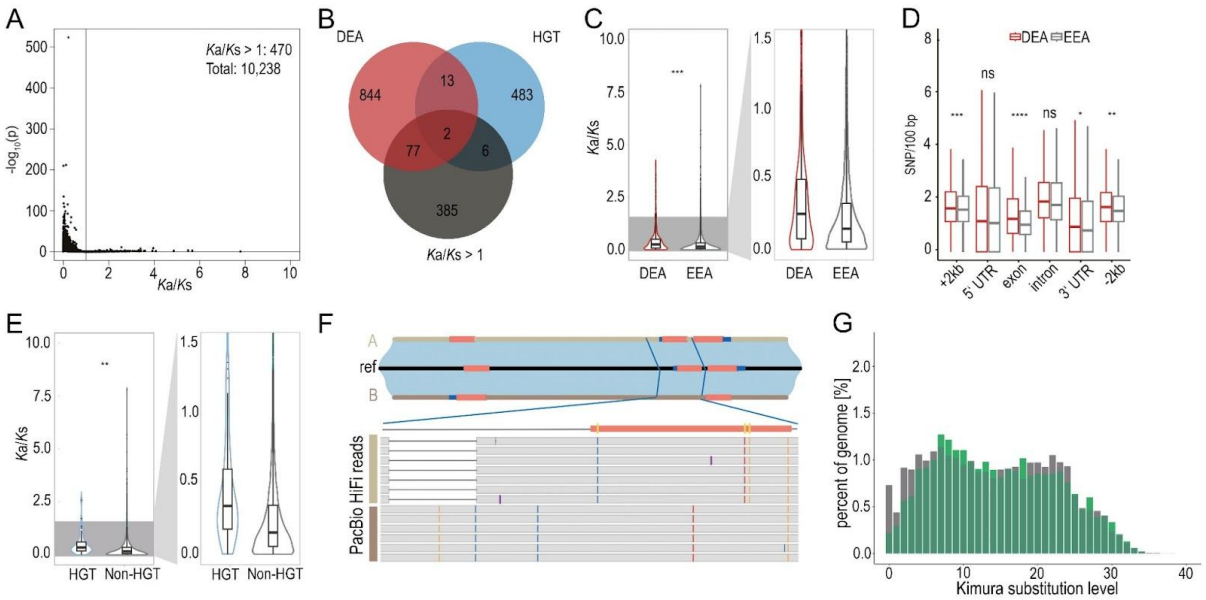


Fig. 3. Independent haplotypes contribute to evolvability. (**A**) Pairwise comparison of the K_a/K_s distribution for allelic genes. (**B**) Venn diagram showing numbers of differentially expressed alleles (DEA), horizontally transferred genes (HGT), and genes with $K_a/K_s > 1$. (**C**) Violin plots show levels of K_a/K_s of differentially expressed alleles and equivalently expressed alleles (EEA), highlighting $K_a/K_s < 1.5$. (**D**) Boxplots calculated for SNPs/100 bp for genic and neighboring regions, including the 5' and 3' untranslated regions (UTR), exons, introns, and up-/ and downstream regions (+2 kb and -2 kb respectively) for differentially (DEA) and equivalently (EEA) expressed alleles. (**E**) Violin plots show levels of K_a/K_s of HGTs and non-HGTs, highlighting $K_a/K_s < 1.5$. (**F**) Differentially expressed HGT (*Ppr_hap0_g20454*) as an example of the molecular changes contributing to divergent haplotypes. The two haplotypic blocks A and B differ in the number of exons (red bars). The highlighted region shows HiFi reads supporting allelic divergence (brown bars) by a deletion (white bar) and three non-synonymous substitutions (yellow stripe) in the second exon (red bar). Colored stripes in the reads indicate non-reference base substitutions. (**G**) Transposable element divergence landscape of the largest homologous haplotypic blocks of chromosome five (indicated by gray and green). Statistics were calculated using the Student's t-test between two groups (DEAs (n=879) and EEAs (n=9,359) and between HGTs (n=62) and non-HGTs (n=10,176)) (*: p-values < 0.05, **: < 0.01, ***: < 0.001).

Summary

In summary, we provide support for at least 20 million years of asexual evolution in multiple geographically separated populations of the oribatid mite *P. peltifer*. Several lines of evidence indicate that haplotypic independence provides the substrate for evolvability and evolutionary innovation. Thus far, the best evidence for the Meselson effect and ancient asexuality in parthenogenetic animals only

exists in oribatid mites (this study and also (*1*)), suggesting that the rarity of long-term asexuals might be explained by the requirement for haplotypic independence as a source for novelty to adapt and diversify in the absence of sex.

Acknowledgments:

Mark Maraun and Katja Wehner for electron microscopy picture of *Platynothrus peltifer*, Marcel D. Solbach for the *P. peltifer* drawing. Mélanie Jean for collection of mites in Canada. Nobuhiro Kaneko (Fukushima University) provided information about his previous study site as a habitat for *P. peltifer*.

This work was supported by the DFG Research Infrastructure West German Genome Center (407493903) as part of the Next Generation Sequencing Competence Network (project 423957469). Computational support of the Zentrum für Informations- und Medientechnologie, especially the HPC team (High Performance Computing) at the Heinrich-Heine University is acknowledged.

Funding:

German Research Foundation Emmy Noether grant BA 5800/4-1 (JB).

NSERC discovery grant RGPIN-2019-05744 (AMDC)

Author contributions:

Conceptualization: JB

Resources: HÖ, VL, AMDC, SS

Methodology: JB, HÖ, SG, NG, AB, YZ, KAB, VB

Investigation: HÖ, LO, CB, KB

Visualization: HÖ, SG, NG, AB, YZ, VB, LB

Funding acquisition: JB

Project administration: HÖ, JB

Supervision: JB, HÖ, NG, SG

Data curation: NG, SG

Formal Analyses: HÖ, SG, AB, NG, YZ, VB, LB, KAB

Writing – original draft: JB, HÖ, with input from all authors

Writing – review & editing: JB, HÖ

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

All code for the analyses can be found at https://github.com/TheBastLab/Ppr_evolution.

The genome, together with its annotation is available at NCBI and the reads at the SRA under Bioproject PRJNA1003031 (haplotypic blocks A: PRJNA1019947; haplotypic blocks B: PRJNA10199

Chapter 3

Ancient asexual speciation in the oribatid mite *Platynothrus peltifer*

Öztoprak, H.*, Guiglielmoni, N., Borgschulte L. and Bast, J.

in prep. (2023)

*Corresponding author

Abstract

In theory, long-term disadvantages of asexual reproduction are plentiful, deeming the absence of sex an evolutionary dead-end. Yet, certain oribatid mite species evade their detrimental fate and are known as 'ancient asexual scandals'. They not only persist but also diversify over evolutionary time. One prominent example is *Platynothrus peltifer*, which has been asexual for millions of years and diversified into various known evolutionary species. Here we provide haplotype-resolved reference genomes of five asexual *P. peltifer* populations from Germany, Russia, Italy, Japan, and Canada. We identify i) conserved genome synteny among the populations, ii) shared nucleotide diversity, iii) population-specific regions of genetic differentiation, and iv) distinct mitochondrial lineages of worldwide populations. Finally, we conclude that the Japanese *P. peltifer* population underwent speciation and highlight the affinity of cryptic species in this clade.

Introduction

Our understanding of how speciation proceeds has changed from one that emphasizes slow, steady processes controlled by different types of natural selection (Darwin, 1859) to a more contemporary perspective that acknowledges dynamic and potentially rapid mechanisms (Butlin et al., 2012). Key criteria for defining speciation revolve around reproductive isolation, genetic cohesion, common ancestry, and lineage distinguishability. Most research on speciation has primarily focused on the formation of reproductive barriers in sexual organisms. The concept of genetically and morphologically distinct clusters in asexual organisms however is not new, albeit somewhat controversial (Barraclough et al., 2003; De Queiroz, 2007). The emergence of species concepts involving asexual species thus protects the biodiversity of these units. As such the evolutionary species concept states that a lineage maintaining its identity from other organisms, with its own evolutionary tendencies and historical fate is a species, thus including asexuals (Simpson, 1951; Wiley, 1981). However, the processes of asexual

speciation remain poorly understood. Establishing the genomic basis of asexual population dynamics will unravel the vast biodiversity that is currently overseen. Evolutionary consequences of asexuality suggest that asexual organisms struggle to adapt to rapidly changing environments, thus occupying the tips of phylogenies only. However, parthenogenesis is more common than previously recognized, with many parthenogens occurring in larger geographic regions than their related sexual species (van der Kooi et al., 2017). Remarkably, certain eukaryotic taxa have maintained asexual reproduction for millions of years (Brandt et al., 2021; Öztoprak et al., 2023). They persisted in the absence of sex despite the consequences of asexuality and are known as "ancient asexual scandals" (Judson & Normark, 1996).

The Taxonomy of clonal organisms has been a subject of significant discourse, emphasizing the need to distinguish between various types of asexual organisms, their origins, histories, and modes of reproduction (Martens et al., 1998; Simpson, 1961). Moreover, population genetic theory has demonstrated that asexual organisms can potentially undergo speciation, either through diversifying selection or prolonged geographical isolation (Barraclough et al., 2003). Hence, various species concepts have emerged over time. Each strives to delimit concrete entities of organisms, which live in a dynamic and ever-changing natural world.

Concepts such as the Evolutionary Species Concepts (EvSC; Hey, 2006; Simpson, 1961) were among the first to include asexual organisms, as asexuals encompass (i) independent evolving lineages, (ii) monophyly in taxonomy, and (iii) phenotypic reflection of genotypes (Barraclough, 2010). The Evolutionary Genetic Species Concept (EgSC) provides a framework to assign asexual organisms into species based on the degree of genetic differentiation, using the K/θ method (commonly referred to as 4X-Rule; Birky et al., 2010; Birky & Barraclough, 2009).

An extensively studied system characterized by vast morphological diversity and distinct asexual clades are the bdelloid rotifers (Fontaneto et al., 2007). They are species under the EvSC, coherent with genetic clusters that evolve independently under geographic isolation. Another notable group exhibiting asexual speciation, under the EvSC, is found in oribatid mites. A study of the oribatid mite *Platynothrus peltifer* identified seven distinct clades within this species, supported by genetic differences exceeding 2% among lineages and averaging 56%. Evidence suggests that they were physically isolated by continental drift and mountain uplift (Heethoff et al., 2007). These clades met the criteria of species according to the 4X rule (Schön et al., 2009), although the extent to which they differ morphologically and ecologically remains uncertain. Recently, an in-depth study of worldwide populations of *P. peltifer* provided evidence for its long-term asexuality and showed that its haplotypes evolved independently for at least 20 million years (Öztoprak et al., 2023). These findings raise questions about how haplotypic independence influences diversification and speciation.

In this study, we aimed to investigate the global diversity of *P. peltifer* and to identify the underlying mechanisms of asexual speciation on a genomic scale. Therefore, we collected samples from Germany, Italy, Russia, Japan, and Canada and assembled four *de novo* genomes of *P. peltifer* individuals and one asexual outgroup, *Nothrus palustris*. With previously published German *P. peltifer* and *Hermannia gibba* assemblies, we performed a comparative analysis of the genomes, exploring population structure and dynamics. Our results characterized the patterns of nuclear and mitochondrial divergence and conserved synteny. In addition, we presented the first genome-wide quantification of genomic variation in multiple populations of this ancient asexual species. Understanding the processes of divergence and speciation in asexual organisms, such as *P. peltifer*, holds the potential to advance speciation research, which has traditionally been founded on sexual reproduction.

Material & Methods

Sample preparation, sequencing and read quality

Details of sample collection and sample preparation are given in Chapter 2. In short: five different populations of *Platynothrus peltifer* were collected from Germany, Russia, Italy, Japan and Canada. The German population of *P. peltifer* and its sexual outgroup *Hermannia gibba* were sampled from the same spot in Dahlem (North Rhine-Westphalia). Additionally, an asexual outgroup to *P. peltifer*, *Nothrus palustris* was acquired from Göttingen (Lower Saxony, Germany).

For details on sequencing see Chapters 2 and 4. In short: for the *de novo* assembly of each population of *P. peltife*r, *H. gibba* and *N. palustris* one reference individual was long-read sequenced with PacBio HiFi (SMRTbell® Libraries from Ultra-Low DNA Input) and linked-read sequenced with TELL-seq™ whole-genome-sequencing (WGS) library. To generate population data four individuals from each population were sequenced with TELL-seq™ only.

To assess read quality 27-mers in the HiFi reads were analyzed using KAT v2.4.2 (Mapleson et al., 2017) with the modules kat hist and kat gcp (default parameters). Ploidy was further investigated using kmc v3.2.1 with parameters '-k 27 -ci 1 -cs 10000' and Smudgeplot (v0.2.5) (Rhyker Ranallo-Benavidez et al., 2020) with default parameters. To obtain a frequency count of *k*-mers Jellyfish v2.2.8 was used as input to estimate genome heterozygosity and repeat content with GenomeScope (http://qb.cshl.edu/genomescope/).

Assembly of *N. palustris* and additional *P. peltifer* populations

Assemblies. Assemblies of *N. palustris* and the Russian, Italian, Japanese and Canadian *P. peltifer* populations were conducted following this pipeline: HiFi reads were assembled using hifiasm v0.16.1-r375 (Cheng et al., 2021) with default parameters into initial collapsed haploid contigs. Flye v2.9 (Kolmogorov et al., 2019) was used with default parameters to phase contigs. To purge artificial duplications minimap2 v2.24 (Li, 2021) with parameter '-x map-hifi' and purge_dups v1.2.5 (Guan et al., 2020) were used. Short-read data from TELL-Seq was used to scaffold reads with scaff10x and ragtag. Gaps were filled with TGS-Gapcloser v.1.1.1. Finally, the scaffolds were polished after mapping Hifi reads using minimap 2 v2.24 with parameter '-x map-hifi' and HyPo v.1.0.3.

Assembly evaluation. To assess ortholog completeness Benchmarking Universal Single-Copy Orthologs (BUSCO v.5.0.0; Simão et al., 2015) was used against the Arthropoda odb10 lineage (1,066 orthologs) and the Arachnida odb10 lineage databases (2,934 orthologs). To evaluate *k*-mer completeness KAT v2.4.2 was used with the module kat comp with default parameters.

Genome annotation. For details on genome annotation see Chapter 4. In short: A repeat library including transposable elements (TEs) was built. The hardmasked assemblies were converted to softmasked assemblies. Details on RNA-seq mapping of *N. palustris* followed the German *P. peltifer* pipeline. Functional annotation followed the Funannotate v.1.8.13 pipeline, using trimmed RNA-seq reads from each species. RNA-seq data from German *P. peltifer* was used for the other *P. peltifer* populations.

Genome dynamics. Gene synteny of protein sequences of the collapsed haploid assembly and the haplotypic blocks were aligned using BLAST v2.6.0 with parameters '-evalue 1e-10 -outfmt 6'. In order to find blocks of gene synteny, MCScanX (commit 97e74f4) was run between the collapsed haploid assembly (Hap0), of each population with default parameters. Collinearity was visualized with SynVisio (Bandi et al., 2022) with a minimum match score of 2812.74 (CAvsJP), 2953.26 (ITvsRU), 3577.26 (RUvsJP) and 3873.95 (DEvsIT). 2,653,942 matches imported and 36,971 pairwise comparisons among all populations.

Mitochondrial genome assembly. To generate the mitochondrial genome trimmed TELL-seq reads were first assembled using MitoFinder v1.4.1 (Allio et al., 2020) with parameter '-MEGAHIT' (Li et al., 2015) for invertebrate mitochondria. To check for heteroplasmy a second assembly with NOVOPlasty v4.3.1 was conducted (Dierckxsens et al., 2017; Dierckxsens et al., 2020). The mitochondrial genome of *Steganacarus magnus* was used as seed.

Mitochondrial genome annotation. To identify protein coding, tRNA and rRNA genes, the mitochondrial genomes were annotated using MiTFi v0.1 and additionally using MITOS, MITOS2 (both available at http://mitos2.bioinf.uni-leipzig.de/index.py; Bernt et al., 2013; Donath et al., 2019), ARWEN v1.2.3 and tRNAscan-SE v2.0 (Chan et al., 2021). tRNAscan-SE was run with COVE cutoff '-20'. Geneious (Biomatters Ltd.) was used to manually curate annotations and find consensus predictions. tRNAs were selected based on their predicted secondary structure and their minimum free energy, computed using the RNAfold web server (http://rna.tbi.univie.ac.at/).

Population genetic analyses

Single nucleotide polymorphisms (SNPs) were used to investigate population dynamics within populations of *P. peltifer* from Germany, Russia, Italy, Canada, and Japan. Phased population data were generated by mapping trimmed raw reads generated by TELL-read v1.0.3 to the collapsed haploid reference genome.

Mapping. TELL-seq reads of each individual were mapped against the soft masked collapsed haploid assembly (Hap0) of its population using *bwa mem* v0.7.15 with default parameters. The resulting alignment was sorted using samtools v1.11 and duplications were removed using Picard MarkDuplicates (v2.26.2 Broad Institute). Coverage was calculated with samtools flagstat with default parameters.

Variant calling. Variants were called using the Genome Analysis ToolKit (GATK v.4.1.9.0) pipeline. GVCFs were produced using HaplotypeCaller and then merged using CombineGVCFs. Variants were detected with GenotypeGVCFs. SNPs were selected with SelectVariants. SNPs were filtered with VariantFilteration following hard-filtering criteria of GATK in variant quality 'QD < 2.0, QUAL < 30.0, strand bias estimated by the Symmetric Odds Ratio test SOR > 3.0, fisher strand filter FS > 60.0, mapping quality MQ < 40.0, MQRankSum < -12.5' and distance of ALT reads from the end of the reads 'ReadPosRankSum < -8.0'. Multiallelic SNPs and indels were removed.
After filtering, our final data set involved 25 samples with 15,105,721 sites for *P. peltifer*, and five samples with 7,479,671 sites for *H.gibba. Nothrus palustris* was excluded from this analysis.

Nucleotide diversity. To identify whether asexual populations have less diversity than sexual populations and further characterize genetic distances among the populations of *P. peltifer*, we calculated the nucleotide diversity of each population. Therefore we obtained per-site nucleotide diversity with the '--site-pi' function in VCFtools (Danecek et al., 2011) and calculated their mean and

standard deviation. The distribution of nucleotide diversity was calculated in 1 Mb windows with the '--windowed' function.

Genetic differentiation. Genetic differentiation was assessed through pairwise F_{st} comparisons among the populations. Pairwise F_{st} values were calculated for each pair with the '--weir-fst-pop' function of VCFtools (Danecek et al., 2011) in a sliding window with 50-kb windows in 10-kb steps. To detect selection signatures regions in the 5 % right tail of the empirical distribution of F_{st} values were considered to be candidate regions under selection.

Tajima's D. To shed light on whether DNA is evolving non-random processes, we calculated Tajima's D in non-overlapping sliding windows of 50-kb size using the '--TajimaD' function of VCFtools (Danecek et al., 2011).

Phylogenetic Analyses of cytochrome oxidase I

A *cytochrome oxidase I* (*cox1*) sequence was amplified from each isolate for details see Chapter 2. For the phylogenetic analyses, a reference database containing 65 *cox1* sequences of *P. peltifer* was obtained from NCBI GenBank database. These sequences add similar and more distinct populations of *P. peltifer* (last date of accession: October 2023). Additionally, *cox1* sequences from *Heminothrus* spp. were obtained from NCBI and isolated from *Heminothrus* spp. individuals occurring at the same sample spot in Japan. Sequences of the generated population data of *Hermannia gibba* and *Nothrus palustris* were added as outgroups. All *cox1* sequences obtained for this study were manually checked and corrected for sequencing errors using Chromas v.2.6.5 and aligned in SeaView v.4.6 (Gouy et al., 2010). They were aligned in mafft using the local-pair algorithm (Katoh & Standley, 2013). The phylogenetic tree was calculated in Genious (Biomatters Ltd. Geneious Prime 2023.1.2: Biomatters Ltd.). Following previous studies on oribatid mite mitochondria based on the Akaike informative criterion, the Neighbor-Joining clustering method with the genetic-distance model HKY was used. Confidence values were added as percentages obtained through 1000 replicates bootstraps to the best scoring tree.

Results & Discussion

De novo Platynothrus peltifer genomes of worldwide populations compared to chromosome-scale genomes of sexual outgroup *Hermannia gibba* and asexual outgroup *Nothrus palustris*

As the basis for analyzing genomic patterns of differentiation, we *de novo* assembled four draft genomes of worldwide populations of *Platynothrus peltifer* (based on criteria including geographic distribution and mitochondrial divergence) and their asexual outgroup *Nothrus palustris*. Adding the chromosome-level reference genome of *P. peltifer* (Öztoprak et al., 2023), and of its sexual outgroup *Hermannia gibba* (Wulsch et al., 2023), allows for comparative analysis.

For these assemblies, a single individual from each population was sequenced using PacBio HiFi and TELLseq™ (Illumina). The resulting assemblies of *P. peltifer* varied in size. The combined genome size ranged from a minimum of 195 Mb in the Japanese assembly to a maximum of 292 Mb in the Canadian assembly (Table 1). Generally, the genome sizes of all populations were largely similar. Except for the Canadian *P. peltifer*, which exhibited a genome size excess of more than 50%. However, this is most likely due to the lower quality of the genome assembly and the resulting poor annotation (Table 1). The assembled genome size of *Nothrus palustris* comprised 255 Mb. The scaffold N_{50} and N_{90} of the assemblies were 0.99 Mb and 0.15 Mb for the Canadian, 9.3 Mb and 2.1 Mb for the Russian, 11.8 Mb and 5.3 Mb for the Japanese and 18.8 Mb and 9.4 Mb for the Italian *P. peltifer*. For *N. palustris* N_{50} and N_{90} were 6.9 Mb and 2.5 Mb respectively. Genome assembly completeness was high with BUSCO completeness scores in the Arachnida dataset ranging from 96.6 % to 96.8 % (Table 1) for all genomes. Genome annotation yielded total gene predictions per genome, ranging from 19,722 to 28,499 (mean= 24,326.8, standard deviation= 2,252.64) for *P. peltifer*. BUSCO analysis of the annotated genes showed less completeness for all populations, with a mean of 90.2 ± 1.72 % in the Arachnida dataset (Table 1). Finally, repeats (including transposable elements and unknown repeats) comprise 7.5 % in Japanese, 7.7 % in Russian, and 26.9 % in Italian assembly. Because high continuity of the assembly is essential for TE identification, Canadian repeats were not annotated, as identification is expected to yield even fewer TEs compared to Japan and Russia. Thus, variations in genome size are associated with the relative TE content of TEs, explaining the majority of assembly size differences across *P. peltifer*.

Table 1: Genome and annotation statistics of collapsed haploid assembly of *Platynothrus peltifer* (Ppr) populations, sexual *Hermannia gibba* (Hga) and asexual *Nothrus palustris* (Nps). arachnida_odb10 including n:2934 and arthropoda_obd10 n:1013 orthologs. C: complete BUSCSOs, S: Complete and single-copy BUSCOs, D: Complete and duplicated BUSCOs, F: Fragmented BUSCOs, M: Missing BUSCOs.

	Canadian Ppr	Japanese Ppr	Italian Ppr	Russian Ppr	German Ppr	Hga	Nps
Dataset	PacBio HiFi, TELLseq	PacBio HiFi, TELLseq	PacBio HiFi, TELLseq	PacBio HiFi, TELLseq	PacBio HiFi, TELLseq Omni-C	PacBio HiFi, Arima Hi-C	PacBio HiFi, TELLseq
Size of assembled scaffolds [Mb]	292	195	208	219	219	352	255
Number of scaffolds	1023	111	49	70	9	8	139
Largest scaffold [Mb]	4.2	18	26	14.4	32.6	75.9	22.9
Scaffold N50 [Mb]	0.99	11.8	18.8	9.3	23.4	43.9	6.9
Scaffold N90 [Mb]	0.15	5.3	9.4	2.1	19.5	30.9	2.5
BUSCO completeness [%] (arthropoda odb10)	C:95.8% [S:88.3%, D:7.5%], F:1.3%, M:2.9%	C:95.2% [S:93.6%, D:1.6%], F:1.7%, M:3.1%	C:95.8% [S:93.8%, D:2.0%], F:1.3%, M:2.9%	C:95.7% [S:93.5%, D:2.2%], F:1.4%, M:2.9%	C:96.0% [S:93.9%, D:2.1%], F:1.0%, M:3.0%	C:94.3% [S:91.4%, D:2.9%], F:2.1%, M:3.6%	C:95.4% [S:92.8%, D:2.6%], F:1.4%, M:3.2%
BUSCO completeness [%] (arachnida odb10)	C:96.7% [S:87.9%, D:8.8%], F:1.0%, M:2.3%	C:96.6% [S:93.9%, D:2.7%], F:1.0%, M:2.4%	C:96.8% [S:94.4%, D:2.4%], F:1.0%, M:2.2%	C:96.8% [S:94.4%, D:2.4%], F:1.0%, M:2.2%	C:96.7% [S:93.7%, D:3.0%], F:1.0%, M:2.3%	C:96.1% [S:92.7%, D:3.4%], F:0.9%, M:3.0%	C:96.8% [S:93.2%, D:3.6%], F:1.1%, M:2.1%
Repeats [%]	-	7.5	26.9	7.7	31.9	48.5	6.4
GC content [%]	28.59	28.54	28.56	28.67	28.69	-	-
Number of genes	28449	23300	19722	25254	24909	-	-

Synteny among the assemblies of the five populations of *Platynothrus peltifer*

Taking together the four newly assembled lineages of *P. peltifer* from this study and the one previously published, the generated data exhibit a range from various populations in the northern hemisphere. Owing to the old age of *P. peltifer* populations and their geographical isolation, genome-wide structural rearrangements are expected. No large-scale rearrangements were prominent, indicating conserved gene order among the populations (Figure 1). All populations shared 78.56% collinear genes, with a match score of 50. In chromosome (Chr) 6 of the German Assembly, a region of approx. 11 Mb is relocated in Chr 6 of the Italian assembly. Following this region, we see that it is partly relocated to Chr 7 in the Russian assembly. Considering the previous in-depth analysis of the German assembly, this region is poor in coding sequences (Öztoprak et al., 2023). Thus, we conclude that this translocation likely did not affect many genes. Further translocations were observed between Russian and Japanese assemblies (Figure 1)

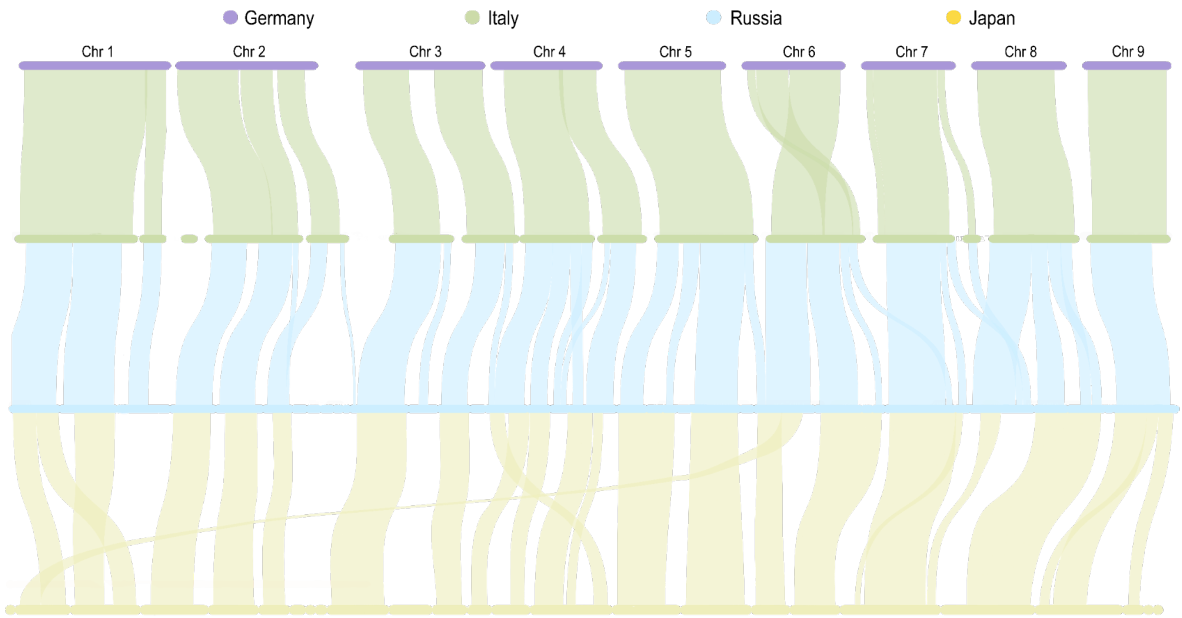


Figure 1: Syntenic regions of collinear genes among *Platynothrus peltifer* populations show only a few large-scale rearrangements. Note: Canadian assembly was excluded due to low continuity. Showing distinctly separated scaffolds for Russian and Japanese assemblies is difficult due to the high scaffold number.

Genetic diversity and comparative population analysis

As the assembly quality varied, *k*-mer analysis of PacBio HiFi reads was conducted to compare heterozygosity differences independent from assembly quality. Japan and Canada exhibited higher homozygous content than other populations, indicating less heterozygosity (suppl. Figure 1). This is consistent with findings demonstrating the Meselson effect in the German, Russian, and Italian populations, but not in the Japanese and Canadian populations of *P. peltifer* (Öztoprak et al., 2023).

To allow the genetic comparison of population statistical metrics among the five geographically distinct *P. peltifer* genomes (DE, RU, IT, JP, and CA), aligning the genome sites to the chromosome-scale assembly of the German *P. peltifer* was required. A set of five individuals per population was used for population diversity and differentiation analyses. Nucleotide diversity (pi) suggests comparably high diversity among global *P. peltifer* populations and their sexual outgroup *Hermannia gibba* (suppl. Figure 2). Moreover, while fluctuations in nucleotide diversity along the chromosomes were shared, the overall amount of diversity differed among the *P. peltifer* populations (Figure 2). Compared to the other populations, diversity in the Japanese and Italian populations was reduced, indicating more homogenous populations in Japan and Italy. Furthermore, comparable trends were observed between the observed regions of lower diversity and the previously observed regions of lower heterozygosity (Öztoprak et al., 2023). However, regions of remarkably high and narrow peaks of nucleotide diversity with neighboring lower values, which would indicate recombination, were absent.

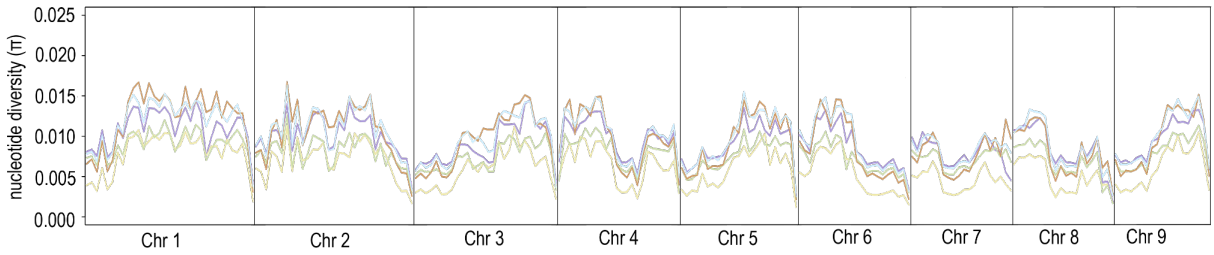


Figure 2: Mean nucleotide diversity (π) between haplotypes within individuals for each population of *Platynothrus peltifer* along the nine chromosomes, separated into 1 Mb genomic blocks. Color code for the populations: purple=Germany, blue=Russia, green= Italy, yellow=Japan, brown=Canada.

However, it is worth noting that the calculated diversity is underestimated, as all sites were used for calculation. This suggests that the actual diversity is even higher than our estimates. This intriguing similarity in diversity patterns between sexual and asexual species raises questions about how *P. peltifer* manages to maintain such high genetic diversity despite its asexual reproduction mode.

Regions of differentiation are unevenly distributed and population-specific

To identify specific regions contributing to population differentiation, we measured the fixation index (F_{st}). The F_{st}-values among the populations were variable, including values from 0.16±0.06 between the German and Russian populations, up to 0.42±0.08 between the Japanese and Italian populations (Figure 3). This aligns with differences in heterozygosity between the European and Pacific clades (Öztoprak et al., 2023). Pairwise comparisons among all populations showed genomic regions with high and low differentiation unevenly distributed along the chromosomes. For each population, region above the 5% tail of the empirical distribution are distributed along the whole genome. This might be due to the linkage of genes in asexuals. Still, there are most prominent population-specific regions of differentiation at the end of Chr 1 for the Japanese, Chr 3 for the Canadian, two consecutive regions in Chr 7 for the German, Chr 8 for the Russian and Chr 9 for the Italian population. Some pairwise comparisons did not exhibit this differentiation or were observed at a different location: Canada vs. Germany in Chr 3, Russia vs. Japan in Chr 8, and Russia vs Italy in Chr 9. The regions of differentiation may be mainly attributed to the differences in mutation rate heterogeneity among the populations. Moreover, differences in the frequency of gene conversion events may also contribute to lineage-specific differentiation. Genes within these regions will be identified in the future.

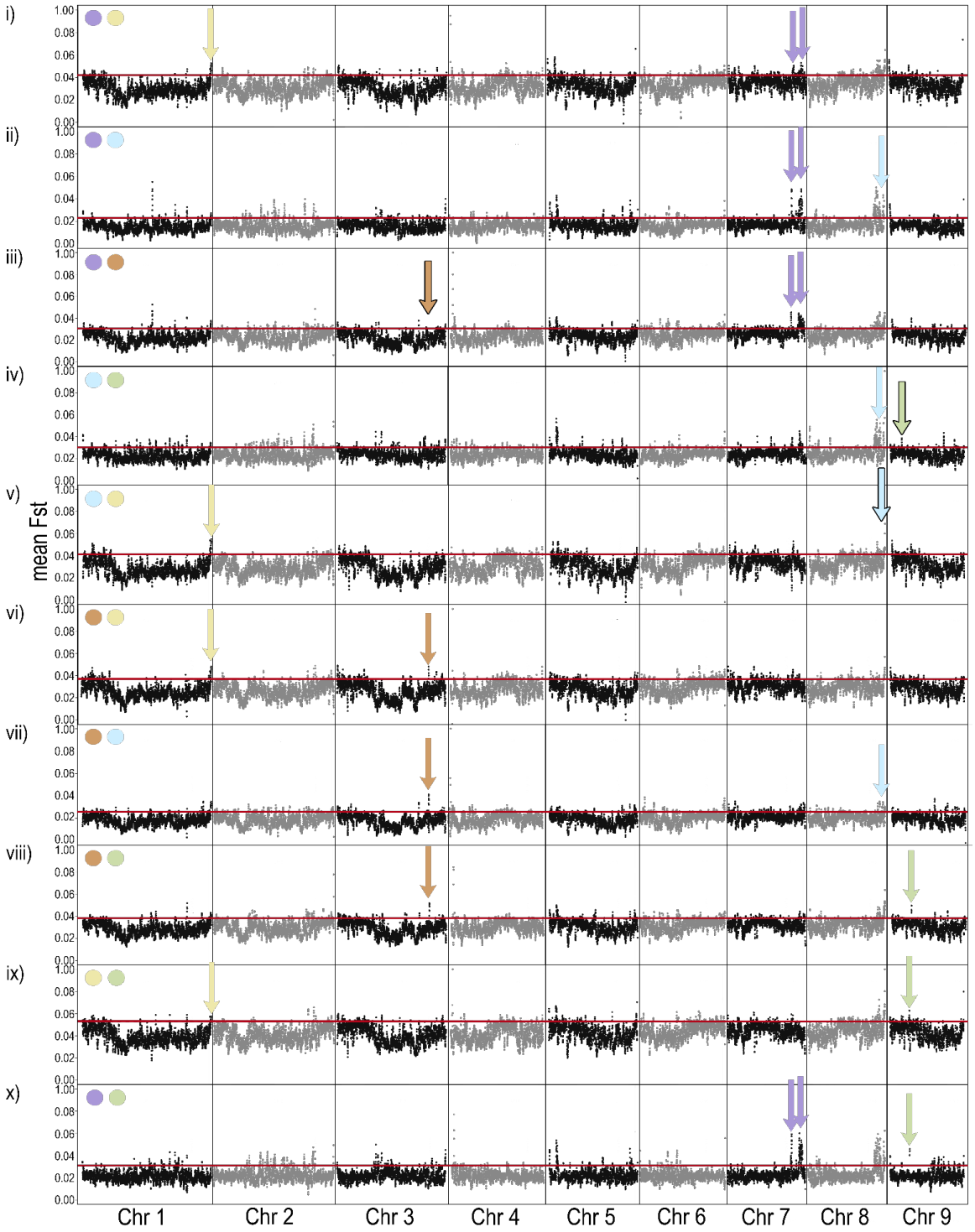


Figure 3: Distribution of population differentiation (F_{st} values) from pairwise comparisons of *Platynothrus peltifer* populations. The region above the red line are in the 5% tail of the empirical distribution: i) F_{st}(DE vs. RU)=0.157±0.057; ii) F_{st}(DE vs. JP)=0.319±0.072; iii) F_{st}(DE vs. CA)=0.232±0.059; iv) F_{st}(DE vs. IT)=0.222±0.058; v) F_{st}(IT vs. RU)=0.234±0.057; vi) F_{st}(JP vs. RU)=0.301±0.081; vii) F_{st}(JP vs. CA)=0.270±0.071; viii) F_{st}(CA vs. RU)=0.188±0.056; ix) F_{st}(CA vs. IT)=0.301±0.064; x) F_{st}(JP vs. IT)=0.418±0.077. Arrows indicate population-specific high differentiation. Arrows with grey borders indicate low differentiation in regions with high differentiation in pairwise comparisons with all other populations.

Mitochondrial genomes

Mitochondrial (mt) genes exhibit higher mutation rates than nuclear genes, hence the mito-nuclear mismatch increases over evolutionary time in asexuals (Havird et al., 2015). We therefore additionally assembled mitochondrial (mt) genomes from short linked reads. This allowed next to genomic diversification and differentiation, the analysis of recent evolutionary changes in the form of mitochondrial divergence. While mt genome assembly size was consistent within populations, it varied among populations between 14,891 bp (German) to 14,930 bp (Japan). Their variation in size but also the intrapopulation variation is consistent with previous assemblies of *P. peltifer* and other oribatid mites (Domes-Wehner, 2009; Li & Xue, 2019). For *N. palustris* assembly size was consistently 14,546 bp and for *H. gibba* 14,340 bp (suppl. Table 1). Annotation revealed 13 protein-coding genes in all species. The gene order among all the species was generally conserved (Figure 4, suppl. Table 2). From the expected 22 tRNAs (Zhan et al., 2021) 21 were annotated in the German *P. peltifer* (Ppr DE1; Öztoprak et al. 2023). Further, some tRNAs were systematically missing in all individuals from each species: trnaA (Alanine) for *P. peltifer and N. palustris.* However, previous studies identified tRNA for Alanine in both species (Li & Xue, 2019; Zhan et al., 2021). This also applies to the not retrieved trnR (Arginine) in *H. gibba*.

To compare the nucleotide diversity of the mt genomes in the populations, we examined the percentage of identical sites. For the single populations of *N. palustris* and *H. gibba*, we found 99.1 % and 99 % identical sites, respectively, among the mt genomes in their populations (suppl. Figure 3). For *P. peltifer*, we identified differences for each population, with 70.4 % of the sites shared among all populations (suppl. Figure 3). The highest intraindividual variation in mt genomes was observed among Canadian, German, and Russian individuals with only approximately 84 % shared sites. Here, we found two distinct mitochondrial lineages, where the individuals in the lineage shared all sites, but between the lineages, we found divergence. This phenomenon is frequent among oribatid mites, with multiple parthenogenetic mites in various habitats showing significant intraspecific genetic divergence (von Saltzwedel et al., 2014).

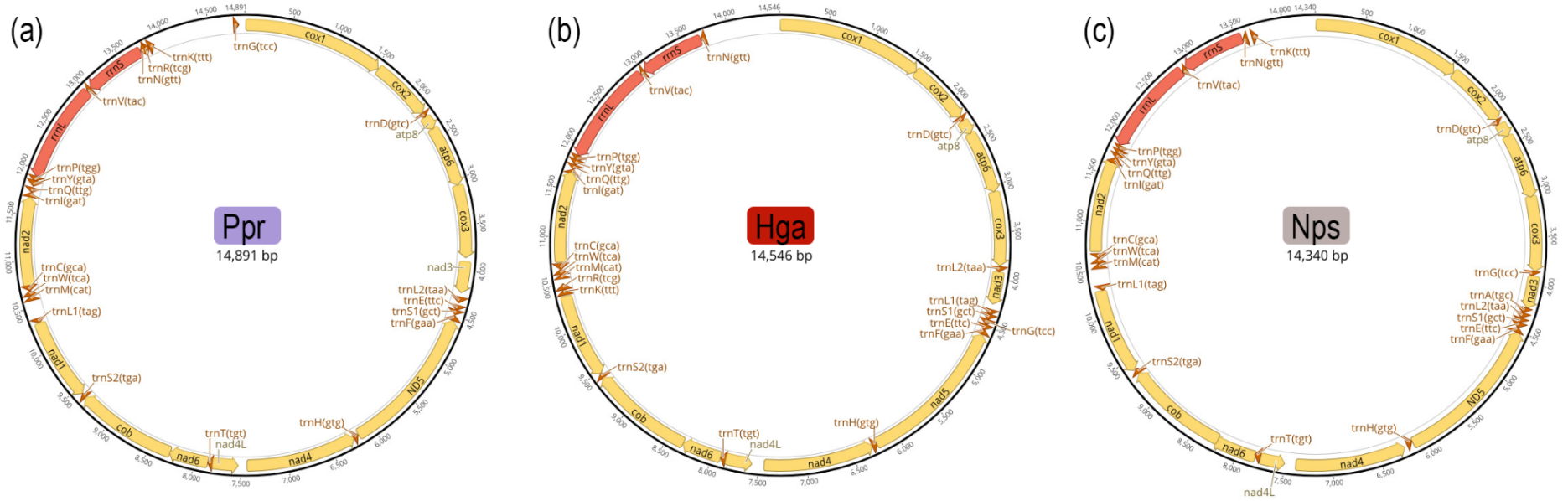


Figure 4: Schematic representation of mitochondrial genomes of German *Platynothus peltifer, Hermannia gibba* and *Nothrus palustris*. Genes are highlighted in yellow, tRNAs in brown and rRNAs in red. The size of the circular contig is shown in the center.

Mitochondrial phylogeny

Mitochondrial genes like cytochrome oxidase I (*cox1*) are employed in phylogenetic analyses to elucidate speciation (Armstrong & Ball, 2005; Hajibabaei et al., 2007; Hebert et al., 2003). To illustrate interindividual divergence among worldwide *P. peltifer, cytochrome cox1* sequences were obtained. We were able to generate nearly full-length *cox1* sequences for 23 of the 25 individuals. The sequences ranged from 468 to 660 nucleotides (suppl. Table 1). A combined analysis of these sequences and sequences from a previous study of global *P. peltifer* from the NCBI database were subjected to phylogenetic analyses (Heethoff et al., 2007) to allow for more data and thus a taxonomically broader approach. As oribatid mite species are known to exhibit cryptic species (Lienhard & Krisper, 2021; Pfingstl et al., 2019; Schäffer et al., 2019), we also included sequences of closely related *Heminothrus* species to allow for more genetic references. The analysis revealed, next to known geographical *cox1* lineages of *P. peltifer* (Domes-Wehner, 2009; Heethoff et al., 2007; Rosenberger, 2022), novel clades in the family Nothridae (Figure 5). *Hermannia gibba* and *Nothrus palustris* individuals grouped monophyletic and basal to all other sequences with full support. All sequences of *P. peltifer* and *Heminothrus* spp. each formed a distinct monophylum. The *P. peltifer* clade exhibited seven main clades with full support, roughly following their geographical origin. A basal clade with two sequences from Canada was followed by a fully supported clade with mainly sequences from Northern Tyrolia, Italy, exhibiting longer branch lengths, hence more diversity. However, sequences obtained from this study from Russia and one from Germany were also grouped in this clade. Their exact location is unclear as bootstrap support within this monophylum is low. All Japanese sequences built a monophyletic clade with full bootstrap support. Paraphyletic to the Japanese phylum was a clade separated into the Northern/Central Europe clade, including newly obtained sequences from Germany but also one Russian sequence from a previous study. Nevertheless, these Russian sequences were grouped with one sequence from Germany, and both were separated from all the other sequences with full support. Next to the Northern/Central Europe clade, we found two clades with full support, one inducing the remaining three Canadian sequences and one with Northern Tyrolia sequences from a previous study intertwined with two of the newly obtained Russian sequences. Although these seven subclades exhibited full support, their basal branching ordination remains unclear, indicating unresolved diversity in the family Nothridae.

Paraphyletic to the *Platynothrus peltifer* clade was the *Heminothrus* clade. This clade included *Heminothrus* targionii, *Heminothrus longisetosus*, as well as two distinct but unknown *Heminothrus* species from NCBI with full support. It also showed the newly obtained *Heminothrus* specimen from this study grouping basally in a fully supported monophylum. As the ‘American’ *P. peltifer* sequences, which were obtained from a previous study, grouped within the *Heminothrus* clade, we suggest misidentification of these sequences.

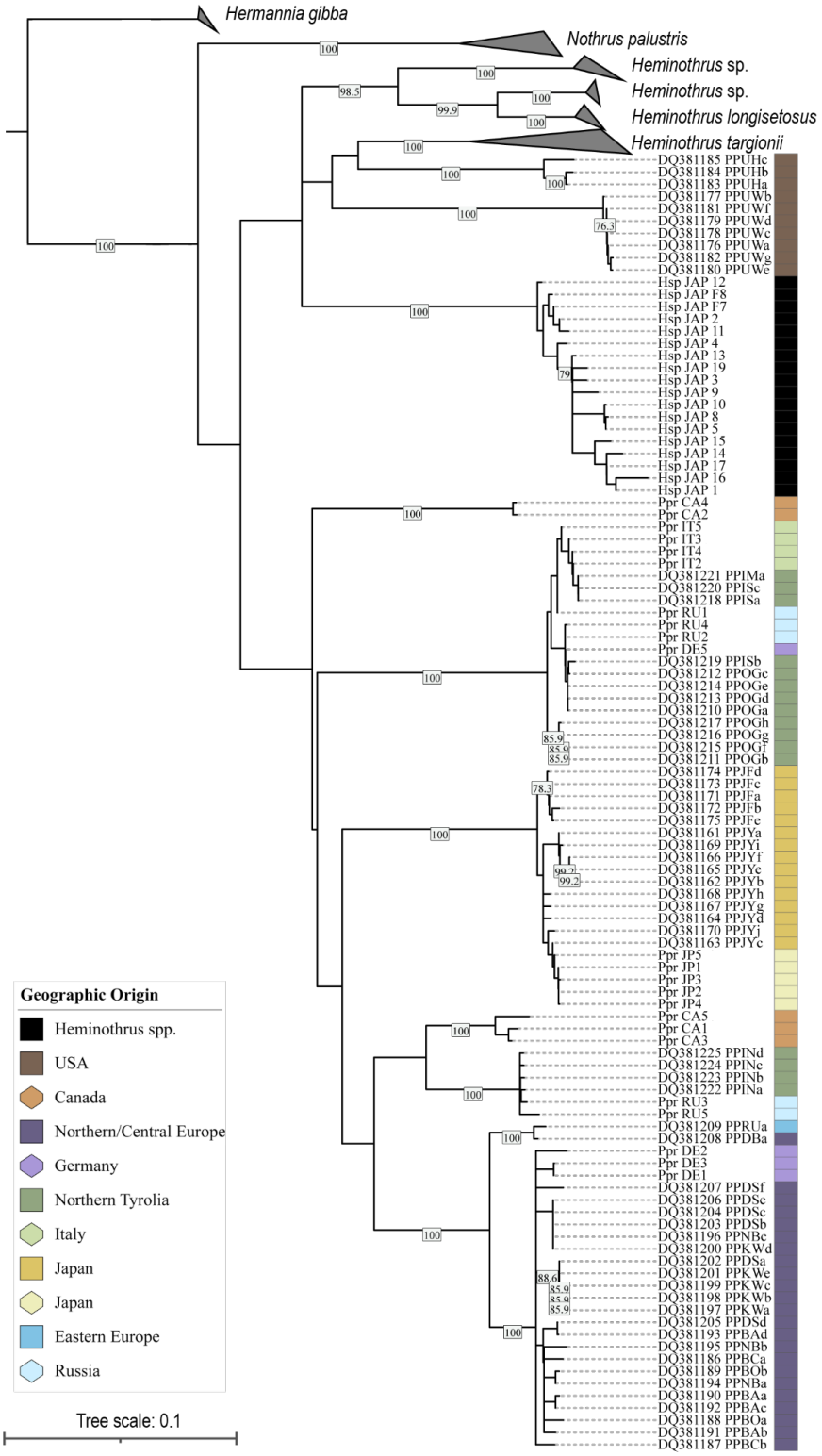


Figure 5: *Cytochrome oxidase I* phylogeny of *Platynothrus peltifer* with sequences of *Heminothrus* spp. and *Hermannia gibba* and *Nothrus palustris* as outgroups. Shown is the Neighbor-Joining tree including 153 sequences. Support values are given in bootstraps ≥75 % from 1000 replicates. Color indicates geographical origin, hexagons indicate newly obtained sequences and squares indicate sequences from NCBI. *Heminothrus* spp. sequences obtained from NCBI and outgroups are collapsed.

Employing species criteria

To investigate whether speciation criteria of evolutionary genetic (Eg) species are met in *P. peltifer* lineages, the 4X rule is conducted on the clades identified through the phylogenetic analysis. The rule states that nucleotide diversity among the clades should be four times greater than the maximum diversity within the clade (Birky et al., 2005). The highest within-clade diversity is 2.8 % observed in the Japanese clade. This is 14 times greater than within the whole *P. peltifer* clades overall (suppl. Table 3). Yet, all seven clades within *P. peltifer* exhibit this phenomenon, which indicates all clades as species following the EgSC.

To check whether these differences result in coding variation we translated the alignment for *cox1* and *cob*, revealing a pairwise identity of 99.9 % and 96.3 % respectively for all *P. peltifer* (suppl. Table 4). Only one amino acid in the Japanese mt genomes differed in *cox1* from the other populations. However, for *cob1* translational differences of 3.7 % among individuals were observed.

Species delimitation in ancient asexual *Platynothrus peltifer*

Genome synteny and whole genome nucleotide diversity are intended to be re-evaluated once the data has undergone necessary improvements. For now, we therefore, focus on general population differentiation and heterozygosity patterns, basing them on mitochondrial phylogeny to discuss possible species delimitation in *P. peltifer*. The use of mitochondrial phylogeny as our basis allows us to discuss already available data on worldwide populations and expand the dataset.

A distinct separation of the Japanese population becomes evident. Their lower heterozygosity compared to the European clades (Öztoprak et al., 2023) and distinct *cox1* phylogeny indicate independent species boundaries. Moreover, morphological differences between the Japanese specimen compared to the European specimen were identified (suppl. Figure 4). Teratology, i.e., morphological abnormalities are generally observed in oribatids (Arroyo & Iturrondobeitia, 2007). Still, these differences were consistently exhibited within the Japanese population (personal communication, Dr. Fernando Villagomez, October 2023). Taking this into account we suggest the Japanese population be an independent subspecies of *P. peltifer.* The separation of Japan from the mainland was 25-15 Ma (Maruyama et al., 1997), perpetuating the effects of geographical isolation and allopatric speciation on this lineage.

The Canadian sequences did not form a monophyletic group, although the basal bootstrap support, in general, is low. Their lower heterozygosity but comparably higher nucleotide diversity to the European clade might suggest a cryptic lineage. However, this remains inconclusive at this point. Improving the Canadian genome will allow for improved nuclear comparisons to the other populations and within the population. The European sequences lack clear separation. Two distinct clades include Italian and Russian sequences. Despite their shared phylogeny, Italy and Russia have been geographically isolated

by the uplift of the Alps at least since 44 Ma (van Hinsbergen et al., 2020). This period exceeds the separation of Japan from the mainland, indicating that 'enough' time should have passed to differentiate. However, other factors like mutation rate heterogeneity and environmental factors could also influence these differences. Indeed, pairwise comparisons do show distinct regions of differentiation along the genomes of these two populations. Among all, the Italian population lacks a morphological record. Considering, that the Italian population had no mitochondrial diversity within the population and comparably lower nuclear diversity than the other populations, speciation might be ongoing. The Morphological analysis becomes indispensable to uncover the relationship between the Italian population to the other European populations. How this mechanism facilitates this effect remains to be investigated in the future. Another possibility stems from the idea that high genetic diversity in parthenogenetic species suggests the possibility of multiple colonization by genetically diverse clones. The wind-mediated dispersal of Russian, German, and Italian clones may have contributed to their phylogenetic polyphyly. This could lead to re-colonization of different lineages in the same habitat. Considering that most of the European-clades variation stems from haplotypic divergence, we suggest a mechanism working on a global (population) scale facilitating cohesion among geographical lineages.

Conclusion and comments

Asexual speciation is a longstanding enigma in evolutionary biology, challenging conventional notions of sexual reproduction and species formation. This study highlights the relevance of applying genetic and phylogenetic tools to organisms with 'unconventional' modes of reproduction, shedding light on various levels of diversity in nuclear and mitochondrial genomes.

Although this study has provided valuable insights into the genomics of asexual speciation in *Platynothrus peltifer*, it is important to note that this study is an ongoing endeavor. As the genomes are continuously improved, there will be a need to add, revisit and refine certain genomic analyses. Additionally, future investigations are planned to explore further aspects of asexual speciation in this intriguing species, such as e.g. the McDonald Kreitman test to compare species variation to the divergence between species or a genome-wide association study (GWAS) to identify genetic markers associated with particular phenotypes. Studies on *P. peltifer*'s ecological and morphological data are also ongoing (Heethoff Lab, University of Darmstadt, Dr. Fernando Villagomz, personal communication, October 2023). We advise integrating whole genome, mitochondrial and morphological results to delimit new species. We anticipate that these upcoming analyses will contribute to a deeper and more comprehensive understanding of the complex processes underlying asexual speciation.

Chapter 4

A female heterogametic ZW sex-determination system in Acariformes

Wulsch S.[*], **Öztoprak H.**, Guiglielmoni N., Jeffries D.L.[†], Bast J.[†]

bioRxiv (2023)

*Corresponding author

† These authors contributed equally to this work

Abstract

Sexual reproduction, while often associated with separate sexes, is an ancient and widespread feature of multicellular eukaryotes. While a diversity of sex-determination mechanisms exist, for many organisms, which of these mechanisms is used remains unknown. Exploring sex-determination mechanisms in Acariformes, among the oldest chelicerate clades, is intriguing due to its potential to unveil conserved sex-determination systems. This insight can have implications for understanding sex chromosome evolution and its broader impact on higher taxa. To identify the mechanism of sex determination in Acari, i.e., oribatid mites, we generated a high-quality chromosome-level genome assembly of *Hermannia gibba* (Koch, 1839) by combining PacBio HiFi and Hi-C sequencing. Coverage and allele-frequency analyses on pools of male and female individuals suggest a female-heterogametic ZW sex-determination system with little degeneration of the W chromosome. To date, this represents the only documented case of a ZW system in Acariformes. Further comparative studies in H. *gibba* will reveal how old the ZW system is and whether it exhibits conservation or polymorphism.

Introduction

The vast majority of metazoans reproduce sexually whereby two separate sexes mix their genomes via meiosis and fusion of gametes to generate offspring (Bell 1982). Despite this being such a widely shared and fundamental feature, the mechanisms that determine whether individuals develop into males or

females are highly diverse (Bachtrog et al. 2014; Beukeboom and Perrin 2014). Two (not mutually exclusive) main types of sex determination occur in nature: environmental sex determination (ESD), where environmental cues trigger sex development; and genotypic sex determination (GSD), where sex-specific genotypes determine the individual's sex, such as sex chromosomes. Sex chromosome systems have evolved independently numerous times throughout eukaryotes. The spectrum ranges from stable sex chromosomes, e.g. in birds or mammals (Bull 1984; Zhou et al. 2014) to rapidly evolving sex chromosome pairs with frequent turnovers, e.g. in frogs and fish (Takehana et al. 2014; Jeffries et al. 2018). Why the stability of sex chromosomes and the prevalence of transitions differ so much across the tree of life remains one of the unsolved mysteries in the field of reproductive systems evolution. To answer this question, it is imperative to characterize the diversity of sex chromosome systems across a diversity of organisms (Bachtrog et al. 2014; Consortium et al. 2014; Palmer et al. 2019).

Acariformes are one of the major groups of Acari including over 11,500 species with many more undescribed (Subías 2023). They likely originated during the Precambrian and were among the first to colonize land (Schaefer et al. 2010; Lozano-Fernandez et al. 2019). Despite their age and diversity, relatively little is known of their sex-determination system. Generally, there is considerable debate and contrasting reports about sex determination in Acariformes based on karyotypes (Heethoff et al. 2006). The ancestral state in Acariformes is believed to be diplodiploidy, i.e. the same karyotype for females and males without distinct sex chromosomes (Sokolov 1954; Norton et al. 1993; Wrensch et al. 1994). Most previous studies on the major Acari groups (Parasitiformes and Acarifomres) investigated sex determination based on indirect evidence or karyotypes: in Parasitiformes 76 cases with XO or XY sex chromosomes and 106 with some form of haplodiploid, whereas in Acariformes 17 cases of XO or XY are reported and 106 with a type of haplodiploidy (Heethoff et al. 2006; Consortium et al. 2014).

Oribatid mites (Acariformes, Acari), lack differentiated sex chromosomes and are diploid (mostly 2n=18) (Norton et al. 1993). Due to conflicting evidence hinting at the potential existence of haplodiploidy or parahaploidy, the exact sex-determination mechanism present in oribatid mites remains elusive (Oliver 1983; Norton et al. 1993; Wrensch et al. 1994). The sex ratios of some sexual oribatid mites are balanced (1:1) or they tend to show a female bias (Grandjean 1941; Webb and Gw 1979; Smelansky 2006; Wehner et al. 2014). Moreover, the presence of diverse sex-determination systems within the group cannot be ruled out, given its large size, diversity, and the paucity of prior investigations into sex-determination mechanisms. In most species where balanced sex ratios are observed, sex is genetically determined through the mechanism of GSD and often involves sex chromosomes (Uller et al. 2007). Some oribatid mite species show an equal 1:1 sex ratio, whereas for most species female-biased (72%-69%) sex ratios are observed (Domes et al. 2007). Interestingly, a large number of asexual

species exist in the Oribatida that reproduce in the absence of sex over a substantial amount of time (Brandt et al. 2021; Wehner et al. 2021; Öztoprak et al. 2023).

Hermannia gibba (Hermanniidae, Sarcoptiformes), a widely distributed sexual species found in diverse forest soils, is presumed to employ genetic sex determination due to the unstable nature of its soil habitat with various microhabitats (Smrž 2010; Liana and Witaliński 2012; Lóšková et al. 2013). However, a comprehensive understanding of its sex-determination mechanisms remains elusive. In this study, we investigate the sex-determination system of the oribatid mite *H. gibba*. Prerequisites for the identification of the sex-determining region (SDR) are high-quality genome and sequence data of pools from male and female individuals.

Thus, we first generated a chromosome-level genome assembly from a female *H. gibba* individual up to current standards (Guiglielmoni et. al 2022). Second, we conducted whole-genome sequencing of pools of female and male individuals. We used two complementary approaches to identify regions of sex linkage (Palmer et al. 2019): scans of i) coverage differences and ii) allele frequencies along the genome of male and female mite pools. Further for additional verification, we performed a restriction enzyme digestion of a predicted female-specific site and confirmed the adequacy between morphological and molecular identifications (Methods and Figure S4). Our results identify a ZW female-heterogametic sex-determination system with a sex-linked region of 14.54 Mb on Chromosome 1 in *H. gibba*. To date, this is the first documented case of a ZW system in Acari.

Results

Genome assembly and annotation

The *de novo* genome assembly of *Hermannia gibba* was performed using a combination of Pacific Biosciences (PacBio) HiFi reads from a single female individual, and Hi-C reads from a population pool, as described in **Figure 1a** (see Methods for details). PacBio HiFi reads were first analyzed with *k*-mer spectra, which showed that sequencing data had little contamination and the genome was predicted as diploid (Supplementary Figures S1, S2). The initial assembly resulted in 934 contigs with a total size of 537.1 Mb, which was decreased to 415 contigs and a total size of 371.1 Mb after haplotig purging. Subsequent Hi-C scaffolding yielded eight chromosome-level scaffolds. The final assembly has a size of 352.2 Mb, 61 gaps, and with few duplicated regions (**Figure 1b**). The Benchmarking Universal Single-Copy Orthologs (BUSCO) tool was used to assess completeness, resulting in a score of 96.1 % against the Arachnida odb10 database (92.7 % single-copy orthologs, 3.4 % duplicated orthologs). The eight chromosome candidates, ranging from 29.5 to 75.9 Mb in size, can be identified in the Hi-C contact map (**Figure 1c**). 55.92 % of the assembly was identified as repetitive content, including 29,072 telomeric repeats of the sequence 'AACCT'. Genome annotation yielded 25,869 predicted protein-coding genes with a BUSCO score of 95.5 % against the Arachnida odb10 database (91.5 % single-copy orthologs, 4.0 % duplicated orthologs).

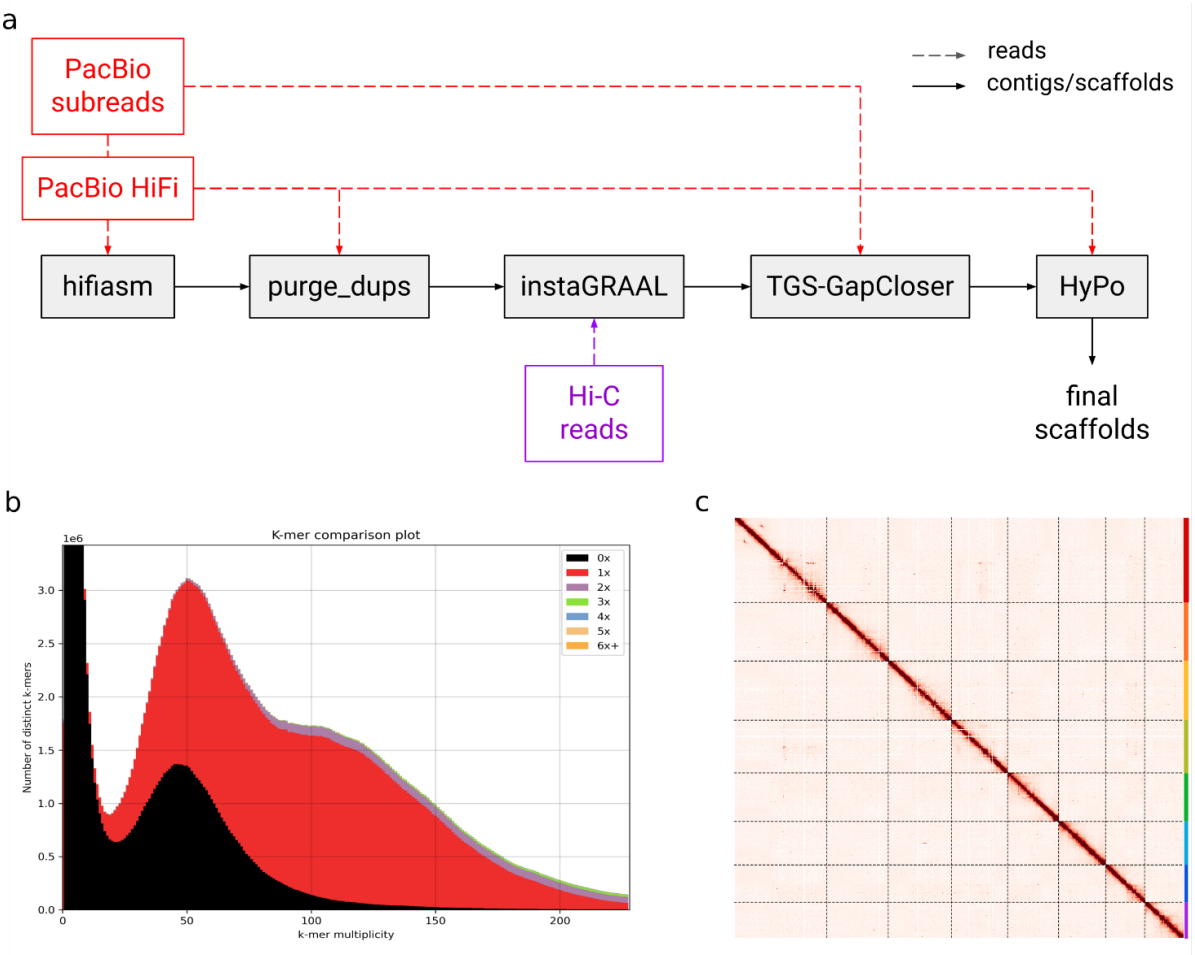


Figure 1 - Genome assembly of female *Hermannia gibba*. a) Assembly workflow. b) *k*-mer analysis of the final assembly. Low-multiplicity *k*-mers are absent from the assembly (0X). The heterozygous peak and the homozygous peak are respectively identified around 50X and 110X. As haplotypes are collapsed in the assembly, homozygous *k*-mers are represented once and only half of heterozygous *k*-mers are included in the assembly. There are limited artifactual duplications (2X in the homozygous peak). c) Hi-C contact map of the eight chromosome candidates.

Searching for degenerated sex-linked regions using read coverage

To identify the SDR, we searched for genomic differences between females and males by sequencing pools of 30 individuals of each sex from *H. gibba* (for details see methods, Figure S1).

Sex chromosomes often lose recombination between gametologs (i.e. X/Y or Z/W) in the region surrounding the sex determination locus. These non-recombining regions are expected to accumulate sequence divergence over time, as well as deleterious mutations, and repetitive sequences. Given enough time, the sex-limited chromosome (Y or W) eventually loses large numbers of genes (Charlesworth et al. 2005; Jeffries et al. 2018), resulting in regions of the X or Z that are haploid in the heterogametic sex. To search for such highly diverged and/or degenerated sex-linked regions, we aligned the pooled sequencing data against the female repeat-masked reference genome, yielding comparable mapping

rates of the female pool (96.76 %) and the male pool (98.11 %). We then compared coverage between male and female pools in 1 kb windows along the genome. If degenerated sex-linked regions exist, we would expect to find a male/female coverage ratio $\Delta = 0.5$ for a XY system, or $\Delta = 2$ for a ZW system. While autosomal, pseudoautosomal, or undegenerated sex-linked regions are expected to have a male/female coverage ratio of $\Delta = 1$. The ratio of male-to-female coverage along the genome of *H. gibba* did not show any differences, indicating either a lack of sex chromosomes or the absence of a degenerated sex-linked region (**Figure 2a**).

Identification of sex-linked regions using allele frequency calculation

Nucleotide sequences in SDRs can diverge more rapidly between the gametologs than large changes in chromosome structure. Thus, in sex-linked regions that recently lost recombination and have not yet degenerated to the point of exhibiting coverage differences, differences in allele frequencies (AF) between males and females can be more informative. We found 4,663,024 polymorphic sites between the two pools which we used to search for loci with SNPs specific to the sex-limited chromosome. Such loci are expected to possess a major allele frequency of 1 in the homogametic sex, and 0.5 in the heterogametic sex. We therefore filtered for loci where one of the sexes is fixed (with a major allele frequency of >0.95) and the other to be half (~0.5; **Figure 2b,c,d)**. As we did not know the system of heterogamety before performing this analysis, we did this twice, once to search for an XY system by filtering for a major allele frequency of >0.95 in females and ~0.5 in males, and once to test for a ZW system with a major allele frequency of >0.95 in males, and ~0.5 in females.

Only in the female pool, we observed loci that exhibited allele frequencies ranging from 0.51 to 0.69, calculated based on the local maxima and minima. Among these 7,649 discovered SNPs, 7,029 (92 %) were located on Chromosome 1, forming a contiguous region of 14.54 Mb from the start of 6.3 Mb to the end of 20.8 Mb (**Figure 2e**). As these loci are homozygous in males and heterozygous in females, we inferred the presence of a ZW female heterogametic genotypic sex-determination system. To test this hypothesis, we searched for the presence of a sex-specific locus in the sex-linked region that is exclusively present in females and absent in males. We performed restriction enzyme digestion, specifically targeting the female-specific locus located within a HindIII restriction site (Figure S4).

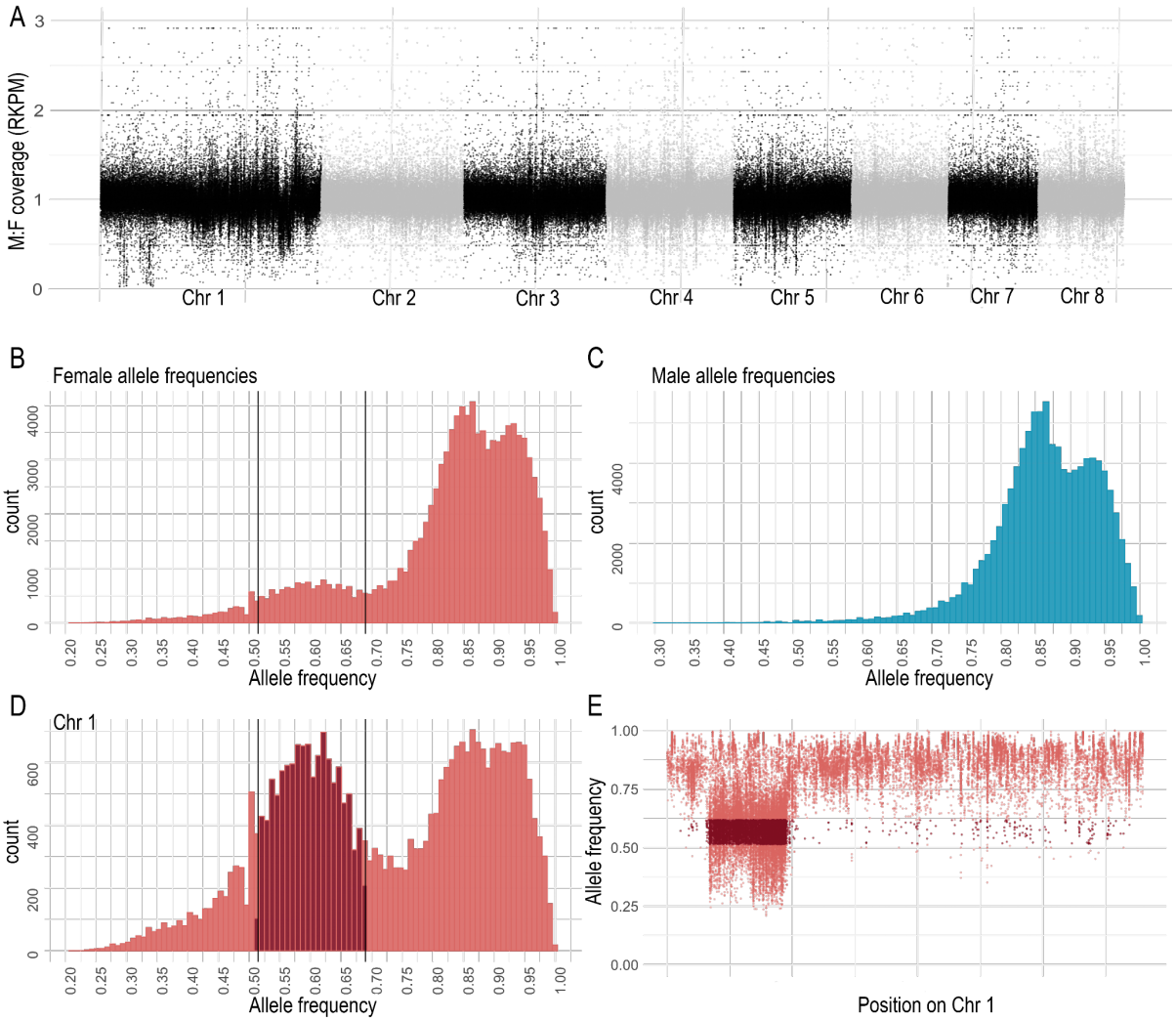


Figure 2 - The sex determination region is not highly degenerated in *Hermannia gibba*. RPKM normalized coverage differences along the eight chromosomes in 1 kb blocks do not reveal any significant differences in coverage between male and female genomes (A). The histograms (B) and (C) provide an overview of the allele frequency distributions within each sex across the genome. The x-axis represents the allele frequency, and the y-axis represents the number of alleles corresponding to each frequency. Only alleles with an allele frequency (AF) greater than 0.95 in the other sex are included. The chromosome-specific histogram (E) shows the allele frequencies on Chromosome 1, with allele frequencies falling within the calculated threshold of 0.51 to 0.69 displayed in dark red. Scatter plot (E) visualizes the position of allele frequencies falling within the calculated threshold range on Chromosome 1. These allele frequencies indicate a sex-linked region spanning a continuous region of 14.54 Mb.

Discussion

Historically, sex determination in oribatid mites relied mainly on cytological approaches, which led to contradicting results (Heethoff et al. 2006). Our study represents the first time a sequencing-based approach has been used to identify the sex-determination system within the Acariformes group, while also uncovering the first documented case of a ZW system in Acariformes. Our investigation revealed

a lack of extensive degeneration of the *Hermannia gibba* sex chromosome. This may reflect a relatively recent loss of recombination in the sex-linked region. However, the age of the sex chromosome remains unknown as there is a large variation in the rate at which sex chromosomes lose recombination, and in the rate at which they degenerate once recombination has ceased (Wright et al. 2016).

The relatively large size (75.9 Mb) of the identified sex chromosome compared to autosomes might indicate chromosomal rearrangements. Oribatid mites possess holocentric chromosomes, which, unlike monocentric chromosomes, offer unique advantages that may impact recombination patterns (Wrensch et al 1994, Izraylevich 1995). Holocentric chromosomes are recognized for preserving chromosomal regions that might otherwise degrade or disappear due to recombination suppression (Lachange, 1967). This preservation is attributed to their ability to retain kinetic activity along the whole chromosomes thus withstanding fragmentation. Notably, it was demonstrated that both in butterflies (which also exhibit a ZW sex-determination system) and in nematodes (which share holocentric chromosomes), inverted meiosis can drive the evolution of new karyotypes (Lukhtanov et al. 2018). Oribatid mites are suggested to exhibit inverted meiosis (Wrensch et al. 1994), thus if these rearrangements happened recently, the sex chromosomes in *H. gibba* might not have had enough time to degenerate.

In our study, the presence of distinct hard edges in the allele frequency plot suggests a singular event, most likely an inversion. Considering the significant role of inversions in contributing to recombination suppression and the formation of sex chromosomes, they align with two possible models i) the sex antagonistic selection model and ii) the deleterious mutation load model.
Sex antagonistic selection is known to be the main driver of recombination suppression in sex-chromosome evolution (Wright et al. 2016). However, a recently proposed model suggests that this might be an overrepresentation (Ironside 2010; Jay et al. 2022). Besides sex antagonistic selection, alternative hypotheses deal with the effect of meiotic drive and genetic drift (Svedberg et al. 2018). However, oribatid mites are known to exhibit large and stable population sizes (Maraun and Scheu 2000) and testing for meiotic drive is difficult (Ponnikas et al. 2018). Nonetheless, insights from comparative genomic analyses in fungi challenge the notion that inversions primarily cause halted recombination on sex chromosomes (Bergero and Charlesworth 2009). Instead, these analyses suggest that recombination suppression was the ancestral state, with inversions emerging as a consequential, rather than primary, factor in recombination suppression. Many forces drive recombination suppression, and identifying a single event is difficult since interactions of different circumstances can result in a loss of recombination forming heterozygous sex chromosomes (Charlesworth et al. 2005; Bergero and Charlesworth 2009; Ponnikas et al. 2018).

In summary, our analysis indicated the absence of highly differentiated regions through coverage analysis. Allele frequencies identified distinct patterns indicative of chromosomal rearrangements, possibly inversions. We hypothesized a relatively recent loss of recombination. Nevertheless, how and why the sex chromosomes might not have differentiated is a complex process influenced by various genetic and structural factors. To obtain a more profound understanding of sex chromosome evolution in mites, it is imperative to expand our investigations to include additional mite species for comprehensive data collection. Considering Acariformes' position as a basal group within Chelicerata, where diverse mechanisms are observed, our research contributes to enhancing clarity on sex-determination system evolution in basal groups.

Materials & Methods

Since there is no apparent external dimorphism in these mites, sex was determined by examination of the genitalia as previously described in the literature (Palmer and Norton 1991). The reference genome was from a single individual, and reads were generated from sex-specific pools to perform coverage analysis and allele frequency calculation to determine genomic differences between the sexes.

Sampling

Animals were collected from ground litter in a coniferous forest in Dahlem, Germany (50.389204 N, 6.568780 E) in April 2021. Specimens were isolated using heat gradient extraction (Kempson, Lloyd, and Ghelardi 1963). *Hermannia gibba* was morphologically identified after Weigmann (Weigmann 2006) and genetically confirmed by *cytochrome oxidase I* (CO1) sequencing.

Sample preparation

To minimize contamination from gut contents, mites were starved for at least one week before DNA extraction. Mites were cleansed with a brush in sterile water, in a solution of detergent/water (1:20) (fit GmbH, Zittau, Germany), 70 % ethanol (EtOH), and 0.05 % bleach (DonKlorix; CP GABA GmbH, Hamburg, Germany), and then rinsed with sterile water. To prevent DNA degradation, mites were stored at -80°C prior to sexing.

Identification of individuals sexes

Following standard practice for morphological sexing in oribatid mites, the presence of an ovipositor is used for identification of females; spermatophore and missing ovipositor were used to identify male specimens (Figure S3). To perform morphological sexing, the mites were first quickly frozen and then

placed onto a microscope slide in 5 µl cooled TNES buffer (400 mmol NaCl, 50 mmol Tris pH 8, 0.5% SDS, 20 mmol EDTA). To expose the ovipositor, genital plates were removed with a dissecting needle, and specimens were dissected completely. To ensure no misidentification and therefore bias to pool data, only clearly identified specimens were used. Sexed individuals were transferred to 190 µl of TNES and frozen at -80°C until *g*DNA extraction. In total 155 specimens were examined with 72 females (~46 %) and 83 males (~54 %). 30 individuals were pooled to an equimolar amount and sequenced at 3-4X coverage per individual (see Supplementary Table 1). We obtained a total of 38 Gigabases (Gb) of raw data for each pool, with an average of $10.19*10^6$ reads for the female pool and $9.97*10^6$ reads for the male pool.

gDNA extraction (for extra low input)

We followed the high-molecular-weight *g*DNA extraction protocol for oribatid mites with chitinase digestion (Öztoprak & Bast 2023). In short: chitinase was added to each individual and incubated at 37°C for 60 min. Then Proteinase K was added followed by t yeast tRNA (Invitrogen) and 5 M NaCl and 96 % ethanol. The solution was incubated at -20°C for 1h or overnight. DNA purification was conducted by washing the pellet twice with fresh and chilled 70 % ethanol and eluted in TE buffer. Samples were left to homogenize at 4°C overnight. RNase digestion was conducted and DNA concentration was measured using Qubit Fluorometer v. 4 with the Qubit dsDNA HS Assay kit (ThermoFisher Scientific, Waltham, MA). For pool-sequencing DNA of n = 30 individuals of each sex are pooled to an equimolar amount.

Long-read sequencing

Single-individual high-molecular-weight DNA was sequenced at Genomics & Transcriptomics Lab (GTL) in Düsseldorf with an ultra-low input protocol for long-read sequencing. Sequencing yielded in total 45.8 Gb of Pacific Biosciences Circular Consensus HiFi reads (Wenger et al. 2019) with an N_{50} of 15 kilobases (kb).

Hi-C sequencing

Thirty whole mites were crosslinked in 3 % formaldehyde for 1 hour at room temperature with low agitation, followed by quenching in 250 mM glycine for 30 minutes at room temperature with low agitation. The Hi-C library was prepared using the Arima Hi-C+ kit (Arima Genomics); after the addition of lysis buffer, the sample was frozen in liquid nitrogen to crush the mites with a pestle, and the protocol was strictly followed thereafter. The library was sequenced on an Illumina NovaSeq 6000

at the Cologne Center for Genomics (CCG; Cologne, Germany), which generated 88 million pairs of 150-bp reads.

RNA sequencing

A Direct-zolTM RNA MicroPrep kit (Zymo Research, Irvine, CA) was used to extract total RNA from 10 adults using TRIzol reagent that had been DNase I-treated. With the use of TruSeq Stranded Total RNA with Ribo-Zero Globin, the CCG built an RNA-seq library and sequenced $35x10^6$ pairs of 100-bp reads.

Pool sequencing

We used n = 30 individuals from each sex for pool sequencing. To meet sample requirements of 3-5-fold coverage for Illumina NovaSeq 6000 Genomic DNA (PCR free -350bp) paired-end DNA sequencing, concentrations were normalized for each individual (Table S1). Genome sequencing and associated library preparation were performed by Novogene Europe (Cambridge, United Kingdom).

Preliminary analyses of HiFi reads

k-mer spectra of the PacBio HiFi reads were built using KAT v2.4.2 (Mapleson et al. 2016) with the modules kat hist and kat gcp (default parameters, k=27). Ploidy was estimated using KMC v3.2.1 (Kokot et al. 2017) with parameters -k 27 -ci 1 -cs 10000 and Smudgeplot v0.2.5 (Ranallo-Benavidez et al. 2020) with default parameters.

Genome assembly and scaffolding

PacBio HiFi reads were assembled using hifiasm v0.16 (Cheng et al. 2021) with parameter '-l 2 ' Primary contigs were purged of artifactual duplications using minimap2 v2.24 (Li 2021) with parameter '-x map-hifi' and purge_dups v1.2.5 (Guan et al. 2020) with thresholds '-l 6 -m 89 -u 327'. Hi-C reads were first trimmed using cutadapt v1.15 (Martin 2011) with parameters '-m 10 -a AGATCGGAAGAG -A AGATCGGAAGAG' and then pre-processed using BWA v0.7.17 (Li 2013) and hicstuff v3.1.1 (Matthey-Doret et al. 2020) with parameters '-e DpnII,HinfI -m iterative -a bwa'. The resulting Hi-C sparse matrix was used to scaffold the assembly using instaGRAAL (Baudry et al. 2020) on Galaxy (Galaxy Version 0.1.6; (Galaxy Community 2022) with parameters '-l 5 -n 50 -c 1 -N 5'. These scaffolds were curated using instaGRAAL-polish with parameters '-m polishing -j NNNNNNNNNN'. Gaps were filled using TGS-GapCloser v1.1.1 (Xu et al. 2020) and the non-CCS PacBio reads with parameters '–tgstype pb –ne'. The scaffolds were finally polished by mapping the PacBio HiFi reads using minimap2

v2.24 with parameter '-x map-hifi' and running HyPo v1.0.3 (Kundu et al. 2019). Assembly was curated using hicstuff for iterative mapping, PretextMap v0.1.9 and PretextView v0.2.5.

Assembly evaluation

Ortholog completeness was assessed using the tool Benchmarking Universal Single-Copy Orthologs (BUSCO) v.5.0.0 (Manni et al. 2021) against the Arthropoda odb10 lineage (1,066 orthologs) and the Arachnida odb10 lineage (2,934 orthologs). *k*-mer completeness was evaluated using KAT v2.4.2 (Mapleson et al. 2016) and the module kat comp with default parameters. Hi-C reads were mapped to the assembly using BWA (Li 2013) and hicstuff (Matthey-Doret et al. 2020) as previously described. The contact map was generated using the module hicstuff view with the parameter '-b 500'. The mapped PacBio HiFi reads were mapped to the final scaffolds using minimap2 v2.24 (Li 2021) with parameters '-ax map-hifi' and the mapped reads were sorted with SAMtools v1.11. The final scaffolds were aligned against the nucleotide database using the Basic Local Alignment Search Tool (BLAST) v2.6.0 (Altschul et al. 1990) with parameters '-outfmt "6 qseqid staxids bitscore std sscinames scomnames" -max_hsps 1 -evalue 1e-25'. The outputs of minimap2, BLAST, and BUSCO (against the Arachnida odb10 lineage) were provided as input to BlobTools2 v2.3.3 (Challis et al. 2020). Sequences identified as bacteria were subsequently removed.

Repeat annotation

Transposable elements (TE) were annotated with the EDTA pipeline v1.9.4 (Bell et al. 2022). Long terminal repeats (LTR) were predicted by LTRharvest v1.6.1 (Ellinghaus et al. 2008) and LTR_FINDER v1.07 (Xu and Wang 2007) and then filtered using LTR_retriever v2.9.0 (Ou and Jiang 2018). Helitron transposons were identified using HelitronScanner (Xiong et al. 2014), and other repeats were detected using Generic Repeat Finder v1.0 (Shi and Liang 2019) and TIR-learner (Su et al. 2019). The final TE library was produced after further filtering and repeat annotation with RepeatModeler v2.0.1 (Flynn et al. 2020). The output hardmasked assembly was converted into a softmasked assembly. In addition, telomeres were masked after identification using Telomic Identifier v0.2.1 with parameter '-l 5'.

Gene prediction and functional annotation

RNA-seq reads were trimmed using cutadapt v1.15 (Martin 2011) with parameters '-m 10 -a AGATCGGAAGAG -A AGATCGGAAGAG' and assembled into transcripts using Trinity v2.14 (Grabherr et al. 2011) with default parameters. The genome assembly was then annotated using the pipeline Funannotate v1.8.13 using the trimmed RNA-seq reads and the transcriptome assembly as input. First, the module 'funannotate train' pre-processed the RNA-seq reads by mapping them using

HISAT2 v2.2.1 (Kim et al. 2019) and providing the output to StringTie v.2.2.1 (Shumate et al. 2022), and processed the transcripts through PASA v2.5.2 (Haas et al. 2003). The initial predictions were provided as input to 'funannotate predict' which combined them with predictions from Augustus v3.3.3 (Stanke et al. 2008) using EVidenceModeler v1.1.1 (Haas et al. 2008).

Data preprocessing

Detailed scripts can be found on Github at https://github.com/TheBastLab/sexdetermination.

Raw reads were first checked for quality with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/). Specific adapter sequences of paired-end reads were provided (see Github) and removed and a quality phred score cut off of '30' was performed with TrimGalore v0.6.5 (Krueger et al. 2023). For the allele we used the unmasked reference genome. Trimmed reads were mapped against both versions of the in-house reference genome using BWA v0.7.17 (Li 2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM (Li 2013). Using SAMtools v1.11 (Danecek et al. 2021) SAM files were converted into BAM files and sorted using GATK SortSam v4.1.9.0 (McKenna et al. 2010). Duplicates were then marked using Picard (http://broadinstitute.github.io/picard/) and reindexed.

Coverage analysis

We used deepTools bamCoverage v3.5.1 (Ramírez et al. 2014) to calculate the coverage as the number of reads per bin. We set a window size of 1 kb and performed the normalization according to the scaling factor *reads per kilobase transcript per million mapped reads* (RPKM; García et al., 2016). The results from the two pools were then merged for further analysis.

Allele frequency calculation

Allele frequencies (AF) were calculated using MAPGD -pool (Lynch et al. 2014; Ackerman et al. 2017), which uses a maximum likelihood approach to estimate allele frequencies from pooled population genomic data. Probabilities of polymorphism are calculated using the log-likelihood ratio test (LLR) by testing the null hypothesis of no polymorphism. The LLR threshold for identifying a polymorphism corresponds to a significant level of p = 0.00001 (-a 22) (Lynch et al. 2014; Guirao-Rico and González 2021). Allele frequencies were calculated for 4,663,024 million sites, with frequencies calculated based on the major allele in the female and male pools. According to common practice, loci are assumed to be monomorphic/fixed for AF in the range of 0.95 to 1 (Chakraborty et al. 1980).

Restriction enzyme digestion

Restriction sites were identified using the online tool "GenScript Restriction Enzyme Map Analysis Tools". (https://www.genscript.com/tools/restriction-enzyme-map-analysis, last accessed on January 21, 2023). Restriction enzyme HindIII was selected for chromosome 1, as the site is exclusive to the female sex. Primers were designed using primer3 with default values (forward: 'ACACGACAACGCGTCTTTAAT', reverse: 'AAAACTCCTGGTTCGCAGTTT'). For PCR we used 1 µl DNA, 10 µl DreamTaq Green PCR Master Mix (2X) (ThermoFisher), 2 µl forward, 2 µl reverse primer and 5 µl ddH_2O. Each sample was PCR-amplified in a thermal cycler (BentoLab) with initial denaturation at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 60°C and 1 min at 72°C. Following ThermoFisher instructions for DNA digestion with a single restriction enzyme, we used 1 µl PCR product (~2-4 ng/µl), 2 µl 10x Puffer R, 1 µl HindIII, 16 µl ddH_2O. The reaction was incubated at 37°C for 1.5 hours. The samples were run on a 1.5 % agarose gel with 12 µl SYBR™ Green I (ThermoFisher) and TBE 1×

Funding: German Research Foundation Emmy Noether grant BA 5800/4-1 and core funding of group (JB).

Author contributions:

Conceptualization: JB

Resources: HÖ, SW

Methodology: JB, HÖ, SW, NG, DLJ

Investigation: SW, HÖ

Visualization: SW, NG, HÖ

Funding acquisition: JB

Project administration: JB, DLJ, HÖ

Supervision: JB, DLJ, HÖ, NG

Data curation: NG

Formal Analyses: SW, NG

Writing – original draft: SW, HÖ, with input from all authors

Writing – review & editing: SW, JB

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All code for the analyses can be found at https://github.com/TheBastLab/sexdetermination. Data will be available shortly.

Chapter 5

Science Outreach and Empowerment: Bridging the Gap for Diverse Audiences

As the scientific process usually excludes the public and addresses mainly other scientists with a specific scientific language, research awareness is mainly confined to the respective scientific field. This poses barriers to communication with the public. The rise of misinformation and distrust, exacerbated by the COVID-19 pandemic, challenges the credibility of the scientific process (Nasr 2021). Many headlines such as 'Why don't people trust science?' (2022) from the European Commision (CORDIS), 'Why so many people distrust science and how we can fix it' from Greenfield (2022) in University World News and 'Widespread distrust in science: Is the way we communicate to blame?' from Boyle (2022) in AAMC Communities Network, were circulating in media during the pandemic. To counter these misconceptions, scientific outreach plays a vital role by providing insight into scientific processes in an accessible and engaging way for diverse audiences. Especially among children, this outreach can foster curiosity and critical thinking and establish a foundation for informed decision-making and potential career inspirations in science. Thus, scientific outreach can contribute to broader demographic diversity in higher education, through the empowerment of children. Academic achievement gaps are driven by factors like race, ethnicity, family socioeconomic status, and parental education level (Ahram, Fergus, and Noguera 2011; Pfeffer 2018). A German study found that the likelihood of achieving a doctorate drops by half if only one parent is non-German (Krempkow 2017). The biggest difference in transitional quota was observed early on between children aged nine and ten, where 37 % of children with migration background went on to schools that allow college admissions, whereas for children without migration background 57 % did. Early sustained intervention in the form of educational programs can help reduce the achievement gap between children (Temple, Ou, and Reynolds 2022). The consistent recognition of the importance of diversity in science underscores the ongoing global lack of diversity in higher education (Bernard and Cooperdock 2018; Judd and McKinnon 2021; Forrester 2020; *Nature* 2018).

Frontiers for Young Minds is a unique journal that facilitates collaboration between children and scientists in the peer-review process of articles tailored for young readers. The journal involves 6500 young reviewers aged 8-15, who give insights to enhance article accessibility for their peers. Supporting them are 819 science mentors, who are in direct contact with the young reviewers guiding them through peer review. Articles cover new research or shed light on core concepts of various fields such as 'Biodiversity', 'Engineering and Technology', 'Astronomy and Physics', and more (https://kids.frontiersin.org/about/journal, last access: 13.04.23 13:44).

My article, 'Having Babies in Soil: Is Sex Really Necessary?' was published as part of the collection in soil biodiversity aiming to fight the misconception that soil is "just dirt below our feet" (https://blog.frontiersin.org/2020/04/28/children-in-science-soils-are-alive/ last access: 13.04.23 13:44). The article introduces diverse reproductive modes of soil organisms, highlighting parthenogenesis in oribatid mites. It explains core concepts of evolutionary biology like the two-fold cost of sex and mutational meltdown. The collaborative effort involved researchers from the terrestrial ecology working group at the University of Cologne, the J.F. Blumenbach Institute of Zoology and Anthropology of the University of Göttingen, and the Department of Ecology and Evolution of the University of Lausanne. This article was translated into German making it accessible to an even broader audience. Further translations into French, Japanese, and Turkish are in progress. It also forms the basis for workshops conducted with Stiftung Wissen, an organization by the Sparkasse KölnBonn promoting education, science and research (https://www.stiftung-wissen-koelnbonn.de/, last access 7.08.23). Moreover, it was integrated into the biology curriculum for secondary school in North-Rhine Westphalia shaping scientific knowledge acquisition ("Wissenschaftliche Erkenntnisgewinnung") for students aged 10 to 16 ("Kernlehrplan für Sekundarstufe I NRW (Biologie)") in the winter term 2022 and summer term 2023, across two schools around Cologne.

Investing in science outreach remains crucial to ensure quality science education for children and engaging exposure to science.

Having babies in soil: is sex really necessary?

Öztoprak H.[*], Brandt A., Solbach M.D., Bast J. & Schaefer I.

Front. Young Minds 9:611659 (2021)

*Corresponding author

Key words: asexual, sex, parthenogenesis, reproductive mode, cost of sex

ABSTRACT

Finding a partner and having sex to produce babies is a common way to reproduce. Yet, upon closer look, we see that nature provides many ways for reproduction. What about a world without males? What first sounds impossible is the reality for many organisms that reproduce asexually, meaning without having sex. Females produce daughters that are clones of themselves, so no partner is required and males are dispensable. An Example of such all-female societies are several species of oribatid mites, which live in soils. These mites were already on earth long before the dinosaurs. Have oribatid mites always been asexual? Why do they reproduce without males? Does asexual reproduction have any advantages? Keep reading to learn about asexual reproduction and why oribatid mites are a key organism to investigate the question, "Why sex?"

REPRODUCTION IN SOIL ORGANISMS

All living organisms reproduce to generate new offspring. Almost all organisms, including humans, use some form of sex to reproduce. In **sexual reproduction**, an egg cell produced by a female and a sperm cell produced by a male fuse. The result is a **zygote** that develops into a unique offspring. Each offspring is a mixture of its parents, since it inherits half of its DNA from its mother and half from its father. The new offspring grows, becomes an adult, finds a partner, and finally produces offspring itself. This is the circle of life.

Many kinds of organisms live in soil, and they have a variety of different ways to reproduce. For example, earthworms are **hermaphrodites**, which means that one worm has both male and female reproductive organs. Earthworms reproduce sexually – if two earthworms meet, they exchange sperm cells that fuse with the egg cells of the other earthworm (**Figure 1A**). Since they have both male and

female sex organs, earthworms do not have to worry about finding a partner of the opposite sex because there is always a match. Earthworms belong to a group of soil organisms called the macrofauna, which includes all animals larger than 2 mm. Macrofauna are soil giants compared to most soil organisms.

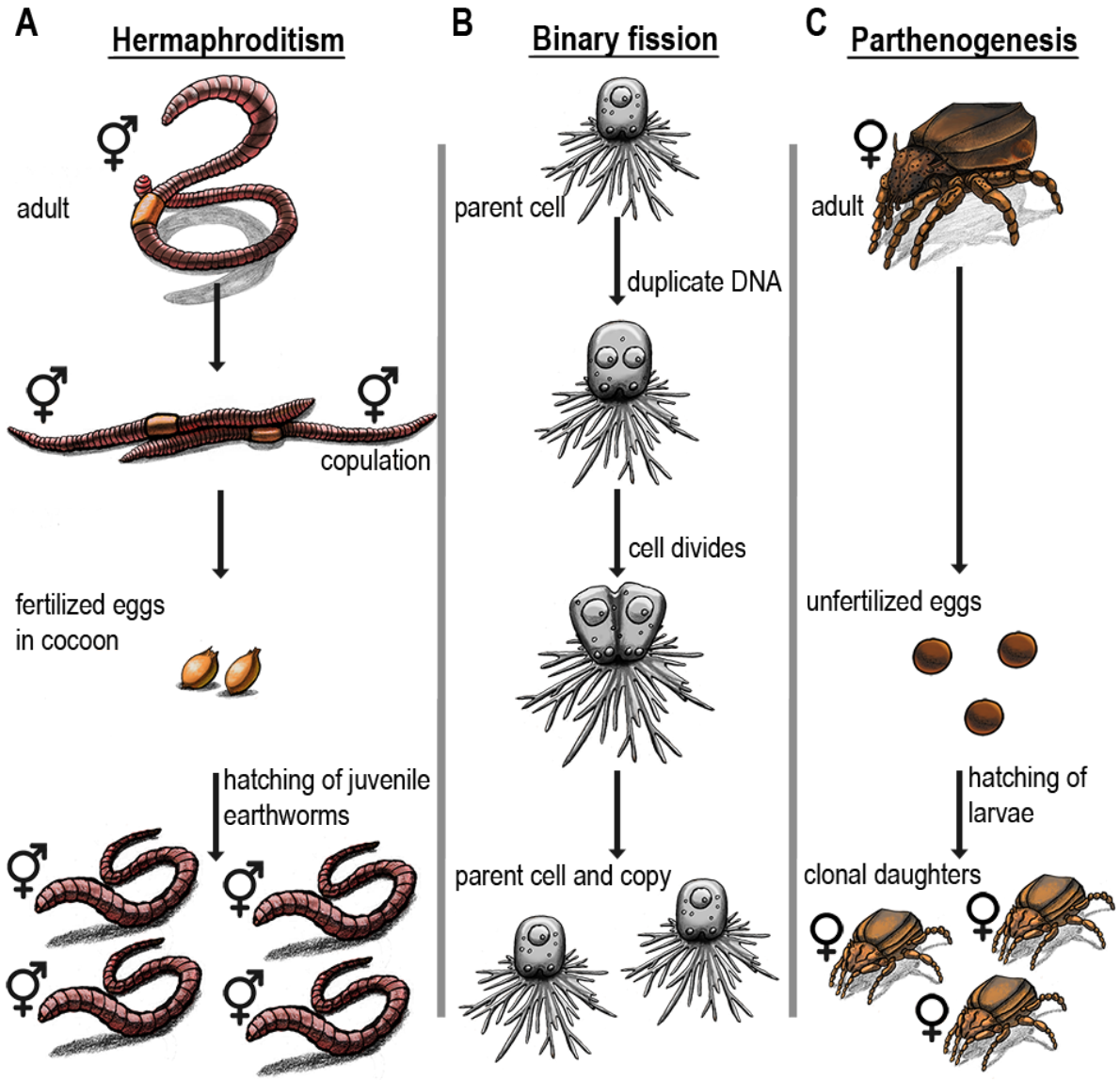


Figure 1 – Overview of different reproductive modes described in the text. (**A**) In hermaphroditism, each individual possesses both male and female reproductive organs; earthworms fertilize each other and each individual lays eggs. (**B**) In binary fission, a cell divides into two cells of equal size after duplicating its genetic material. Two individuals are produced from a single parent cell. This method is used by many Protists, for example amoebae. (**C**) Parthenogenesis is a form of asexual reproduction in which an offspring develops from an unfertilized egg cell. Some oribatid mites reproduce via parthenogenesis.

The greatest number of soil organisms belongs to the microfauna, which consists of organisms smaller than 0.1 mm. Most are single-celled organisms called **protists**. They are so small that one handful of soil contains more protists than there are humans in the entire world. For reproduction, protists do not need a partner at all. Some protists reproduce via a type of **asexual reproduction**, by making exact copies of themselves through a process called **binary fission**. First, they duplicate all their genetic material, and then one cell divides in two (**Figure 1B).** If repeated several times, many identical copies of one individual independently populate the soil.

Another essential group of soil organisms is known as the mesofauna, which includes all soil animals between 0.1 and 2 millimeters in size. The soil mesofauna includes springtails and oribatid mites. These organisms are very common and play a key role in the soil food web. They shred dead organic matter from plants, making the nutrients available to other organisms, including bacteria and fungi, which are then eaten by other belowground organisms. Many oribatid mites live in all-female populations and have done so for millions of years. They do not need sex for reproduction, because they can lay eggs that develop without being fertilized by a male, in a process called **parthenogenesis** (**Figure 1C**). Each egg contains the DNA only of the mother, which means that the offspring are clones of the mother.

ADVANTAGES OF ASEXUAL REPRODUCTION

Copying or cloning oneself seems to be much easier than finding a partner, but there are even more advantages of asexual reproduction. If you follow a sexual and an asexual population over time, two major differences appear. An asexual female only produces daughters, and these daughters produce only daughters again when *they* reproduce. The sexual females, however, must produce sons to fertilize the eggs – but only daughters can produce offspring. So, even if sexual and asexual females produce the same number of offspring, the asexual female has more daughters, which means more offspring that can reproduce.

Over time, the asexual population grows much faster and may outcompete the sexual population just by numbers (**Figure 2**). Scientists call this the “cost of males”. In addition to providing greater population growth, asexual reproduction seems to have other advantages: no sexually transmitted diseases, no energy loss, and no chance of being eaten by predators while trying to find a mate.

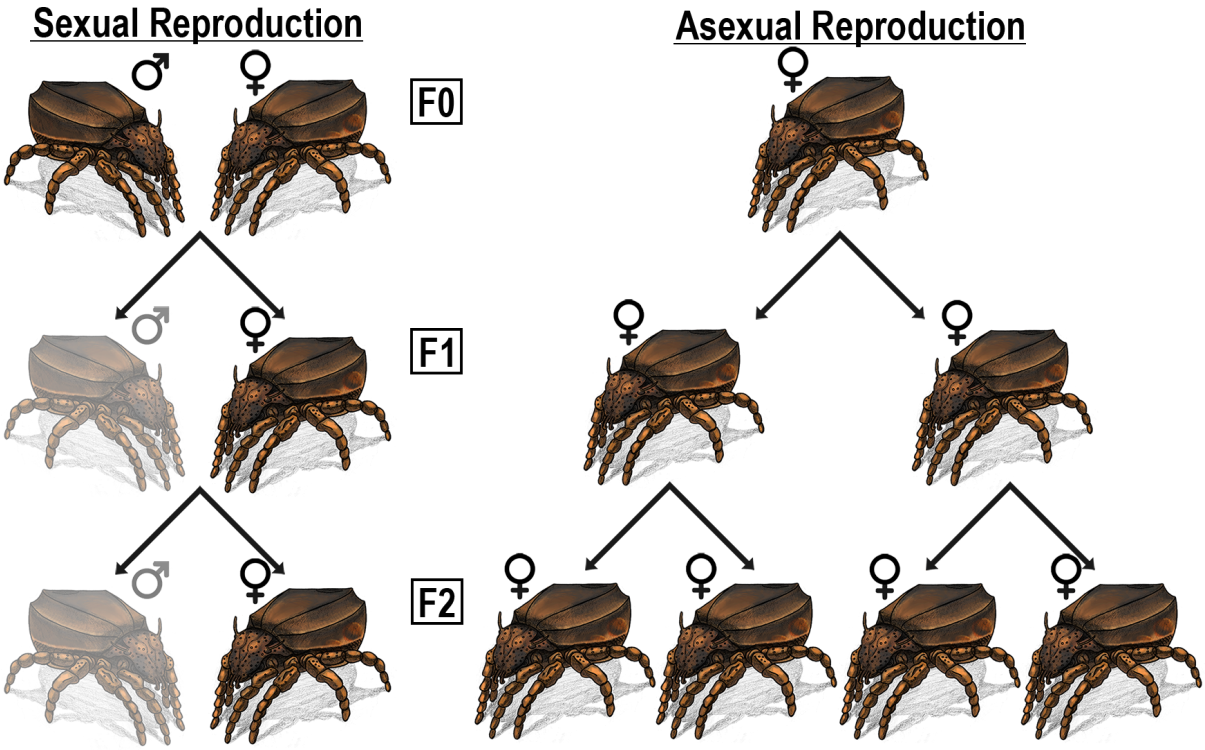


Figure 2 – Predicted population growth for sexual and asexual populations. In this example, each female produces two offspring. The size of the sexually reproducing population remains constant over time, because males are needed to fertilize the eggs of the females, but they cannot have offspring themselves. In asexual reproduction, the female produces twice as many childbearing offspring (females), leading to exponential population growth. (F0: parental generation, F1: first set of offspring from parents, F2: next generation of offspring from F1)

WHAT IS THE USE OF SEX IF REPRODUCTION CAN HAPPEN WITHOUT IT?

If there are successful ways of asexual reproduction, why do **eukaryotes** bother with complex, risky, and costly sexual reproduction? More than 98 % of all animals use sex to reproduce. This means that sexual reproduction must have clear advantages over asexual reproduction. Therefore, scientists try to explain the benefits of sex mainly by looking for potential problems if sex is not used. One disadvantage that scientists proposed for asexual reproduction has to do with mutations. Mutations are changes in the DNA that are an important cause of diversity among organisms. Occasionally, mutations are beneficial – sometimes they are very harmful – and most of the time they are slightly harmful. When an organism copies itself all the time, slightly harmful mutations keep adding up over generations, causing more and

more harm. Once they build up enough, these mutations could lead to the extinction of the species. This does not happen in sexual organisms, because the harmful mutations of one parent can be compensated for by the unmutated DNA from the other parent. Think of two bikes, one with a flat tire and one with a broken pedal. By combining parts from the two, you could still have a functional bike. It would be preferable to have one fully functional instead of two semi-functional ones. Some scientists suggest that this repair mechanism is an essential advantage of sexual over asexual reproduction.

Furthermore, just making copies or clones means that an organism will stay the same for many generations. This leads to problems when the environment changes. For example, the availability of resources such as food could change over time, e.g. due to climate change or the presence of other organisms competing for food. Food may develop a defense strategy, like running faster or becoming poisonous. Parasites can also be a problem. Since all individuals in an asexual population are very similar, they will not have the tiny differences that could help them adapt quickly enough to the ever-changing environment, so they would eventually die out. This means that offspring produced through sexual reproduction have higher chances of survival simply because they are different from those that came before them. This seems to tell us that all asexual organisms should go extinct in the long term. But, knowing that so many asexual organisms thrive in soil, we asked whether asexual organisms truly are doomed to die.

GENES CAN TELL US ABOUT ASEXUAL REPRODUCTION

Our research groups work with soil-living oribatid mites because many asexual species have survived without males for millions of years. We analyze the genes of sexual and asexual oribatid mite species because genes can tell us what happened to organisms in the past. Imagine genes as a captain's logbook, in which important things that happened to an organism in the past were recorded and passed on to future generations. We are still able to see the disadvantages of asexual reproduction in these logbooks. If we compare logbooks of oribatid mites that reproduce sexually to those that reproduce asexually, we can figure out which of the above-mentioned problems occurred and how they were solved. Since we are comparing two very similar species of mites, most of the genes are very similar. However, for some genes we can identify certain differences that must be caused by the consequences of the different reproductive modes.

In contrast to what scientists theorized, we found that asexual species do *not* accumulate more harmful mutations than sexual species (**Figure 3**) [1]. They do not need to combine two broken bikes to get a functional bike, because they keep the bikes functional. We also found that the asexual mites *do*

maintain variability in their large populations [2]. The genes in separate mother-daughter lines are as different from each other as are individuals that mix genes by reproducing sexually. Last, we found that the genes of two (or more) populations of asexual oribatid mites can be as varied as the genes of sexual species. This means that asexual oribatid mites *do not* stay the same for many generations, so they *can* adapt to new environments and are even able to diverge into new species [3].

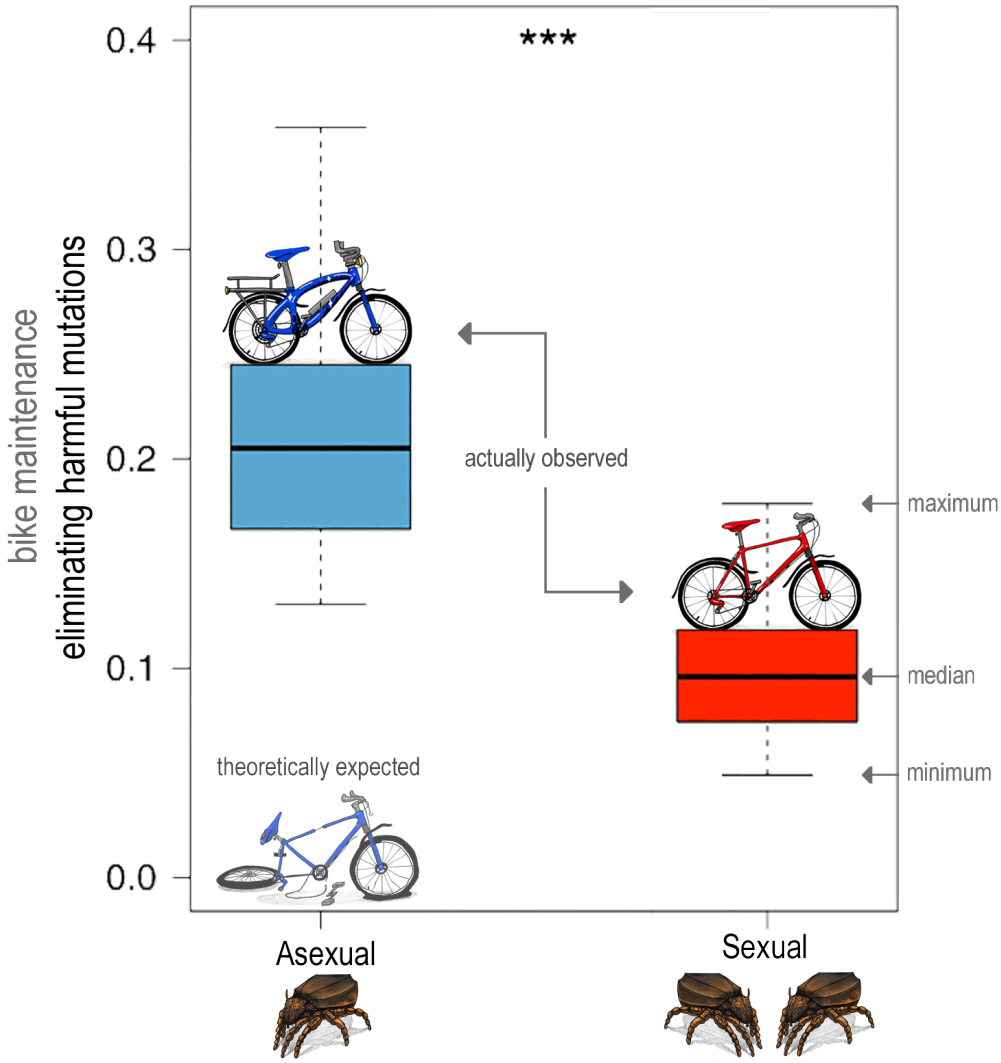


Figure 3 – Asexual oribatid mites get rid of slightly harmful mutations even more effectively than sexual oribatid mites [1]. Asexuals are even better at maintaining healthy genes (indicated by the fancy blue bike) than the sexual species (which still have decent red bikes). A boxplot shows the middle 50% of the data, each with a line in the middle representing the median (center of data). The two lines outside the box indicate the highest (maximum) and lowest (minimum) observations. The three stars on top indicate that this result is statistically significant, meaning it is very unlikely to observe this just by chance.

From our research, it appears that asexual oribatid mites have no disadvantages compared with the ones that reproduce sexually. They keep mutations in check, maintain genetic variability, and adapt over time, all without sex!

WHY IS THIS IMPORTANT?

There are many ways to produce offspring, and they do not always include sex. In theory, asexual species are expected to go extinct. In nature, however, there are various asexual organisms that have persisted over time. There must be ways to overcome the disadvantages of asexuality, as we saw in the asexual oribatid mites. But how do they do it? Do they have special mechanisms that fix mutations? Do they have to maintain a large population size to keep genetic variability? Is it easier to be asexual if you feed on a dead food source, which cannot develop defense strategies that you must adapt to? One future challenge will be to find the mechanisms that most asexual organisms have in common. Scientists continue to work on pairs of related sexual and asexual organisms to shed light on the variety of reproduction modes and to attempt to answer the question, "Why is sex so common?"

Glossary

Zygote: A cell resulting from the fusion of a male's sperm cell with a female's egg cell i.e., a zygote is a fertilized egg.

Hermaphrodite: A single individual that can produce both egg cells and sperm cells. It is basically both a female and a male.

Eukaryotes: can be animals, plants, single-celled organisms like protists and fungi. These are complex cells, where the genetic information is organized in a nucleus.

Parthenogenesis: A form of asexual reproduction in which an offspring develops from an unfertilized egg cell.

Sexual reproduction: Reproductive mode that requires an egg cell of a female and sperm cell of a male individual. The offspring is a unique individual that carries half of each parent's DNA.

Asexual reproduction: Any type of reproductive mode that produces offspring without the fusion of egg and sperm cells, also called gametes. The offspring are identical to the parent and/or to each other.

Binary fission: A cell divides into two cells of equal size after duplicating its genetic material. Two individuals are produced from a single parent cell i.e., one daughter and one parent cell.

Protist: Single celled organism that are eukaryotes. Many feed on bacteria and release excess nitrogen, providing essential nutrients for plants and other organisms.

Acknowledgments

We thank Mascha Gavrik for valuable comments on the manuscript. This study was supported by funding from the German Research Foundation fellowship BA 5800/3-1 to JB and core funding of the Scheu lab.

References

[1] A. Brandt, I. Schaefer, J. Glanz, T. Schwander, M. Maraun, S. Scheu and J. Bast, “Effective purifying selection in ancient asexual oribatid mites,” *Nat. Commun.*, vol. 8, no. 1, p. 873, Oct. 2017.

[2] S. C. Palmer and R. A. Norton, “Genetic diversity in thelytokous oribatid mites (Acari; Acariformes: Desmonomata),” *Biochem. Syst. Ecol.*, vol. 20, no. 3, pp. 219–231, Apr. 1992.

[3] M. Heethoff, K. Domes, M. Laumann, M. Maraun, R. A. Norton, and S. Scheu, “High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida),” *J. Evol. Biol.*, vol. 20, no. 1, pp. 392–402, Jan. 2007.

General Discussion

Importance of Single Individual Genomics from Natural Populations

Natural populations are crucial for understanding evolution's complex tapestry and validating evolutionary theories. These theories often arise from a blend of conceptual frameworks, natural observation, and laboratory experiments. Charles Darwin's observations of finches on the Galapagos Islands provide one of the most famous examples of natural populations informing evolutionary theory (Darwin 1859; Grant 2017). Evidence of natural selection was gathered during the Industrial Revolution in England, where the change in the moth population's coloration was observed over just 30 years (Kettlewell 1958; Cook 2003). The paradox of sex remains a complex and enigmatic question in evolutionary biology (Hörandl 2013; Smith and Maynard-Smith 1978; Neiman et al. 2018). Obligate asexuality is predicted to be associated with numerous effects on genome evolution (Jaron et al. 2021; Glémin and Galtier 2012; Hörandl et al. 2020). One of the predictions is the divergence of alleles in the absence of sex over evolutionary time also known as the Meselson effect (Welch and Meselson 2000). Evidence of the Meselson effect in parthenogenetic animals stems from the oribatid mite *Oppiella nova* (Brandt et al. 2021). Yet, this study was not able to exclude the possibility of nonconical forms of sex outside the regions featuring the Meselson effect, because of their scaffold-rich assemblies.
This thesis established a high-molecular-weight *g*DNA extraction protocol (for oribatid mites) and thus transformed ancient asexual mite research. It addresses the gap in high-quality phased genomes of non-model organisms (Chapter 1). It provided the chromosome-level reference genome of a single *Platynothrus peltifer* individual and its sexual outgroup *Hermannia gibba* from the same location. As well as additional genomes representing worldwide populations, building *P. peltifer* into a new model organism for evolutionary biology (Chapters 2-3).

Unraveling Ancient Asexuality

One long-standing open question was whether *P. peltifer* is indeed reproducing without sex over evolutionary time. Indeed, we can confirm our hypothesis: *P. peltifer* is an ancient asexual. It has been reproducing parthenogenetically since possibly around 20 million years (Chapter 2). Its haplotypes evolve independently, which is reflected by the Meselson effect. Our second hypothesis can also be confirmed as this independent haplotype evolution provides the means for adaptation and evolution in the absence of sexual reproduction. The evolutionary success of oribatid mites follows in the footsteps of bdelloid rotifers, which have been integral in ancient asexual evolution research thus far. Theories about oribatid mite ecology, mainly about their resource acquisition and high population sizes have

been proposed to explain some of their evolutionary success (Brandt et al. 2017; Maraun et al. 2012). With the gained insights from this thesis we can now conclude more holistically: their persistence is likely a combination of their cytological predisposition, genomic innovation, and probably their broad ecological spectrum: (i) oribatid mite cytology is largely unknown. The suggested automictic thelytoky, with terminal fusion and 'inverted meiosis' in parthenogenetic oribatid mite species, lacks definitive evidence (Bergmann et al. 2018). However, it must be a form of automixis that preserves heterozygosity over generations, as such 'effective clonality' is a requirement for the found haplotypic divergence. It further provides the foundation for the shared fluctuations of heterozygosity and nucleotide diversity observed among various populations (Chapters 2 and 3). These patterns are probably also facilitated by similar mutation rates and/or purifying selection among populations. (ii) Haplotypic independence observed in *P. peltifer* harbors the potential of genomic innovation as it mimics whole-genome duplication. Across various metazoans, whole-genome duplication events enabled neofunctionalization and subfunctionalization (Kenny et al. 2016; Ohno 1970). Evolutionary innovation after whole genome duplication was shown as functional diversification of genes (Storz et al. 2013), increased tolerance to environmental stress (Nowell et al. 2018), and enhanced adaptive radiation and speciation (cichlid fishes; Van de Peer et al. 2009). Similarly, *P. peltifer* exhibits allelic diversification and functionalization (Chapter 2), which might facilitate the conservation of function in one haplotype, thus increasing mutational robustness, while simultaneously maintaining and facilitating genetic variation (Fares 2015). (iii) Oribatids are key organisms in the soil food web. Their role as primary decomposers during soil processes is essential in regulating propagating microorganisms (Coleman et al. 2017). The association between the high frequency of parthenogenetic oribatids and their high densities supports predictions of the Structured Resource Theory of Sex (Maraun et al. 2012; Pequeno et al. 2022; Scheu and Drossel 2007). This theory suggests that the abundance of dead organic matter in soil facilitates asexuality, as resource acquisition is not threatened by biotic factors.

Unraveling Asexual Speciation

Asexual speciation remains a topic of debate as it is regarded as relatively rare compared to speciation through sexual reproduction. However, in some taxa like the oribatids this phenomenon might occur more frequently, as their ecology and evolutionary persistence provide a suitable foundation for evolutionary innovation, as presented in this thesis. Still, there is a contrast between the continuous process of evolution and the delimitation of definite entities, we call 'species'. The complexity of species delimitation is a constant hot topic in "species" related research. It raises the question of whether species delimitation is introducing more misconceptions than it is providing merits to research.

Untangling the genomic basis of asexual speciation will enhance our understanding of species and species-related research.

Most oribatids are cosmopolitan, with parthenogenetic oribatid mite species exhibiting a larger range than sexual species (Maraun et al. 2022). While *P. peltifer* is considered a single species, their globally distributed populations consist of a range of distinct genotypes. Long-term asexuality gives rise to multiple evolutionary lineages in the ancient asexual *P. peltifer*. This indicates that the persistence of ancient asexuals allows for diversification. The mitochondrial analysis supports seven different lineages of *P. peltifer* (Chapter 3), while genomic data indicate two separate lineages most likely following independent transitions to asexuality (Chapter 2). The divergence among the Japan-Canada and European clades raises questions about their common ancestry. Observed genetic variation in the *P. peltifer* nuclear genome (e.g. shared heterozygosity, conserved synteny) is mainly explained by haplotype divergence (Chapter 2,3). Still, we partially reject our previous hypothesis that asexual adaptation and diversification into novel species proceed differently from sexual species, as it does not involve large-scale rearrangements of structural variation. However, identifying smaller rearrangements and comparing synteny once all assemblies reach chromosome-level, is imminent to draw a conclusion. How the mechanisms leading to speciation differ between sexuals and asexuals needs to be investigated further. Still, the evolutionary history shown by mitochondrial genes (such as *cox1*) alone is incomplete and demonstrates how deceptive it can be to base population genetic research on just one marker (Noor and Feder 2006; Anderson 2001). Integrating further information from various sources such as morphology and ecology will boost our understanding of speciation in the ancient *P. peltifer*.

Discoveries in Sex Determination

The exploration of the *Hermannia gibba* genome, an important outgroup of the asexual cluster, led to the discovery of a ZW female heterogametic system in *H. gibba* (Chapter 4). This unexpected finding opens new avenues of inquiry, such as determining the frequency of ZW systems in oribatids and Acariformes and understanding their contribution to transitions to asexuality. Understanding sex determination in Acariformes provides further insights into ancestral sex-determination mechanisms within the Chelicerate lineage, which subsequently diversified into various sex-determination systems in higher taxa (Král et al. 2011).

Bridging Science and Education

The preparation of a manuscript addressing children exceeded the realm of academic research and transgressed to the public (Chapter 5). General concepts of evolutionary biology and the vast

biodiversity in soil reached younger audiences and were even taught in schools. This engagement fosters a broader understanding of the intricate world of oribatid mites and the significance of reproduction and soil biodiversity.

Conclusion

The genomic analyses of natural populations of *Platynothrus peltifer* and related organisms in this thesis offer insights into the consequences of asexual reproduction, asexual speciation, sex chromosome dynamics, and thus allow a glimpse into the tapestry of life's evolutionary journey without sex.

General Outlook

The success of establishing a reference assembly with phased haplotypes of the German *Platynothrus peltifer* and improving the assemblies of the other populations with Hi-C data, will allow chromosome-level comparison of e.g. genomic rearrangements to further untangle their intricacies. Further, investigating population-specific mutation rates has the potential to give insight into the effects shaping their different heterozygosity levels. Genome fluidity can be evaluated through pan/core-genomes as it will also highlight the conserved evolutionary history in *P. peltifer* populations. This will allow us to expand our understanding of asexual speciation. Moreover, incorporating multiple lines of evidence is imperative when delimiting species. Especially in cryptic species with conserved morphology but high diversity as in the oribatid mites an integrative approach for species delimitation is indispensable. To not underestimate species diversity and increase robustness in *P. peltifer*, morphological and ecological data should be acquired and critically discussed with the gained genomic knowledge; not only for species delimitation but also for the implications of their evolutionary trajectory.

Haplotype divergence is not universally observed in asexual organisms and thus far only conclusively documented in oribatid mites. As it is now possible to study the evolutionary innovation of *P. peltifer* it will be imperative to test other oribatids and compare replicates of independent transitions to asexuality among asexuals and sexual relatives to exclude lineage-specific patterns. Identifying similarities among the evolutionary replicates not only allows differentiation between species-specific effects but also identifies general signatures of asexual evolution in oribatid mites. Such further in-depth analyses will shed light on all the innovations buried in these ancient asexual genomes.

References

General References

Anderson, T. J. 2001. "The Dangers of Using Single Locus Markers in Parasite Epidemiology: Ascaris as a Case Study." *Trends in Parasitology* 17 (4): 183–88.

Barraclough, Timothy G. 2019. "The Evolutionary Biology of Species". *Oxford University Press*.

Barraclough, Timothy G., C. William Birky Jr, and Austin Burt. 2003. "Diversification in Sexual and Asexual Organisms." *Evolution; International Journal of Organic Evolution* 57 (9): 2166–72.

Jens Bast, Kamil S Jaron, Donovan Schuseil, Denis Roze, Tanja Schwander. 2019. "Asexual reproduction reduces transposable element load in experimental yeast populations" *eLife* 8:e48548

Bell, G. 1982. "The Masterpiece of Nature: The Evolution and Genetics of Sexuality. 1982." *London, United Kingdom: CroomHelm*.

Bergmann, Paavo, Michael Laumann, Roy A. Norton, and Michael Heethoff. 2018. "Cytological Evidence for Automictic Thelytoky in Parthenogenetic Oribatid Mites (Acari, Oribatida): Synaptonemal Complexes Confirm Meiosis in Archegozetes Longisetosus." *Acarologia* 58 (2): 342–56.

Birky, C. William, Jr, Joshua Adams, Marlea Gemmel, and Julia Perry. 2010. "Using Population Genetic Theory and DNA Sequences for Species Detection and Identification in Asexual Organisms." *PloS One* 5 (5): e10609.

Birky, C. W., Jr. 1996. "Heterozygosity, Heteromorphy, and Phylogenetic Trees in Asexual Eukaryotes." *Genetics* 144 (1): 427–37.

Brandt, Alexander, Ina Schaefer, Julien Glanz, Tanja Schwander, Mark Maraun, Stefan Scheu, and Jens Bast. 2017. "Effective Purifying Selection in Ancient Asexual Oribatid Mites." *Nature Communications* 8 (1): 873.

Brandt, Alexander, Patrick Tran Van, Christian Bluhm, Yoann Anselmetti, Zoé Dumas, Emeric Figuet, Clémentine M. François, et al. 2021. "Haplotype Divergence Supports Long-Term Asexuality in the Oribatid Mite *Oppiella Nova*." *Proceedings of the National Academy of Sciences of the United States of America* 118 (38).

Cavalli, Giacomo, and Tom Misteli. 2013. "Functional Implications of Genome Topology." *Nature Structural & Molecular Biology* 20 (3): 290–99.

Coleman, David C., Mac A. Callaham, and D. A. Crossley Jr. 2017. *Fundamentals of Soil Ecology*. Academic Press.

Cook, Laurence M. 2003. "The Rise and Fall of the Carbonaria Form of the Peppered Moth." *The Quarterly Review of Biology* 78 (4): 399–417.

Coyne, Jerry A., H. Allen Orr. 2004. *Speciation*. Vol. 37. Sinauer Associates Sunderland, MA.

Danchin, Etienne G. J., Marie-Noëlle Rosso, Paulo Vieira, Janice de Almeida-Engler, Pedro M. Coutinho, Bernard Henrissat, and Pierre Abad. 2010. "Multiple Lateral Gene Transfers and Duplications Have Promoted Plant Parasitism Ability in Nematodes." *Proceedings of the National Academy of Sciences of the United States of America* 107 (41): 17651–56.

Darwin, Charles. 1859. "On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life." 1909. *The Origin of Species*. PF Collier & son New York.

Debortoli, Nicolas, Xiang Li, Isobel Eyres, Diego Fontaneto, Boris Hespeels, Cuong Q. Tang, Jean-François Flot, and Karine Van Doninck. 2016. "Genetic Exchange among Bdelloid Rotifers Is More Likely Due to Horizontal Gene Transfer Than to Meiotic Sex." *Current Biology: CB* 26 (6): 723–32.

Engelstädter, Jan. 2008. "Constraints on the Evolution of Asexual Reproduction." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 30 (11-12): 1138–50.

Faddeeva-Vakhrusheva, Anna, Ken Kraaijeveld, Martijn F. L. Derks, Seyed Yahya Anvar, Valeria Agamennone, Wouter Suring, Andries A. Kampfraath, et al. 2017. "Coping with Living in the Soil: The Genome of the Parthenogenetic Springtail *Folsomia Candida*." *BMC Genomics* 18 (1): 493.

Fares, Mario A. 2015. "The Origins of Mutational Robustness." *Trends in Genetics: TIG* 31 (7): 373–81.

Felsenstein, J. 1974. "The Evolutionary Advantage of Recombination." *Genetics* 78 (2): 737–56.

Fisher, Sir Ronald Aylmer. 1999. The Genetical Theory of Natural Selection: *A Complete Variorum Edition*. *OUP Oxford.*

Flot, Jean-François, Boris Hespeels, Xiang Li, Benjamin Noel, Irina Arkhipova, Etienne G. J. Danchin, Andreas Hejnol, et al. 2013. "Genomic Evidence for Ameiotic Evolution in the Bdelloid Rotifer *Adineta Vaga*." *Nature* 500 (7463): 453–57.

Gandolfi, Andrea, Ian R. Sanders, Valeria Rossi, and Paolo Menozzi. 2003. "Evidence of Recombination in Putative Ancient Asexuals." *Molecular Biology and Evolution* 20 (5): 754–61.

Glémin, Sylvain, and Nicolas Galtier. 2012. "Genome Evolution in Outcrossing Versus Selfing Versus Asexual Species." In *Evolutionary Genomics: Statistical and Computational Methods, Volume 1*, edited by Maria Anisimova, 311–35. Totowa, NJ: Humana Press.

Grandjean, F. 1950. "Observations Ethologiques Sur *Camisia Segnis* (Herm.) et *Platynothrus Peltifer* (Koch)(Acariens)." *Bull. Mus. Nat. Paris*.

Grant, Peter R. 2017. Ecology and Evolution of Darwin's Finches *(Princeton Science Library Edition): Princeton Science Library Edition*. Princeton University Press.

Hammer, M., and Wallwork Ja. 1979. "A Review of the World Distribution of Oribatid Mites (Acari: Cryptostigmata) in Relation to Continental Drift."

Heethoff, M., K. Domes, M. Laumann, M. Maraun, R. A. Norton, and S. Scheu. 2007. "High Genetic Divergences Indicate Ancient Separation of Parthenogenetic Lineages of the Oribatid Mite *Platynothrus Peltifer* (Acari, Oribatida)." *Journal of Evolutionary Biology* 20 (1): 392–402.

Hernández-Hernández, Tania, Elizabeth C. Miller, Cristian Román-Palacios, and John J. Wiens. 2021. "Speciation across the Tree of Life." *Biological Reviews of the Cambridge Philosophical Society* 96 (4): 1205–42.

Hill, W. G., and A. Robertson. 1966. "The Effect of Linkage on Limits to Artificial Selection." *Genetical Research* 8 (3): 269–94.

Hörandl, Elvira. 2013. "Meiosis and the Paradox of Sex in Nature." *Meiosis* 1.

Hörandl, Elvira, Jens Bast, Alexander Brandt, Stefan Scheu, Christoph Bleidorn, Mathilde Cordellier, Minou Nowrousian, et al. 2020. "Genome Evolution of Asexual Organisms and the Paradox of Sex in Eukaryotes." In *Evolutionary Biology—A Transdisciplinary Approach*, edited by Pierre Pontarotti, 133–67. Cham: Springer International Publishing.

Jalil, M. 1972. "A Note on the Life Cycle of Platynothrus Peltifer." *Journal of the Kansas Entomological Society* 45 (3): 309–11.

Jaron, Kamil S., Jens Bast, Reuben W. Nowell, T. Rhyker Ranallo-Benavidez, Marc Robinson-Rechavi, and Tanja Schwander. 2021. "Genomic Features of Parthenogenetic Animals." *The Journal of Heredity* 112 (1): 19–33.

Judson, O. P., and B. B. Normark. 1996. "Ancient Asexual Scandals." *Trends in Ecology & Evolution* 11 (2): 41–46.

Keightley, Peter D., and Sarah P. Otto. 2006. "Interference among Deleterious Mutations Favours Sex and Recombination in Finite Populations." *Nature* 443 (7107): 89–92.

Kenny, N. J., K. W. Chan, W. Nong, Z. Qu, I. Maeso, H. Y. Yip, T. F. Chan, et al. 2016. "Ancestral Whole-Genome Duplication in the Marine Chelicerate Horseshoe Crabs." *Heredity* 116 (2): 190–99.

Kettlewell, H. B. Do. 1958. "A Survey of the Frequencies of Biston Betularia (L.)(Lep.) and Its Melanic Forms in Great Britain." *Heredity* 12 (1): 51–72.

Král, J., T. Kořínková, M. Forman, and L. Krkavcová. 2011. "Insights into the Meiotic Behavior and Evolution of Multiple Sex Chromosome Systems in Spiders." *Cytogenetic and Genome Research* 133 (1): 43–66.

Lehtonen, Jussi, Michael D. Jennions, and Hanna Kokko. 2012. "The Many Costs of Sex." *Trends in Ecology & Evolution* 27 (3): 172–78.

Maraun, Mark, Paul S. P. Bischof, Finn L. Klemp, Jule Pollack, Linnea Raab, Jan Schmerbach, Ina Schaefer, Stefan Scheu, and Tancredi Caruso. 2022. "'Jack-of-all-trades' Is Parthenogenetic." *Ecology and Evolution* 12 (6).

Maraun, Mark, Roy A. Norton, Roswitha B. Ehnes, Stefan Scheu, and Georgia Erdmann. 2012. "Positive Correlation between Density and Parthenogenetic Reproduction in Oribatid Mites (Acari) Supports the Structured Resource Theory of Sexual Reproduction." *Evolutionary Ecology Research* 14 (3): 311–23.

Mark Welch, D., and M. Meselson. 2000. "Evidence for the Evolution of Bdelloid Rotifers without Sexual Reproduction or Genetic Exchange." *Science* 288 (5469): 1211–15.

McDonald, Michael J., Daniel P. Rice, and Michael M. Desai. 2016. "Sex Speeds Adaptation by Altering the Dynamics of Molecular Evolution." *Nature* 531 (7593): 233–36.

McElroy, Kyle E., Stefan Müller, Dunja K. Lamatsch, Laura Bankers, Peter D. Fields, Joseph R. Jalinsky, Joel Sharbrough, Jeffrey L. Boore, John M. Logsdon, and Maurine Neiman. 2021. "Asexuality Associated with Marked Genomic Expansion of Tandemly Repeated rRNA and Histone Genes." *Molecular Biology and Evolution* 38 (9): 3581–92.

Morgan, Andrew D., Rob W. Ness, Peter D. Keightley, and Nick Colegrave. 2014. "Spontaneous Mutation Accumulation in Multiple Strains of the Green Alga, *Chlamydomonas Reinhardtii*." *Evolution; International Journal of Organic Evolution* 68 (9): 2589–2602.

Muller, H. J. 1964. "The Relation of Recombination to Mutational Advance." *Mutation Research* 106 (May): 2–9.

Neiman, Maurine, Patrick G. Meirmans, Tanja Schwander, and Stephanie Meirmans. 2018. "Sex in the Wild: How and Why Field-Based Studies Contribute to Solving the Problem of Sex." *Evolution; International Journal of Organic Evolution* 72 (6): 1194–1203.

Neiman, Maurine, Stephanie Meirmans, and Patrick G. Meirmans. 2009. "What Can Asexual Lineage Age Tell Us about the Maintenance of Sex?" *Annals of the New York Academy of Sciences* 1168 (June): 185–200.

Neiman, M., T. F. Sharbel, and T. Schwander. 2014. "Genetic Causes of Transitions from Sexual Reproduction to Asexuality in Plants and Animals." *Journal of Evolutionary Biology* 27 (7): 1346–59.

Noor, Mohamed A. F., and Jeffrey L. Feder. 2006. "Speciation Genetics: Evolving Approaches." *Nature Reviews. Genetics* 7 (11): 851–61.

Norton, R. A., and S. C. Palmer. 1991. "The Distribution, Mechanisms and Evolutionary Significance of Parthenogenesis in Oribatid Mites." In *The Acari: Reproduction, Development and Life-History Strategies*, edited by Reinhart Schuster and Paul W. Murphy, 107–36. Dordrecht: Springer Netherlands.

Nowell, Reuben W., Pedro Almeida, Christopher G. Wilson, Thomas P. Smith, Diego Fontaneto, Alastair Crisp, Gos Micklem, Alan Tunnacliffe, Chiara Boschetti, and Timothy G. Barraclough. 2018. "Comparative Genomics of Bdelloid Rotifers: Insights from Desiccating and Nondesiccating Species." *PLoS Biology* 16 (4): e2004830.

Ohno, Susumu. 1970. *Evolution by Gene Duplication*. Springer-Verlag: Berlin, Germany.

Ollivier, M., T. Gabaldón, J. Poulain, F. Gavory, N. Leterme, J-P Gauthier, F. Legeai, D. Tagu, J. C. Simon, and C. Rispe. 2012. "Comparison of Gene Repertoires and Patterns of Evolutionary Rates in Eight Aphid Species That Differ by Reproductive Mode." *Genome Biology and Evolution* 4 (2): 155–67.

Pequeno, Pedro Aurélio Costa Lima, Elizabeth Franklin, and Roy A. Norton. 2022. "Hunger for Sex: Abundant, Heterogeneous Resources Select for Sexual Reproduction in the Field." *Journal of Evolutionary Biology* 35 (10): 1387–95.

Ricci, Claudia N. 1987. "Ecology of Bdelloids: How to Be Successful." In *Rotifer Symposium IV*, 117–27. Dordrecht: Springer Netherlands.

Saltzwedel, Helge von, Mark Maraun, Stefan Scheu, and Ina Schaefer. 2014. "Evidence for Frozen-Niche Variation in a Cosmopolitan Parthenogenetic Soil Mite Species (Acari, Oribatida)." *PloS One* 9 (11): e113268.

Scheu, S., and B. Drossel. 2007. "Sexual Reproduction Prevails in a World of Structured Resources in Short Supply." *Proceedings. Biological Sciences / The Royal Society* 274 (1614): 1225–31.

Schön, Isa, Giampaolo Rossetti, and Koen Martens. 2009. "Darwinulid Ostracods: Ancient Asexual Scandals or Scandalous Gossip?" In *Lost Sex: The Evolutionary Biology of Parthenogenesis*, edited by Isa Schön, Koen Martens, and Peter Dijk, 217–40. Dordrecht: Springer Netherlands.

Schwander, Tanja. 2016. "Evolution: The End of an Ancient Asexual Scandal." *Current Biology: CB* 26 (6): R233–35.

Schwander, Tanja, and Bernard J. Crespi. 2009. "Twigs on the Tree of Life? Neutral and Selective Models for Integrating Macroevolutionary Patterns with Microevolutionary Processes in the Analysis of Asexuality." *Molecular Ecology* 18 (1): 28–42.

Seehausen, Ole, Roger K. Butlin, Irene Keller, Catherine E. Wagner, Janette W. Boughman, Paul A. Hohenlohe, Catherine L. Peichel, et al. 2014. "Genomics and the Origin of Species." *Nature Reviews. Genetics* 15 (3): 176–92.

Signorovitch, Ana, Jae Hur, Eugene Gladyshev, and Matthew Meselson. 2015. "Allele Sharing and Evidence for Sexuality in a Mitochondrial Clade of Bdelloid Rotifers." *Genetics* 200 (2): 581–90.

Smith, J. M., and J. Maynard-Smith. 1978. "The Evolution of Sex." *Genetics Research* 32 (3): b1–9.

Sobel, James M., Grace F. Chen, Lorna R. Watt, and Douglas W. Schemske. 2010. "The Biology of Speciation." *Evolution; International Journal of Organic Evolution* 64 (2): 295–315.

Storz, Jay F., Juan C. Opazo, and Federico G. Hoffmann. 2013. "Gene Duplication, Genome Duplication, and the Functional Diversification of Vertebrate Globins." *Molecular Phylogenetics and Evolution* 66 (2): 469–78.

Taberly, G. 1987. "Researches on the Thelytocous Parthenogenesis of Two Species of Oribatid Mites - *Trhypochthonius Tectorum* (Berlese) and *Platynothrus Peltifer* (Koch). I." *Acarologia* 28 (2): 187–98.

Tran Van, Patrick, Yoann Anselmetti, Jens Bast, Zoé Dumas, Nicolas Galtier, Kamil S. Jaron, Koen Martens, et al. 2021. "First Annotated Draft Genomes of Nonmarine Ostracods (Ostracoda, Crustacea) with Different Reproductive Modes." *G3* 11 (4).

Vakhrusheva, Olga A., Elena A. Mnatsakanova, Yan R. Galimov, Tatiana V. Neretina, Evgeny S. Gerasimov, Sergey A. Naumenko, Svetlana G. Ozerova, et al. 2020. "Genomic Signatures of Recombination in a Natural Population of the Bdelloid Rotifer *Adineta Vaga*." *Nature Communications* 11 (1): 6421.

Van de Peer, Yves, Steven Maere, and Axel Meyer. 2009. "The Evolutionary Significance of Ancient Genome Duplications." *Nature Reviews. Genetics* 10 (10): 725–32.

Wallace, R. L., and T. W. Snell. 1991. "Rotifera. Ecology and Classification of North American Freshwater Invertebrates." Academik Press. Inc.

White, M. J. D. 1977. Animal Cytology and Evolution. *CUP Archive*.

References Chapter 1

Application note "Considerations for using the low and ultra-low DNA input workflows for whole genome sequencing." (2021, August 19). PacBio. https://www.pacb.com/literature/application-note-considerations-for-using-the-low-and-ultra-low-dna-input-workflows-for-whole-genome-sequencing/

Bast, J., Schaefer, I., Schwander, T., Maraun, M., Scheu, S., & Kraaijeveld, K. (2016). No Accumulation of Transposable Elements in Asexual Arthropods. *Molecular Biology and Evolution*, *33*(3), 697–706.

Brandt, A., Van, P. T., Bluhm, C., Anselmetti, Y., Dumas, Z., Figuet, E., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., Labédan, M., Maraun, M., Parker, D. J., Robinson-Rechavi, M., Schaefer, I., Simion, P., Scheu, S., Schwander, T., & Bast, J. (2021). Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(38).

Brückner, A., Barnett, A. A., Bhat, P., Antoshechkin, I. A., & Kitchen, S. A. (2022). Molecular evolutionary trends and biosynthesis pathways in the Oribatida revealed by the genome of *Archegozetes longisetosus*. *Acarologia*, *62*(2), 532–573.

De Coster, W., & Rademakers, R. (2023). NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics*, *39*(5).

Eisenhauer, N., & Hines, J. (2021). Invertebrate biodiversity and conservation. *Current Biology: CB*, *31*(19), R1214–R1218.

Gregory, T. R. (2023). Animal genome size database. *http://www.genomesize.com*.

Heethoff, M., Domes, K., Laumann, M., Maraun, M., Norton, R. A., & Scheu, S. (2007). High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *Journal of Evolutionary Biology*, *20*(1), 392–402.

Jalil, M. (1972). A Note on the Life Cycle of Platynothrus peltifer. *Journal of the Kansas Entomological Society*, *45*(3), 309–311.

Judson, O. P., & Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution*, *11*(2), 41–46.

Kelley, J. L., Peyton, J. T., Fiston-Lavier, A.-S., Teets, N. M., Yee, M.-C., Johnston, J. S., Bustamante, C. D., Lee, R. E., & Denlinger, D. L. (2014). Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nature Communications*, *5*, 4611.

Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., & Clavijo, B. J. (2017). KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics*, *33*(4), 574–576.

Maraun, M., Heethoff, M., Scheu, S., Norton, R. A., Weigmann, G., & Thomas, R. H. (2003). Radiation in sexual and parthenogenetic oribatid mites (Oribatida, Acari) as indicated by genetic divergence of closely related species. *Experimental & Applied Acarology*, *29*(3-4), 265–277.

Marçais, G., & Kingsford, C. (2007). JELLYFISH–fast, parallel *k*-mer counting for DNA. *Bioinformatics*.

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*(3), 1215.

Remén, C., Krüger, M., & Cassel-Lundhagen, A. (2010). Successful analysis of gut contents in fungal-feeding oribatid mites by combining body-surface washing and PCR. *Soil Biology & Biochemistry*, *42*(11), 1952–1957.

Renker, C., Otto, P., Schneider, K., Zimdars, B., Maraun, M., & Buscot, F. (2005). Oribatid mites as potential vectors for soil microfungi: study of mite-associated fungal species. *Microbial Ecology*, *50*(4), 518–528.

Rhyker Ranallo-Benavidez, T., Jaron, K. S., & Schatz, M. C. (2020). GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications*, *11*(1), 1–10.

References Chapter 2

1. A. Brandt, P. Tran Van, C. Bluhm, Y. Anselmetti, Z. Dumas, E. Figuet, C. M. François, N. Galtier, B. Heimburger, K. S. Jaron, M. Labédan, M. Maraun, D. J. Parker, M. Robinson-Rechavi, I. Schaefer, P. Simion, S. Scheu, T. Schwander, J. Bast, Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021).

2. M. Heethoff, K. Domes, M. Laumann, M. Maraun, R. A. Norton, S. Scheu, High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *J. Evol. Biol.* **20**, 392–402 (2007).

3. A. Brandt, I. Schaefer, J. Glanz, T. Schwander, M. Maraun, S. Scheu, J. Bast, Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* **8**, 873 (2017).

4. J. Bast, I. Schaefer, T. Schwander, M. Maraun, S. Scheu, K. Kraaijeveld, No accumulation of transposable elements in asexual arthropods. *Mol. Biol. Evol.* **33**, 697–706 (2016).

5. I. Schaefer, R. A. Norton, S. Scheu, M. Maraun, Arthropod colonization of land - Linking molecules and fossils in oribatid mites (Acari, Oribatida). *Mol. Phylogenet. Evol.* **57**, 113–121 (2010).

6. M. Heethoff, R. A. Norton, S. Scheu, M. Maraun, "Parthenogenesis in oribatid mites (Acari, Oribatida): evolution without sex" in *Lost Sex - The Evolutionary Biology of Parthenogenesis*, I. Schoen, K. Martens, P. van Dijk, Eds. (Springer Press, 2009), pp. 241–257.

7. R. A. Norton, S. C. Palmer, "The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites" in *The Acari: Reproduction, development and life-history strategies*, R. Schuster, P. W. Murphy, Eds. (Springer Netherlands, Dordrecht, 1991), pp. 107–136.

8. M. Maraun, M. Heethoff, K. Schneider, S. Scheu, G. Weigmann, J. M. Cianciolo, R. H. Thomas, R. A. Norton, Molecular phylogeny of oribatid mites (Oribatida, Acari): evidence for multiple radiations of parthenogenetic lineages. *Exp. Appl. Acarol.* **33**, 183–201 (2004).

9. R. Butlin, The costs and benefits of sex: new insights from old asexual lineages. *Nat. Rev. Genet.* **3**, 311–317 (2002).

10. M. Terwagne, E. Nicolas, B. Hespeels, L. Herter, J. Virgo, C. Demazy, A.-C. Heuskin, B. Hallet, K. Van Doninck, DNA repair during nonreductional meiosis in the asexual rotifer *Adineta vaga*. *Sci Adv*. **8**, eadc8829 (2022).

11. G. Taberly, Recherches sur la parthéenogenèse thélytoque de deux espèces d'acariens oribatides: *Trhypochthonius tectorum* (Berlese) et *Platynothrus peltifer* (Koch). IV: Observations sur les males ataviques. *Acarologia*. **29**, 95–107 (1988).

12. S. C. Palmer, R. A. Norton, Genetic diversity in thelytokous oribatid mites (Acari; Acariformes: Desmonomata). *Biochem. Syst. Ecol.* **20**, 219–231 (1992).

13. S. Nokkala, Personal communication (2023).

14. S. C. Palmer, R. A. Norton, Further experimental proof of thelytokous parthenogenesis in oribatid mites (Acari: Oribatida: Desmonomata). *Exp. Appl. Acarol.* **8**, 149–159 (1990).

15. T. Pfingstl, H. Schatz, A survey of lifespans in Oribatida excluding Astigmata (Acari). *Zoosymposia*. **20**, 7–27 (2021).

16. C. W. Birky Jr, Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics*. **144**, 427–437 (1996).

17. J. Engelstädter, Asexual but not clonal: evolutionary processes in automictic populations. *Genetics*. **206**, 993–1009 (2017).

18. B. Dunn, T. Paulish, A. Stanbery, J. Piotrowski, G. Koniges, E. Kroll, E. J. Louis, G. Liti, G. Sherlock, F. Rosenzweig, Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. *PLoS Genet.* **9**, e1003366 (2013).

19. S. P. Otto, Selective interference and the evolution of sex. *J. Hered.* **112**, 9–18 (2021).

20. K. S. Jaron, D. J. Parker, Y. Anselmetti, P. Tran Van, J. Bast, Z. Dumas, E. Figuet, C. M. François, K. Hayward, V. Rossier, P. Simion, M. Robinson-Rechavi, N. Galtier, T. Schwander, Convergent consequences of parthenogenesis on stick insect genomes. *Sci Adv*. **8**, eabg3842 (2022).

21. J. Bast, D. J. Parker, Z. Dumas, K. M. Jalvingh, P. Tran Van, K. S. Jaron, E. Figuet, A. Brandt, N. Galtier, T. Schwander, Consequences of asexuality in natural populations: Insights from stick insects. *Mol. Biol. Evol.* **35**, 1668–1677 (2018).

22. A. Wagner, "Gene duplications and innovation" in *The origins of evolutionary innovations: a theory of transformative change in living systems* (Oxford University Press, 2011).

23. J. L. Payne, A. Wagner, The causes of evolvability and their evolution. *Nat. Rev. Genet.* (2018).

24. L. Boto, Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc. Biol. Sci.* **281**, 20132450 (2014).

25. K. S. Jaron, J. Bast, R. W. Nowell, T. R. Ranallo-Benavidez, M. Robinson-Rechavi, T. Schwander, Genomic features of parthenogenetic animals. *J. Hered.* **112**, 19–33 (2021).

26. J. Bast, K. S. Jaron, D. Schuseil, D. Roze, T. Schwander, Asexual reproduction reduces transposable element load in experimental yeast populations. *Elife*. **8**, e48548 (2019).

27. E. B. Chuong, N. C. Elde, C. Feschotte, Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017).

28. D. W. Zeh, J. A. Zeh, Y. Ishida, Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays*. **31**, 715–726 (2009).

29. D. Kempson, M. Lloyd, R. Ghelardi, A new extractor for woodland litter. *Pedobiologia*. **3**, 1–21 (1963).

30. G. Weigmann, *Hornmilben (Oribatida). Die Tierwelt Deutschlands* (Goecke & Evers, Keltern, 2006).

31. H. Öztoprak, J. Bast, Modified salting out method for high molecular weight *g*DNA extraction (oribatid mites). *protocols.io* (2023)

32. D. Mapleson, G. Garcia Accinelli, G. Kettleborough, J. Wright, B. J. Clavijo, KAT: a *K*-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics*. **33**, 574–576 (2017).

33. T. R. Ranallo-Benavidez, K. S. Jaron, M. C. Schatz, GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).

34. H. Cheng, G. T. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat. Methods*. **18**, 170–175 (2021).

35. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).

36. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. **25**, 1754–1760 (2009).

37. L. Baudry, N. Guiglielmoni, H. Marie-Nelly, A. Cormier, M. Marbouty, K. Avia, Y. L. Mie, O. Godfroy, L. Sterck, J. M. Cock, C. Zimmer, S. M. Coelho, R. Koszul, instaGRAAL: chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. *Genome Biol.* **21**, 148 (2020).

38. M. Xu, L. Guo, S. Gu, O. Wang, R. Zhang, B. A. Peters, G. Fan, X. Liu, X. Xu, L. Deng, Y. Zhang, TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Gigascience*. **9** (2020).

39. R. Kundu, J. Casey, W.-K. Sung, HyPo: Super fast & accurate polisher for long read genome assemblies. *bioRxiv* (2019).

40. D. Guan, S. A. McCarthy, J. Wood, K. Howe, Y. Wang, R. Durbin, Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*. **36**, 2896–2898 (2020).

41. M. Alonge, L. Lebeigle, M. Kirsche, K. Jenike, S. Ou, S. Aganezov, X. Wang, Z. B. Lippman, M. C. Schatz, S. Soyk, Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258 (2022).

42. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. **31**, 3210–3212 (2015).

43. R. Challis, E. Richards, J. Rajan, G. Cochrane, M. Blaxter, BlobToolKit - interactive quality assessment of genome assemblies. *G3*. **10**, 1361–1374 (2020).

44. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience*. **10** (2021).

45. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

46. S. Ou, W. Su, Y. Liao, K. Chougule, J. R. A. Agda, A. J. Hellinga, C. S. B. Lugo, T. A. Elliott, D. Ware, T. Peterson, N. Jiang, C. N. Hirsch, M. B. Hufford, Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 (2019).

47. E. A. Bell, C. L. Butler, C. Oliveira, S. Marburger, L. Yant, M. I. Taylor, Transposable element annotation in non-model species: The benefits of species-specific repeat libraries using semi-automated EDTA and DeepTE *de novo* pipelines. *Mol. Ecol. Resour.* **22**, 823–833 (2022).

48. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. **29**, 15–21 (2013).

49. T. Wicker, F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, A. H. Schulman, A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973–982 (2007).

50. S. J. Teresi, M. B. Teresi, P. P. Edger, TE Density: a tool to investigate the biology of transposable elements. *Mob. DNA*. **13**, 11 (2022).

51. V. Bandi, C. Gutwin, J. N. Siri, E. Neufeld, A. Sharpe, I. Parkin, Visualization tools for genomic conservation. *Methods Mol. Biol.* **2443**, 285–308 (2022).

52. B. D. Ondov, T. J. Treangen, P. Melsted, A. B. Mallonee, N. H. Bergman, S. Koren, A. M. Phillippy, Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **17**, 132 (2016).

53. K. Domes, M. Maraun, S. Scheu, S. L. Cameron, The complete mitochondrial genome of the sexual oribatid mite *Steganacarus magnus*: genome rearrangements and loss of tRNAs. *BMC Genomics*. **9**, 532 (2008).

54. R. Allio, A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, F. Delsuc, MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Mol. Ecol. Resour.* **20**, 892–905 (2020).

55. J. Frank, P. Joern, M. Bernt, A. Donath, M. Middendorf, Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Access*, 1–13 (2011).

56. M. Bernt, A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsch, J. Pütz, M. Middendorf, P. F. Stadler, MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **69**, 313–319 (2013).

57. A. Donath, F. Jühling, M. Al-Arab, S. H. Bernhart, F. Reinhardt, P. F. Stadler, M. Middendorf, M. Bernt, Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Res.* **47**, 10543–10552 (2019).

58. D. Laslett, B. Canbäck, ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*. **24**, 172–175 (2008).

59. P. P. Chan, T. M. Lowe, tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* **1962**, 1–14 (2019).

60. R. Lorenz, S. H. Bernhart, C. Höner Zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, I. L. Hofacker, ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).

61. R. W. Ness, S. A. Kraemer, N. Colegrave, P. D. Keightley, Direct estimate of the spontaneous mutation rate uncovers the effects of drift and recombination in the *Chlamydomonas reinhardtii* plastid genome. *Mol. Biol. Evol.* **33**, 800–808 (2016).

62. L. A. Bergeron, S. Besenbacher, T. Turner, C. J. Versoza, R. J. Wang, A. L. Price, E. Armstrong, M. Riera, J. Carlson, H.-Y. Chen, M. W. Hahn, K. Harris, A. S. Kleppe, E. H. López-Nandam, P. Moorjani, S. P. Pfeifer, G. P. Tiley, A. D. Yoder, G. Zhang, M. H. Schierup, The Mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *Elife*. **11** (2022).

63. P. D. Keightley, A. Pinharanda, R. W. Ness, F. Simpson, K. K. Dasmahapatra, J. Mallet, J. W. Davey, C. D. Jiggins, Estimation of the spontaneous mutation rate in *Heliconius melpomene*. *Mol. Biol. Evol.* **32**, 239–243 (2015).

64. B. Bushnell, BBMap - sourceforge.net/projects/bbmap/ (2015).

65. F. Cabanettes, C. Klopp, D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*. **6**, e4958 (2018).

66. M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A. Drummond, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. **28**, 1647–1649 (2012).

67. M. J. Hubisz, K. S. Pollard, A. Siepel, PHAST and RPHAST: phylogenetic analysis with space/time models. *Brief. Bioinform.* **12**, 41–51 (2011).

68. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

69. R. W. Nowell, P. Almeida, C. G. Wilson, T. P. Smith, D. Fontaneto, A. Crisp, G. Micklem, A. Tunnacliffe, C. Boschetti, T. G. Barraclough, Comparative genomics of bdelloid rotifers: Insights from desiccating and nondesiccating species. *PLoS Biol.* **16**, e2004830 (2018).

70. E. A. Gladyshev, M. Meselson, I. R. Arkhipova, Massive horizontal gene transfer in bdelloid rotifers. *Science*. **320**, 1210–1213 (2008).

71. D. M. Emms, S. Kelly, OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).

72. Z. Zhang, J. Xiao, J. Wu, H. Zhang, G. Liu, X. Wang, L. Dai, ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* **419**, 779–781 (2012).

73. J. Feng, C. A. Meyer, Q. Wang, J. S. Liu, X. Shirley Liu, Y. Zhang, GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics*. **28**, 2782–2788 (2012).

74. W. He, J. Yang, Y. Jing, L. Xu, K. Yu, X. Fang, NGenomeSyn: an easy-to-use and flexible tool for publication-ready visualization of syntenic relationships across multiple genomes. *Bioinformatics*. **39** (2023).

75. E. E. Hare, J. S. Johnston, "Genome size determination using flow cytometry of propidium iodide-stained nuclei" in *Molecular Methods for Evolutionary Genetics*, V. Orgogozo, M. V. Rockman, Eds. (Humana Press, Totowa, NJ, 2011), vol. 772 of *Methods in Molecular Biology*, pp. 3–12.

76. J.-F. Flot, B. Hespeels, X. Li, B. Noel, I. Arkhipova, E. G. J. Danchin, A. Hejnol, B. Henrissat, R. Koszul, J.-M. Aury, V. Barbe, R.-M. Barthélémy, J. Bast, G. a. Bazykin, O. Chabrol, A. Couloux, M. Da Rocha, C. Da Silva, E. Gladyshev, P. Gouret, O. Hallatschek, B. Hecox-Lea, K. Labadie, B. Lejeune, O. Piskurek, J. Poulain, F. Rodriguez, J. F. Ryan, O. a. Vakhrusheva, E. Wajnberg, B. Wirth, I. Yushenova, M. Kellis, A. S. Kondrashov, D. B. Mark Welch, P. Pontarotti, J. Weissenbach, P. Wincker, O. Jaillon, K. Van Doninck, Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature*. **500**, 453–457 (2013).

References Chapter 3

Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B., & Delsuc, F. (2020). MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Molecular Ecology Resources*, *20*(4), 892–905.

Armstrong, K. F., & Ball, S. L. (2005). DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1462), 1813–1823.

Arroyo, J., & Iturrondobeitia, J. C. (2007). *Abnormalities in some species of oribatid mites (Acari, Oribatida) from Burgos (NW Spain)*.

Bandi, V., Gutwin, C., Siri, J. N., Neufeld, E., Sharpe, A., & Parkin, I. (2022). Visualization Tools for Genomic Conservation. *Methods in Molecular Biology*, *2443*, 285–308.

Barraclough, T. G. (2010). Evolving entities: towards a unified framework for understanding diversity at the species and higher levels. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1547), 1801–1813.

Barraclough, T. G., Birky, C. W., Jr, & Burt, A. (2003). Diversification in sexual and asexual organisms. *Evolution; International Journal of Organic Evolution*, *57*(9), 2166–2172.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., & Stadler, P. F. (2013). MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319.

Birky, C. W., Jr, Adams, J., Gemmel, M., & Perry, J. (2010). Using population genetic theory and DNA sequences for species detection and identification in asexual organisms. *PloS One*, *5*(5), e10609.

Birky, C. W., Jr, & Barraclough, T. G. (2009). Asexual speciation. *Lost Sex*, 201–216.

Birky, C. W., Wolf, C., Maughan, H., Herbertson, L., & Henry, E. (2005). Speciation and Selection without Sex. *Hydrobiologia*, *546*(1), 29–45.

Brandt, A., Van, P. T., Bluhm, C., Anselmetti, Y., Dumas, Z., Figuet, E., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., Labédan, M., Maraun, M., Parker, D. J., Robinson-Rechavi, M., Schaefer, I., Simion, P., Scheu, S., Schwander, T., & Bast, J. (2021). Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(38).

Chan, P. P., Lin, B. Y., Mak, A. J., & Lowe, T. M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Research*, *49*(16), 9077–9096.

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, *18*(2), 170–175.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886.

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *45*(4), e18.

Dierckxsens, N., Mardulyn, P., & Smits, G. (2020). Unraveling heteroplasmy patterns with NOVOPlasty. *NAR Genomics and Bioinformatics*, *2*(1), lqz011.

Domes-Wehner, K. (2009). *Parthenogenesis and Sexuality in Oribatid Mites: phylogeny, mitochondrial genome structure and resource dependence* [Phd, Technische Universität]. http://tuprints.ulb.tu-darmstadt.de/1383/

Donath, A., Jühling, F., Al-Arab, M., Bernhart, S. H., Reinhardt, F., Stadler, P. F., Middendorf, M., & Bernt, M. (2019). Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Research*, *47*(20), 10543–10552.

Fontaneto, D., Herniou, E. A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., & Barraclough, T. G. (2007). Independently evolving species in asexual bdelloid rotifers. *PLoS Biology*, *5*(4), e87.

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, *27*(2), 221–224.

Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., & Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, *36*(9), 2896–2898.

Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., & Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics: TIG*, *23*(4), 167–172.

Havird, J. C., Hall, M. D., & Dowling, D. K. (2015). The evolution of sex: A new hypothesis based on mitochondrial mutational erosion: Mitochondrial mutational erosion in ancestral eukaryotes would favor the evolution of sex, harnessing nuclear recombination to optimize compensatory nuclear coadaptation. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *37*(9), 951–958.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings. Biological Sciences / The Royal Society*, *270*(1512), 313–321.

Heethoff, M., Domes, K., Laumann, M., Maraun, M., Norton, R. A., & Scheu, S. (2007). High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *Journal of Evolutionary Biology*, *20*(1), 392–402.

Hey, J. (2006). On the failure of modern species concepts. *Trends in Ecology & Evolution*, *21*(8), 447–450.

Judson, O. P., & Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution*, *11*(2), 41–46.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), 540–546.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676.

Lienhard, A., & Krisper, G. (2021). Hidden biodiversity in microarthropods (Acari, Oribatida, Eremaeoidea, Caleremaeus). *Scientific Reports*, *11*(1), 23123.

Li, H. (2021). New strategies to improve minimap2 alignment accuracy. *Bioinformatics*, *37*(23), 4572–4574.

Li, W.-N., & Xue, X.-F. (2019). Mitochondrial genome reorganization provides insights into the relationship between oribatid mites and astigmatid mites (Acari: Sarcoptiformes: Oribatida). *Zoological Journal of the Linnean Society*, *187*(3), 585–598.

Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., & Clavijo, B. J. (2017). KAT: a *K*-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics*, *33*(4), 574–576.

Marie Curie SPECIATION Network: Butlin, R., Debelle, A., Kerth, C., Snook, R. R., Beukeboom, L. W., Castillo Cajas, R. F., Diao, W., Maan, M. E., Paolucci, S., Weissing, F. J., van de Zande, L., Hoikkala, A., Geuverink, E., Jennings, J., Kankare, M., Knott, K. E., Tyukmaeva, V. I., Zoumadakis, C., Schilthuizen, M. (2012). What do we need to know about speciation? *Trends in Ecology & Evolution*, *27*(1), 27–39.

Martens, K., Rossetti, G., Baltanas, A., & Others. (1998). Reproductive modes and taxonomy. *Sex and Parthenogenesis. Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods*, 197–211.

Maruyama, S., Isozaki, Y., Kimura, G., & Terabayashi, M. (1997). Paleogeographic maps of the Japanese Islands: Plate tectonic synthesis from 750 Ma to the present. *Island Arc*, *6*(1), 121–142.

Öztoprak, H., Gao, S., Guiglielmoni, N., Brandt, A., Zheng, Y., Becker, C., Becker, K., Bednarski, V., Borgschulte, L., Burak, K. A., Dion-Côté, A.-M., Leonov, V., Opherden, L., Shimano, S., & Bast, J. (2023). Haplotype independence contributes to evolvability in the long-term absence of sex in a mite. In *bioRxiv* (2023.09.07.556471).

Pfingstl, T., Baumann, J., & Lienhard, A. (2019). The Caribbean enigma: the presence of unusual cryptic diversity in intertidal mites (Arachnida, Acari, Oribatida). *Organisms, Diversity & Evolution*, *19*(4), 609–623.

Rhyker Ranallo-Benavidez, T., Jaron, K. S., & Schatz, M. C. (2020). GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications*, *11*(1), 1–10.

Rosenberger, M. (2022). *Phylogeography in sexual and parthenogenetic European oribatida* [University Goettingen Repository].

Schäffer, S., Kerschbaumer, M., & Koblmüller, S. (2019). Multiple new species: Cryptic diversity in the widespread mite species Cymbaeremaeus cymba (Oribatida, Cymbaeremaeidae). *Molecular*

Phylogenetics and Evolution, *135*, 185–192.
Schön, I., Martens, K., & van Dijk, P. (2009). Lost sex. *The Evolutionary Biology of Parthenogenesis*, 1–615.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212.

Simpson, G. G. (1951). The Species Concept. *Evolution; International Journal of Organic Evolution*, *5*(4), 285–298.

Simpson, G. G. (1961). *Principles of Animal Taxonomy*. Columbia University Press.
Systematics and the origin of species, from the viewpoint of a zoologist.

van der Kooi, C. J., Matthey-Doret, C., & Schwander, T. (2017). Evolution and comparative ecology of parthenogenesis in haplodiploid arthropods. *Evolution Letters*, *1*(6), 304–316.

van Hinsbergen, D. J. J., Torsvik, T. H., Schmid, S. M., Maţenco, L. C., Maffione, M., Vissers, R. L. M., Gürer, D., & Spakman, W. (2020). Orogenic architecture of the Mediterranean region and kinematic reconstruction of its tectonic evolution since the Triassic. *Gondwana Research*, *81*, 79–229.

von Saltzwedel, H., Maraun, M., Scheu, S., & Schaefer, I. (2014). Evidence for frozen-niche variation in a cosmopolitan parthenogenetic soil mite species (Acari, Oribatida). *PloS One*, *9*(11), e113268.

Wiley, E. O. (1981). Remarks on Willis' Species Concept. *Systematic Biology*, *30*(1), 86–87.

Wulsch S., Öztoprak H., Guiglielmoni N., Jeffries D.L., Bast J. (2023). A female heterogametic ZW sex-determination system in Acariformes. In *bioRxiv* (2023.10.24.563255).

Zhan, X.-B., Chen, B., Fang, Y., Dong, F.-Y., Fang, W.-X., Luo, Q., Chu, L.-M., Feng, R., Wang, Y., Su, X., Fang, Y., Xu, J.-Y., Zuo, Z.-T., Xia, X.-Q., Yu, J.-G., & Sun, E.-T. (2021). Mitochondrial analysis of oribatid mites provides insights into their atypical tRNA annotation, genome rearrangement and evolution. *Parasites & Vectors*, *14*(1), 221.

References Chapter 4

Ackerman MS, Maruki T, Lynch M. 2017. MAPGD: a program for the maximum likelihood analysis of population data.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R, et al. 2014. Sex determination: why so many ways of doing it? PLoS Biol. 12:e1001899.

Baudry L, Guiglielmoni N, Marie-Nelly H, Cormier A, Marbouty M, Avia K, Mie YL, Godfroy O, Sterck L, Cock JM, et al. 2020. instaGRAAL: chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. Genome Biol. 21:148.

Bell EA, Butler CL, Oliveira C, Marburger S, Yant L, Taylor MI. 2022. Transposable element annotation in non-model species: The benefits of species-specific repeat libraries using semi-automated EDTA and DeepTE *de novo* pipelines. Mol. Ecol. Resour. 22:823–833.

Bell G. 1982. The Masterpiece of Nature: The Evolution and Genetics of Sexuality. Los Angeles: Univ of California Press

Bergero R, Charlesworth D. 2009. The evolution of restricted recombination in sex chromosomes. Trends Ecol. Evol. 24:94–102.

Beukeboom LW, Perrin N. 2014. The Evolution of Sex Determination.

Brandt A, Tran Van P, Bluhm C, Anselmetti Y, Dumas Z, Figuet E, François CM, Galtier N, Heimburger B, Jaron KS, et al. 2021. Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. Proc. Natl. Acad. Sci. U. S. A. 118.

Bull JJ. 1984. Sex Determination: Evolution of Sex Determining Mechanisms. Science 224:733–734.

Chakraborty R, Fuerst PA, Nei M. 1980. Statistical Studies on Protein Polymorphism in Natural Populations. III. Distribution of Allele Frequencies and the Number of Alleles per Locus. Genetics 94:1039–1063.

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. 2020. BlobToolKit – Interactive Quality Assessment of Genome Assemblies. G3 Genes|Genomes|Genetics 10:1361–1374.

Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity 95:118–128.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. Nat. Methods 18:170–175.

Consortium TT of S, Ashman T-L, Bachtrog D, Blackmon H, Goldberg EE, Hahn MW, Kirkpatrick M, Kitano J, Mank JE, Mayrose I, et al. 2014. Data from: Tree of sex: a database of sexual systems.: 1–8.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10.

Domes K, Scheu S, Maraun M. 2007. Resources and sex: Soil re-colonization by sexual and parthenogenetic oribatid mites. Pedobiologia 51:1–11.

Ellinghaus D, Kurtz S, Willhoeft U. 2008. LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. BMC Bioinformatics 9:18.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. U. S. A. 117:9451–9457.

Galaxy Community. 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 50:W345–W351.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644–652.

Grandjean F. 1941. Statistique sexuelle et parthenogenese chez les Oribates (Acariens). C. R. Seanc. Ac. Sci. 212:463–467.

Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R. 2020. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics 36:2896–2898.

Guirao-Rico S, González J. 2021. Benchmarking the performance of Pool-seq SNP callers using simulated and real sequencing data. Mol. Ecol. Resour. 21:1216–1229.

Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK Jr, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, et al. 2003. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. 31:5654–5666.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. 9:R7.

Heethoff M, Bergmann P, Norton RA. 2006. Karyology and sex determination of oribatid mites. Acarologia 46:127–131.

Ironside JE. 2010. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. Bioessays 32:718–726.

Jay P, Tezenas E, Véber A, Giraud T. 2022. Sheltering of deleterious mutations explains the stepwise extension of recombination suppression on sex chromosomes and other supergenes. PLoS Biol. 20:e3001698.

Jeffries DL, Lavanchy G, Sermier R, Sredl MJ, Miura I, Borzée A, Barrow LN, Canestrelli D, Crochet P-A, Dufresnes C, et al. 2018. A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. Nat. Commun. 9:4088.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37:907–915.

Kokot M, Dlugosz M, Deorowicz S. 2017. KMC 3: counting and manipulating *k*-mer statistics. Bioinformatics 33:2759–2761.

Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, Hulselmans G, Clamons S. 2023. FelixKrueger/TrimGalore: v0. 6.10-add default decompression path (0.6. 10). Zenodo.

Kundu R, Casey J, Sung W-K. 2019. HyPo: Super Fast & Accurate Polisher for Long Read Genome Assemblies. bioRxiv: 2019.12.19.882506.

Liana M, Witaliński W. 2012. Female and male reproductive systems in the oribatid mite *Hermannia gibba* (Koch, 1839) (Oribatida: Desmonomata). Int. J. Acarology 38:648–663.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv.

Li H. 2021. New strategies to improve minimap2 alignment accuracy. Bioinformatics 37:4572–4574.

Lóšková J, Ľuptáčik P, Miklisová D, Kováč Ľ. 2013. Community structure of soil oribatida (acari) two years after windthrow in the high tatra mountains. Biologia 68:932–940.

Lozano-Fernandez J, Tanner AR, Giacomelli M, Carton R, Vinther J, Edgecombe GD, Pisani D. 2019. Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. Nat. Commun. 10:2295.

Lukhtanov VA, Dincă V, Friberg M, Šíchová J, Olofsson M, Vila R, Marec F, Wiklund C. 2018. Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. Proceedings of the National Academy of Sciences 115:E9610–E9619.

Lynch M, Bost D, Wilson S, Maruki T, Harrison S. 2014. Population-genetic inference from pooled-sequencing data. Genome Biol. Evol. 6:1210–1218.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol. Biol. Evol. 38:4647–4654.

Mapleson D, Garcia Accinelli G, Kettleborough G, Wright J, Clavijo BJ. 2016. KAT: a *K*-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics 33:574–576.

Maraun M, Scheu S. 2000. The structure of oribatid mite communities (Acari, Oribatida): patterns, mechanisms and implications for future research. Ecography 23:374–382.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10–12.

Matthey-Doret C, Baudry L, Bignaud A, Cournac A. 2020. hicstuff: Simple library/pipeline to generate and handle Hi-C data. Zenodo.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303.

Norton RA, Kethley JB, Johnston DE, Oconnor BM. 1993. Phylogenetic perspectives on genetic systems and reproductive modes of mites. In: Wrensch DL, Ebbert MA, editors. Evolution and Diversity of Sex Ratio in Insects and Mites. NY: Chapman & Hall.

Oliver JH. 1983. Chromosomes, Genetic Variance and Reproductive Strategies Among Mites and Ticks. Bull. Entomol. Soc. Am. 29:8–17.

Ou S, Jiang N. 2018. LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. Plant Physiol. 176:1410–1422.

Öztoprak H, Gao S, Guiglielmoni N, Brandt A, Zheng Y, Becker C, Becker K, Bednarski V, Borgschulte L, Burak KA, Dion-Côté AM, Leonov V, Opherden L, Shimano S, Bast J. 2023. Haplotype independence contributes to evolvability in the long-term absence of sex in a mite. bioRxiv: 2023.09.07.556471.

Öztoprak H, Bast J. Modified salting out method for high molecular weight gDNA extraction (oribatid mites). protocols.io (2023)

Palmer DH, Rogers TF, Dean R, Wright AE. 2019. How to identify sex chromosomes and their turnover. Mol. Ecol. 28:4709–4724.

Palmer SC, Norton RA. 1991. Taxonomic, geographic and seasonal distribution of thelytokous parthenogenesis in the Desmonomata (Acari: Oribatida). Exp. Appl. Acarol. 12:67–81.

Picard Toolkit. 2019. Broad Institute, GitHub Repository. https://broadinstitute.github.io/picard/; Broad Institute

Ponnikas S, Sigeman H, Abbott JK, Hansson B. 2018. Why Do Sex Chromosomes Stop Recombining? Trends Genet. 34:492–503.

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 42:W187–W191.

Ranallo-Benavidez TR, Jaron KS, Schatz MC. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat. Commun. 11:1–10.

Schaefer I, Norton RA, Scheu S, Maraun M. 2010. Arthropod colonization of land - Linking molecules and fossils in oribatid mites (Acari, Oribatida). Mol. Phylogenet. Evol. 57:113–121.

Shi J, Liang C. 2019. Generic Repeat Finder: A High-Sensitivity Tool for Genome-Wide *De Novo* Repeat Detection. Plant Physiol. 180:1803–1815.

Shumate A, Wong B, Pertea G, Pertea M. 2022. Improved transcriptome assembly using a hybrid of long and short reads with StringTie. PLoS Comput. Biol. 18:e1009730.

Smelansky IE. 2006. Some Population Characteristics of Oribatid Mites in Steppe Habitats. Acarina.

Smrž J. 2010. Nutritional biology of oribatid mites from different microhabitats in the forest. In: Trends in Acarology. Springer Netherlands. p. 213–216.

Sokolov II. 1954. The chromosome complex of mites and its importance for systematics and phylogeny. Trud. Leningr. Obshch. Estestvois. 72:124–159.

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics 24:637–644.

Su W, Gu X, Peterson T. 2019. TIR-Learner, a New Ensemble Method for TIR Transposable Element Annotation, Provides Evidence for Abundant New Transposable Elements in the Maize Genome. Mol. Plant 12:447–460.

Subías LS, L.S., 2020. Listado sistemático, sinonímico y biogeográfco de los Ácaros Oribátidos (Acariformes: Oribatida) del mundo (excepto fósiles), 15a actualización.

Svedberg J, Hosseini S, Chen J, Vogan AA, Mozgova I, Hennig L, Manitchotpisit P, Abusharekh A, Hammond TM, Lascoux M, et al. 2018. Convergent evolution of complex genomic rearrangements in two fungal meiotic drive elements. Nat. Commun. 9:1–13.

Takehana Y, Matsuda M, Myosho T, Suster ML, Kawakami K, Shin-I T, Kohara Y, Kuroki Y, Toyoda A, Fujiyama A, et al. 2014. Co-option of Sox3 as the male-determining factor on the Y chromosome in the fish *Oryzias dancena*. Nat. Commun. 5:4157.

Uller T, Pen I, Wapstra E, Beukeboom LW, Komdeur J. 2007. The evolution of sex ratios and sex-determining systems. Trends Ecol. Evol. 22:292–297.

Webb NR, Gw E. 1979. Variations between Populations of *Steganacarus magnus* (Acari; Cryptostigmata) In Great Britain.

Wehner K, Scheu S, Maraun M. 2014. Resource availability as driving factor of the reproductive mode in soil microarthropods (Acari, Oribatida). PLoS One 9:e104243.

Wehner K, Schuster R, Simons NK, Norton RA, Blüthgen N, Heethoff M. 2021. How land-use intensity affects sexual and parthenogenetic oribatid mites in temperate forests and grasslands in Germany. Exp. Appl. Acarol. 83:343–373.

Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, Ebler J, Fungtammasan A, Kolesnikov A, Olson ND, et al. 2019. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat. Biotechnol. 37:1155–1162.

Wrensch DL, Kethley JB, Norton RA. 1994. Cytogenetics of Holokinetic Chromosomes and Inverted Meiosis: Keys to the Evolutionary Success of Mites, with Generalizations on Eukaryotes. In: Mites. Springer US. p. 282–343.

Wright AE, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. Nat. Commun. 7:1–8.

Xiong W, He L, Lai J, Dooner HK, Du C. 2014. HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. Proc. Natl. Acad. Sci. U. S. A. 111:10263–10268.

Xu M, Guo L, Gu S, Wang O, Zhang R, Peters BA, Fan G, Liu X, Xu X, Deng L, et al. 2020. TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. Gigascience 9.

Xu Z, Wang H. 2007. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 35:W265–W268.

Zhou Q, Zhang J, Bachtrog D, An N, Huang Q, Jarvis ED, Gilbert MTP, Zhang G. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. Science 346:1246338.

References Chapter 5

Ahram, R., Fergus, E., & Noguera, P. (2011). Addressing Racial/Ethnic Disproportionality in Special Education: Case Studies of Suburban School Districts. *Teachers College Record*, *113*(10), 2233–2266.

Bernard, R. E., & Cooperdock, E. H. G. (2018). No progress on diversity in 40 years. *Nature Geoscience*, *11*(5), 292–295.

Forrester, N. (2020, September 23). *Diversity in science: next steps for research group leaders*. Nature Publishing Group UK.

Judd, K., & McKinnon, M. (2021). A systematic map of inclusion, equity and diversity in science communication research: Do we practice what we preach? *Frontiers in Communication*, *6*.

Krempkow, R. (2017). *Herausforderung chancengerechte Bildung – Von der Grundschule bis zur Promotion. In: Internationalisierung, Vielfalt und Inklusion in der Wissenschaft*.

Nasr, N. (2021). Overcoming the discourse of science mistrust: how science education can be used to develop competent consumers and communicators of science information. *Cultural Studies of Science Education*, *16*(2), 345–356.

Pfeffer, F. T. (2018). Growing Wealth Gaps in Education. *Demography*, *55*(3), 1033–1068.

Science benefits from diversity. (2018). *Nature*, *558*(7708), 5.

Temple, J. A., Ou, S.-R., & Reynolds, A. J. (2022). Closing Achievement Gaps Through Preschool-To-Third-Grade Programs. *Frontiers in Education*, *7*.

Appendix

Supplementary Materials Chapter 2

Materials and Methods

Sample preparation and sequencing

Sample collection. All mite individuals were sampled from natural populations in Germany, Italy, Canada, Japan and Russia. German samples were collected in a coniferous forest in Dahlem (50.39010 N, 6.57162 E) in a 30 cm² square. Italian samples were collected at Montan Southern Tyrol (46.33514 N, 11.29728 E) in a 30 cm² square. Japanese soil and leaf litter was sampled near a larch forest in Yamanashi (35.5449 N 138.2403 E) and sent to Germany for extraction. Canadian samples were collected from 46.03570 N, -64.79896 E and 46.14482 N, -64.76913 E. Specimens of interest were sent to Germany. Russian samples were isolated from a spruce forest in Moscow Oblast (56.02522 N, 38.43074 E). Specimens were isolated out of leaf litter by heat gradient extraction (*29*). *Platynothrus peltifer* was identified morphologically using (*30*) and molecularly confirmed by cytochrome oxidase I (CO1) sequencing.

Sample preparation. Specimens were starved for more than a week and cleansed with a brush in distilled water, distilled water with detergent (fit GmbH, Zittau, Germany), and incubated in NaClO 0.05% (DonKlorix; CP GABA GmbH, Hamburg, Germany) and ethanol 70% for 30 seconds each and rinsed in distilled water again.

High-molecular-weight *g*DNA extraction (for ultra-low input). For high-molecular-weight (HMW) single-individual DNA extraction, we established a modified salting out protocol (*31*). In short: A single individual was submerged in TNES buffer and flash-frozen in liquid nitrogen. The sample was then homogenized using a sterile pestle. After adding Proteinase K, the sample was incubated for at least 1h. Next, yeast tRNA was added, followed by NaCl and 96% ethanol. DNA purification was conducted and the sample was left to homogenize overnight. DNA concentration was measured using Qubit Fluorometer v. 4 with the Qubit dsDNA HS Assay kit (Thermo Fisher Scientific, Waltham, MA).

Cytochrome oxidase I sequencing. The cytochrome oxidase subunit 1 (CO1; ~700 bp) region was amplified using the universal primers LC01490F/HC02198R (Folmer et al., 1994). The PCR reaction was run with 1 ng of DNA, 2x Thermo Scientific DreamTaq Green PCR Master Mix, 1 µM forward and 1 µM reverse primer, and ddH20 to fill until 25 µl.

Amplification was conducted under the following conditions: denaturation at 95°C for 5 min, 35 cycles at 95°C for 30 s, 45°C for 30 s and 72°C for 1 min, and final extension at 72°C for 6 min. PCR products were purified by adding 1 U/ml of Exonuclease and 0.3 U/ml FastAP to 8 µl PCR product then heated for 30 min at 37°C, and subsequently for 20 min at 85°C. For sequencing, the Big dye Terminator Cycle sequencing Kit and an ABI PRISM automatic sequencer were used at the Cologne Center for Genomics (CCG; Cologne, Germany). To molecularly verify the samples, sequences of Cytochrome oxidase I with >99% similarity to *P. peltifer* were identified as such.

DNA quality. HMW *g*DNA sample was assessed at the Genomics & Transcriptomics Laboratory (GTL; Düsseldorf, Germany) using the Agilent Femto Pulse system. HMW *g*DNA with an OD260/280 ratio of approximately 1.8 to 2.0 and fragment size above 15 kb was selected for sequencing.

Long-read and linked-read sequencing for reference assembly. Single-individual HMW DNA was sequenced using two technologies: PacBio HiFi (SMRTbell® Libraries from Ultra-Low DNA Input) and TELL-seq (TELL-seq^{TM} WGS library)TM both with ultra-low DNA input of at least 5 ng. Library preparation and sequencing of the SMRTbellTM templates were conducted by the GTL. PacBio HiFi sequencing yielded 34.7 Gb of reads with an N50 of 15 kb. In addition, TELL-seq libraries were constructed and sequenced by CCG using a TELL-Seq WGS Library Prep Kit (Universal Sequencing Technology, Carlsbad, CA) and yielded 69 Gb of reads.

Omni-C sequencing. To construct the library for Omni-C adult individuals were sampled from two spots in the coniferous forest in Dahlem, Germany in close vicinity. Previous analyses on cytochrome oxidase I from these two spots suggest two distinct mitochondrial lineages to be present in both of these spots. 226 whole adult specimens were flash-frozen with liquid N_2, crushed, vortexed, and the Omni-C Proximity Ligation Assay protocol for insects & marine invertebrates was followed. The library was sequenced on Illumina NovaSeq 6000 by Novogene (Cambridge, United Kingdom), which generated 92.7x10^6 pairs of 150-bp reads.

Linked-read sequencing for populations. In addition to the reference individual, four individuals of the same population were sequenced using Illumina TELL-seq linked paired-end reads only. For details see supplementary data.

RNA sequencing. Total RNA was extracted from ten adult individuals using TRIzol reagent treated with DNase I within a Direct-zolTM RNA MicroPrep kit (Zymo Research, Irvine, CA). An RNA-seq

library was constructed by the CCG using TruSeq Stranded Total RNA with Ribo-Zero Globin and $43x10^6$ pairs of 100-bp reads were sequenced.

***k*-mer analyses.** 27-mers in the HiFi reads were analyzed using KAT v2.4.2 (*32*) with the modules kat hist and kat gcp (default parameters). Ploidy was further investigated using kmc v3.2.1 with parameters -k 27 -ci 1 -cs 10000 and Smudgeplot (v0.2.5) (*33*) with default parameters.

Chromosome-level collapsed and phased assemblies

***De novo* genome assembly.** HiFi reads were assembled using hifiasm (v0.16.1-r375) (*34*) with default parameters into initial collapsed haploid contigs, and using Flye (v2.9) (*35*) with default parameters into phased contigs.

TELL-seq scaffolding. Barcoded TELL-seq linked Illumina reads were generated and corrected from BCL raw data using tell-read (v1.0.2). The barcodes were formatted to 10X Genomics standard by using ust10x (v1.0.2) from TELL-seq conversion_tool. The collapsed haploid contigs (Hap0: 211 Mb, N=33) were scaffolded using Scaff10X (v4.2) (github.com/wtsi-hpag/Scaff10X) with formatted TELL-seq reads and using arguments -longread 1 -gap 100 -matrix 2000 -reads 10 -score 10 -edge 50000 -link 8 -block 50000. Haplotype-resolved contigs were scaffolded using Scaff10X with parameters -longread 1 -gap 100 -matrix 2000 -reads 10 -score 10 -edge 50000 -link 8 -block 50000.

Omni-C scaffolding. Omni-C reads were cleaned using TrimGalore (v0.6.5) (github.com/FelixKrueger/TrimGalore) with parameters '-j 30 -q 30 --fastqc --paired' and mapped to the draft haploid assembly using hicstuff (v3.1.1) and bwa (v0.7.15) (*36*) with parameters --enzyme 100 --iterative --aligner bwa. instaGRAAL (v0.1.6 no-opengl branch) (*37*) was run with parameters --level 5 --cycles 100. The scaffolds were curated with instaGRAAL-polish to reduce misassemblies and add 10 Ns in gaps. 99.75% of the assembly were anchored to nine chromosome-level scaffolds which were selected as chromosome candidates for subsequent analyses.

Gap filling and polishing. HiFi reads were mapped to the haploid assembly using minimap2 with parameters '--secondary=no --MD -ax asm20' and then sorted and indexed using Samtools (v1.11) with default parameters. Gaps in the collapsed scaffolds Hap0 were filled using TGS-GapCloser (v1.1.1) (*38*) with the parameters '-tgstype pb --minmap_arg -x asm20 --ne' and the final haploid (Hap0) assembly was polished using HyPo (v1.0.3) (*39*) with the PacBio HiFi reads.

Haplotype scaffolding. Primary and alternative haplotypes were separated using minimap2 and purge_dups (v1.2.5) (*40*), and are subsequently designated as haplotypic blocks A and B. Haplotypes were reciprocally scaffolded using RagTag (v2.1.0) (*41*) with the other haplotype as reference.

Anchoring of haplotypes. Haplotypic blocks A were mapped against the collapsed assembly (Hap0) using minimap2 (v2.24-r1122) with parameter -x asm5. Haplotypic blocks B were mapped against haplotypic blocks A. Haplotype-resolved scaffolds were attributed to chromosome candidates of Hap0 for which they had the higher number of residue matches.

Assembly evaluation

Completeness. For the chromosome-level assembly (Hap0), ortholog completeness was assessed using the tool Benchmarking Universal Single-Copy Orthologs (BUSCO v.5.0.0) (*42*) against the Arthropoda odb10 lineage (1,066 orthologs) and the Arachnida odb10 lineage (2,934 orthologs). *k*-mer completeness was evaluated using KAT (v2.4.2) and the module kat comp with default parameters.

Omni-C contact map. Omni-C reads were mapped to Hap0 using bwa and hicstuff as previously described. The contact map was generated using the module hicstuff view with the parameter -b 500.

Contaminants. PacBio HiFi reads were mapped to the final Hap0 scaffolds using minimap2 (v2.24-r1122) with parameters -ax map-hifi and the mapped reads were sorted with SAMtools (v1.11) (*43*, *44*). Hap0 was aligned against the nucleotide database using the Basic Local Alignment Search Tool (BLAST v2.6.0) (*45*) with parameters -outfmt "6 qseqid staxids bitscore std sscinames scomnames" -max_hsps 1 -evalue 1e-25. The outputs of minimap2, BLAST, and BUSCO (against the Arachnida odb10 lineage) were provided as input to Blobtools2 (*43*).

Genome annotation

Repeat and transposable element annotation and masking. A repeat library including transposable elements (TEs) was built using EDTA (v1.9.6) (*46*) with parameters '--sensitive 1 --anno 1'. TEs were annotated and classified to the superfamily level using the FasTE pipeline (*47*). The hardmasked assemblies were converted to softmasked assemblies using bedtools (v2.26.0) with mask mode.

RNA-seq mapping. Adapter sequences were removed from RNA-seq raw reads using TrimGalore (v0.6.5) with parameters '-j 30 -q 30 --fastqc --paired'. Trimmed reads were mapped to the softmasked reference haploid assembly Hap0 and the two haplotypic blocks assemblies haplotypic blocks A and haplotypic blocks B using STAR (v2.5.1a) (*48*) with parameters '--readFilesCommand zcat --

outSAMtype BAM SortedByCoordinate --outSAMstrandField intronMotif --outFilterIntronMotifs RemoveNoncanonical'.

Gene Structure Prediction. Trinity (v2.1.1) was employed to assemble the RNA-Seq reads *de novo*, and the resulting 61,067 transcripts were aligned to the assemblies using PASA (v2.5.2). Mapped reads were provided to StringTie (v2.2.0) to predict transcripts. Then TransDecoder (https://github.com/TransDecoder/TransDecoder/wiki; Last accessed November 14, 2018) was used to find the Open Reading Frames (ORF) for each gene. BRAKER2 (v2.1.6) was used to predict gene structures with RNA-seq evidence. These results were provided to EVidenceModeler (v1.1.1) with weight parameters 'ABINITIO_PREDICTION AUGUSTUS 4; TRANSCRIPT assembler-database.sqlite 7; OTHER_PREDICTION transdecoder 8' to select genes. Finally, PASA was used to refine UTR regions. The final annotation was converted into protein sequences using gffread (v0.12.1).

Functional annotation. We utilized Eggnog-mapper (v2.1.6) with protein sequences and generated KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway annotations. Furthermore, we employed InterProScan (version 5.61-93.0) with protein sequences to generate Reactome pathway annotations. The Gene Ontology (GO) annotation was created by combining the results of Eggnog-mapper (v2.1.6) and InterProScan (version 5.61-93.0).

Genome dynamics

TE abundance and landscapes. Quality-cleaned genome-wide repeat abundances were generated by FasTE. To generate the TE divergence landscape a custom R script was created to rename the TE superfamilies after Wicker-classification (*49*) and adjust the data to plot the TE divergence Landscape with the 'Plot Kimura Distance' R script for the Hap0 and the largest haplotypic blocks per chromosome. For more information on the size of regions see table S2.

TE Density. The TE Density pipeline (*50*) was applied to generate the TE Density plots, which are defined as TE-occupied base pairs in a given window, on the collapsed assembly of *P. peltifer*.

Gene synteny. Protein sequences of Hap0 and the haplotypic blocks were aligned using BLAST v2.6.0 with parameters -evalue 1e-10 -outfmt 6. In order to find blocks of gene synteny, MCScanX (commit 97e74f4) was run between Hap0, haplotypic blocks A and haplotypic blocks B with default parameters. Collinearity was visualized with SynVisio (*51*) with a minimum match score >= 3950.895. Heterozygosity between the haplotypes was estimated using Mash (v2.3) (*52*).

Mitochondrial genome assembly and annotation

TELL-seq reads were trimmed using TrimGalore and assembled using MitoFinder v1.4.1 with the mitochondrial genome of *Steganacarus magnus* as seed (*53*, *54*). The assembly was annotated using MiTFi v0.1 (*55*) and then using MITOS (*56*), MITOS2 (*57*)(both available at http://mitos2.bioinf.uni-leipzig.de/index.py), ARWEN v1.2.3 (*58*) and tRNAscan-SE v2.0 (*59*) to find genes and tRNAs that could not be found by MiTFi. tRNAscan-SE was run with COVE cutoff set to -20. Geneious (Biomatters Ltd.) was used to manually curate annotations and find consensus predictions. tRNAs were selected based on their predicted secondary structure and their minimum free energy, computed using the RNAfold web server (http://rna.tbi.univie.ac.at/) (*60*).

Spontaneous mutation rate estimation

Sample preparation. For spontaneous *de novo* mutation rate estimation without rearings and mutation accumulation lines, best practice was followed by sequencing parents and offspring (as in (*61*, *62*)). *Platynothrus peltifer* individuals were sampled from the same population as the reference genome individual (Dahlem, Germany) in early June 2021. For three individual mothers (M1, M2, and M3) their 'offspring daughters', i.e. eggs, were extracted after removing the genital plates. Eggs were cleansed in NaClO 0.05% (DonKlorix; CP GABA GmbH, Hamburg, Germany) with a brush to remove tissue from their mothers. For M1 one daughter (D1), for M2 one daughter (D2), and for M3 three daughters (D3, D4, and D5) could successfully be analyzed.

Sequencing. To avoid PCR-induced read errors and to get highly accurate reads, the 'NEBNext Ultra II FS DNA Library Prep Kit for Illumina' (New England Biolabs, Ipswich, USA) in combination with UMI Adaptor 'NEBNext Multiplex Oligos for Illumina (Unique Dual Index UMI Adaptors DNA Set 1)' (New England Biolabs, Ipswich, USA) was chosen. These libraries were multiplexed and whole-genome sequenced on an Illumina NovaSeq 600 generating paired-end 150 bp reads. This yielded a total of 811.6×10^6 reads with a mean of 102.1×10^6 for the mothers (min 100.3×10^6, max 103.0×10^6) and a mean of 100.9×10^6 (min 81.2×10^6, max 111.0×10^6) for the eggs.

Identifying *de novo* mutations (DNM). Reads were demultiplexed using UMI-tools extract (v.1.1.2). Adapters were trimmed using TrimGalore (v0.6.5) with parameters -q 30 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -a2 AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT --fastqc --paired. Reads were mapped against the hardmasked reference genome of *P. peltifer* using the Burrows-Wheeler Alignment tool *bwa mem* (v0.7.17-r1198-dirty) and filtered with the Genome Analysis ToolKit (GATK v4.1.9). Duplications were removed with UMI-tools (v 1.1.2). Potential variant sites were called and filtered using GATK

(v4.1.9) following an in-house pipeline (all scripts and parameters are outlined in detail at the github). In short, variants were filtered by read depth, phred-scaled likelihood of the genotype, allele depth and homozygous to heterozygous mutations. Additionally, following common practice (*62*, *63*) only candidate sites that were homozygous in mothers and heterozygous in daughters were considered as possible *de novo* mutations. Identified mutation site candidates were manually checked in Integrative Genomics Viewer (IGV v2.8.13) under the following conditions: mothers were filtered for a major allele frequency of 0.95 and the daughters for a minor allele frequency of 0.2. Alternatively, a major allele frequency of 0.9 for the mothers and a minor allele frequency of 0.3 for the daughters also passed the filter. Candidate sites found in all sibling daughters were removed as false positives resulting from erroneous genotype assignment in the mother (see Supplementary Data Table 3).

Spontaneous mutation rate. To calculate the mutation rate per generation, the number of sites for which a new variant could have been detected, i.e. the 'callable genome' was identified. The callable genome includes all homozygous sites in the 3 mothers, similarly filtered as the DNMs. All positions with the variant sites are included in the previously called GVCF-file by GATKs HaplotypeCaller. Each mother's GVCF-file was filtered in read depth (DP), genotype (GT), phred-scaled likelihood of the genotype (PL), and allele depth (AD): DP was filtered like above, sites with DP between 28 to 124 were accepted; for GT, all homozygous sites from the mother were selected; PL for sites must have fulfilled: $PL_{second\ most\ likely} - PL_{first\ most\ likely} < 120$; only sites where AD supported one allele were kept. For details see github.

Population genetic analyses

Variant identification. Single nucleotide polymorphisms (SNPs) were used to investigate population dynamics within populations of asexual *P. peltifer* from Germany, Russia, Italy, Canada and Japan. Phased population data were generated by mapping trimmed raw-reads generated by TELL-read (v1.0.3) using *bwa mem* (v0.7.15) with default parameters to the collapsed soft masked haploid reference genome (Hap0). The resulting alignment was sorted using samtools (v1.11) and duplications were removed using Picard MarkDuplicates (v2.26.2 Broad Institute). Coverage was calculated with samtools flagstat. Following, variants were called using the Genome Analysis ToolKit (GATK v.4.1.9.0) pipeline. GVCFs were produced using HaplotypeCaller and then merged using CombineGVCFs. Variants were detected with GenotypeGVCFs. SNPs were selected with SelectVariants. PLINK (v 1.9) and VCFtools (0.1.15) were used to perform PCA analysis.

Population data coverage filter. To compare population statistic metrics among the different populations (DE, RU, IT, JP, CA), it is required that the respective reference genome sites are matched,

i.e. aligned. Some genomic regions may have only one haplotype being successfully sequenced or mapped, which may lead to artifactual underestimation of heterozygosity and false signals of homozygosity. This becomes apparent with regions mapping with half coverage of the overall median mapped coverage. To mitigate this issue, for each individual the site-wise coverage was filtered, so that only sites that have 75% or higher of the genomic median coverage (for that individual) are used for the population statistics. There is high variation of coverage distribution among individuals (median ranging from 28 in IT3 to 124 in IT1, though most are between 70 and 100) which makes this individual-based coverage filtering necessary.

Basic population statistics. The German reference genome is divided into contiguous 1 Mb blocks for each chromosome, with a total of 220 blocks. Within each block, we calculated the heterozygosity for each individual (see github for scripts used). Because calculation of $\theta\pi$ or θw requires using only sites that passed the coverage filter for every individual and thus the amount of data is too small and potentially biased, we used an alternative method to quantify deviation from sexual equilibrium: the 'unshared-to-shared ratio' by comparing two individuals from the same population at a time, rx = (H_u1+H_u2)/(2*H_sh) where H_u1 is the number of sites that are heterozygous only in the first individual, H_u2 heterozygous only in the second individual, and H_sh heterozygous in both. For each population, all ten possible pairs are used and the mean value is used. This ratio can show if the heterozygosity and site frequency distribution is more similar to a sexual one (closer to 2) or a clonal one (closer to 0), and is comparable in interpretation to the heterozygosity-to-$\theta\pi$ ratio (which is 1 for a sexual equilibrium population and 2 for a long-term clonal one).

Empirical gene conversion estimation. To compare empirical data with the simulated gene conversion data (see simulations), we tallied the distribution of 'homozygosity stretches', i.e., the distance between two consecutive heterozygous sites, from each individual in the German sample. To minimize the effect of mapping artifacts, only stretches of 1000bp or shorter are counted, and the stretch is required to have its mean coverage above the filter threshold. The distribution is calculated separately for each individual and is compared with simulated data under different gene conversion track lengths.

Haplotype-specific analyses

Parallel divergence of haplotypes. Quality trimmed paired-end resequencing reads were mapped against haplotypic block assemblies haplotypic blocks A and haplotypic blocks B simultaneously (competitive mapping) and split according to which haplotype they mapped best to using bbsplit (bbmap v38.63; (*64*)). Reads with ~> 5 noncontiguous substitutions were discarded (minratio=0.9). Reads that mapped to both haplotypic block assemblies were kept (ambiguous2=split) and merged with each the

set of split reads mapping best to haplotypic blocks A and to the set of split reads mapping best to haplotypic blocks B per individual. This was done to avoid biasing the analysis towards regions that are phaseable, and hence highly heterozygous, in most populations and individuals.

The split sets of resequencing reads (haplotypic blocks A+ambiguous reads and haplotypic blocks B+ambiguous reads) were mapped to the collapsed, softmasked genome assembly Hap0 using bwa v0.7.17 and sorted using samtools v 1.15.1 (as described above). Optical and sequencing duplicates were removed using picard MarkDuplicates v2.26.2 (Broad Institute). Variants were called for each split set of resequencing reads, i.e. each haplotype, separately using gatk v4.2.6.1 HaplotypeCaller, CombineGVCFs and GenotypeGVCFs as described above. Indels, multiallelic sites, sites with a quality < 20, genotypes with a depth below 10 and heterozygous genotypes were removed using bcftools. Heterozygous genotypes could result from sequencing error, paralogs and incomplete phasing: either due to the absence of a region from one of the two haplotypic block assemblies such that reads from both haplotypes map to the haplotypic block assembly that is present, or due to low divergence of haplotypes yielding ambiguous mapping results.

Variants were applied to the collapsed assembly using bcftools consensus thereby generating one consensus genome per haplotype per individual. Missing sites and genotypes were applied as Ns leaving a 50 sequence whole haplome alignment (2 haplotypes * 5 individuals * 5 populations). To determine regions that were in-phase, i.e. no phase switching possible within a region, scaffolds of haplotypic blocks assembly haplotypic blocks A were aligned to the collapsed assembly using D-Genies and standard parameters (see Table S3) (*65*). Per chromosome the longest alignment block was extracted from the haplome alignment using geneious (*66*) and subdivided into 1 Mio Mb long bins using msa_split (*67*). Best fitting ML trees were reconstructed for each bin using IQ-tree (version 1.6.1) (*68*) with implemented model testing and 1000 bootstrap replicates. A majority rule consensus tree (i.e. comprises clades that are present in at least 50% of trees) was constructed using geneious (see Figure 2; Figure S9).

Horizontal gene transfer (HGT). To detect horizontally transferred genes from different organisms, a previously developed pipeline (https://github.com/reubwn/hgt) (*69*) was utilized, which applies several lines of evidence for detecting HGT candidates (HGTc). In short, DIAMOND (v0.9.21.122) BLASTP was used in 'sensitive' mode to align the protein sequences from the reference and two haplotypic block assemblies to the UniRef90 database. Then taxids information from the NCBI Taxonomy database and three measures were employed to identify putative candidate HGTs.

The three measures including the (a) HGT Index (hU) using the best-hit bitscores from the BLASTP results and defined '(best-hit bitscore for OUTGROUP) - (best-hit bitscore for INGROUP)'; (b) Alien Index (AI) based on E-values also from the BLASTP results to evaluate HGT defined by 'log10 ((best-

hit E-value for INGROUP) + 1e-200) - log10((best-hit E-value for OUTGROUP) + 1e-200)'(*70*) and then (c) Consensus Hit Support (CHS) using the sum of bitscores across all hits, and the HGTc must be designated as outgroup and the designation of rest hits >=90%. IQ-tree (version 1.6.1) was used to plot gene phylogenies.

Orphan HGTs. To detect HGTc that are only found on one of the two haplotypic blocks ('orphan HGTs'), the protein sequence of the HGTs from one haplotypic block was searched in the other haplotypic block respectively. The potential orphan HGTs were additionally analyzed for coverage, as true missing HGTs should exhibit substantially lower coverage as only one copy (instead of two) and single copy genes from BUSCO results are sequenced. Mapped reads count was divided by gene length for normalization.
Additionally, the orphan HGT candidates were checked for syntenic neighboring gene content.

Variation calling and annotation. Using PacBio HiFi reads from the reference individual, we mapped them to the Hap0 genome using minimap2 (version 2.24-r1122). Subsequently, the resulting BAM file was sorted using Samtools (version 1.11). The identical pipeline as previously employed was then applied and the SNP annotated by ANNOVAR (version 2020-06-07 23:56:37 -0400).

Orthology and Ka/Ks. The protein sequences of annotated genes from the haplotypic blocks and collapsed assemblies were used for detecting the single-copy genes between haplotypes. Among these three, pairwise comparisons were made by OrthoFinder (v 2.5.4) (*71*) to detect single-copy orthogroups genes with default parameters on the haplotypic blocks and collapsed assemblies. The 1:1 orthologous genes between haplotypic blocks A and haplotypic blocks B were additionally filtered by the 1:1 orthology results from the haplotypic blocks A to Hap0 and haplotypic blocks B to Hap0. Then the ParaAT (v 2.0) (*72*) integrated KaKs_Calculator (v 1.2) was employed to calculate the Ka/Ks value for the haplotypic orthologous single-copy genes pairs.

Differential expression of alleles (DEA). The cleaned RNA-Seq reads used for annotation were mapped to the two softmasked haplotypic block assemblies haplotypic blocks A and haplotypic blocks B using STAR (v2.5.1a) with parameters '--readFilesCommand zcat --outSAMtype BAM SortedByCoordinate --outSAMstrandField intronMotif --outFilterIntronMotifs RemoveNoncanonical --outFilterMismatchNmax 3 --outFilterMismatchNoverLmax 0.1 --outFilterMismatchNoverReadLmax 0.5 '.
Differential expression of alleles (DEA) between the two haplotypic blocks haplotypic blocks A and haplotypic blocks B were performed with GFOLD V1.1.4 (*73*). The GFOLD value ≥ | 0.2 | and the

RPKM>1 of all samples were set as cutoff values (The GFOLD value could be considered as a reliable log2 fold change). All identified DEAs were further submitted for functional enrichment with R package clusterProfiler (v4.6.2), using all RPKM>1 single-copy genes as the background with the 'enricher' function.

Wolbachia

Previously published *Wolbachia* genomes on NCBI were utilized ASM1658442v1 as a query. To identify candidate genes integrated into the mite genome, we employed the DIAMOND (v. 0.9.21) Blastp tool. Genes exhibiting an identity of over 40% were deemed potential candidates for integration into the mite genome. The synteny was plotted by using NGenomeSyn (v1.41) (*74*).

Simulations

Single-lineage-based long-term simulation. Considering a lack of exchange of genetic material between individuals in a cloning species, long-termed evolution can be efficiently simulated by tracing the sequences of a pair of haplotypes (i.e., an individual) over time. To determine the long-term effects of cloning and gene conversion on heterozygosity, we simulated a genomic block as a sequence of 5 million sites. Each site can be in a state of homozygosity (represented by value 0) or heterozygosity (represented by value 1).

At the beginning (generation 0), all 5 million sites are set to 0. Mutations occur with a rate of $2\times\mu$ events per site per generation. When it occurs, the site is set to 1 regardless of the original state (in other words, back-mutations from heterozygotes to homozygotes are considered negligible). Gene conversion occurs with a rate of $2\times r_{GC}$ events per site per generation. When it occurs, L_{GC} consecutive sites (starting from a random site) are set to 0, also regardless of the original state. The "2×" in both mutations and gene conversions represents the fact that both types of events can initiate in one or the other haplotype. The mutation rate is set to $\mu = 2.05\times10^{-9}$ (as empirically estimated, see above). The rate of gene conversion is set up in such a way that $r_{GC}\times E(L_{GC}) = \mu/0.015$, which would lead to an equilibrium heterozygosity of 1.5% (to match the average empirical divergence between haplotype of the populations). We set the mean gene conversion track length, $E(L_{GC})$ as a variable with values 50, 100, 200, 500, 1000, 2000 and 5000; and the L_{GC} for each gene conversion event is either constant or geometrically distributed. For each of the 7×2 parameter combinations, 20 replicates are run. Each replicate runs for 5×10^7 generations. To speed up the computation, every 100 generations are completed in the same loop, with the mutations occurring first and gene conversion second. Every 5000 generations, the number of heterozygous sites in the 5 Mb sequence is recorded. From $2\sim5\times10^7$ generations, we also record the distribution of homozygous stretch lengths (numbers of consecutive homozygotes) every 1×10^5 generations.

Genome size estimation

Flow Cytometry. To estimate the genome size of *P. peltifer*, a flow cytometry was utilized following the method of (*75*), supported by personal communication with J. Spencer Johnston. *Drosophila melanogaster* (1 C = 175 Mb) was used as the reference. In short, animals were frozen at -80°C and cut on dry ice. The head of *D. melanogaster* was removed and ground with 15 strokes (1 per second) of a B pestle in 10 µL of Galbraith buffer in a 1 mL Dounce homogenizer, after which 990 µL buffer were added and homogenized. For the mite, the anterior portion of the proterosoma was cut off (as mites do not have distinct 'heads'), typically containing blood cells, muscle and neural cells that will yield 2C nuclei. This part of the mite was similarly ground in 10µL Galbraith buffer, after which 440 µL budder was added. Following, 50 µL of the *D. melanogaster* mix was added to the mite cells and homogenized. The mix containing sample and reference was filtered through a 45-µm nylon mesh and 20 µL of Propidium iodide (50 µg/mL) was added homogenized. The sample was then incubated at 4°C for 2 h prior to running with the Accuri C6 Flow Cytometer (BD Biosciences, Erembodegem, Belgium). To compare fluorescence signals, the sample and reference were run both separately and together to check for 2C and 4C peak signals. The flow cytometry was run slowly and stopped after a minimum of 1000 counts in the 2C peaks. Software gates were used for cleaning and filtering the data. The data was deemed good quality if the peaks were separated and the 4C had the double mean FL3-A value as the 2C peak. Several replicates were run and the best quality peak chosen as result.

Supplementary Text

Genome size estimation with flow cytometry

The relative genome size of *Platynothrus peltifer* was estimated to be 232 Mb. For *P. peltifer* the mean FL3-A for the 2C peak was 395 and for the 4C peak 789. For *D. melanogaster* 298 and 596 respectively. The genome size estimate of flow cytometry matches somewhat closely the cleaned, chromosome-scale genome assembly size (6% difference: 232 Mb vs 219 Mb).

Note that flow cytometric estimates for mites are typically associated with some minor uncertainties, as there is no clear head given and only 'few' cells can be extracted and stained. Additionally, *D. melanogaster* genome sizes are variable to some degree among individuals and strains. Following, the genome assembly is a very suitable representation of the full *P. peltifer* genome.

Mutation rate estimations

In total, two *de novo* spontaneous mutations were identified in daughters compared to mothers based on the filtered and carefully curated variants. These two mutations were found at chromosome 2 in position 5257052 and at chromosome 4 in position 19592709 in D2.

To determine the overall mutation rate, the number of *de novo* mutations is divided by twice (as the species is diploid) the sum of all callable sites from replicate daughters taken together.

$$\mu = \frac{Number\ of\ de\ novo\ mutations}{(2\times\ callable\ genome)} \rightarrow \mu = \frac{2}{(2\times 488{,}839{,}965)} = 2.05 * 10 - 9$$

The estimated spontaneous mutation rate is thus 2.05×10^{-9} per generation for *P. peltifer* and is thus within the range of *Daphnia* and *Drosophila*.

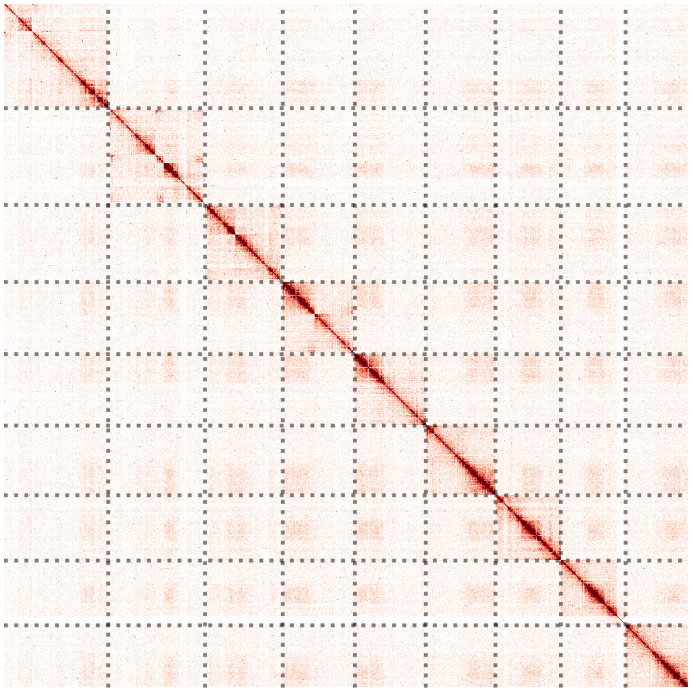
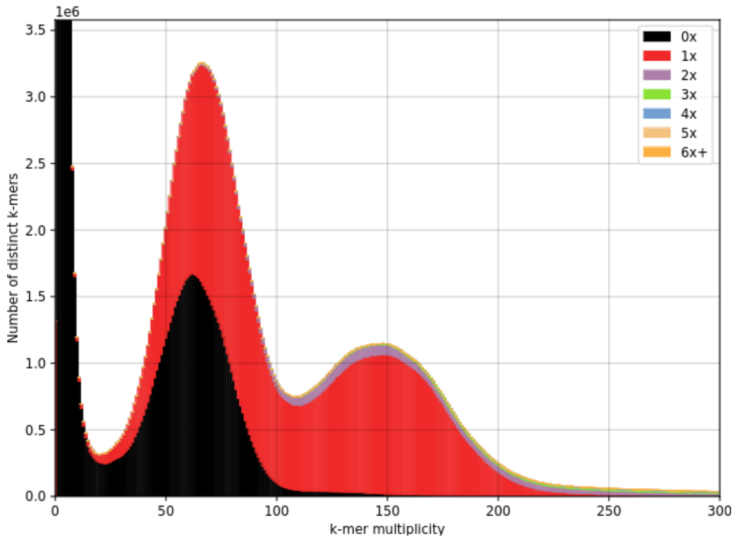
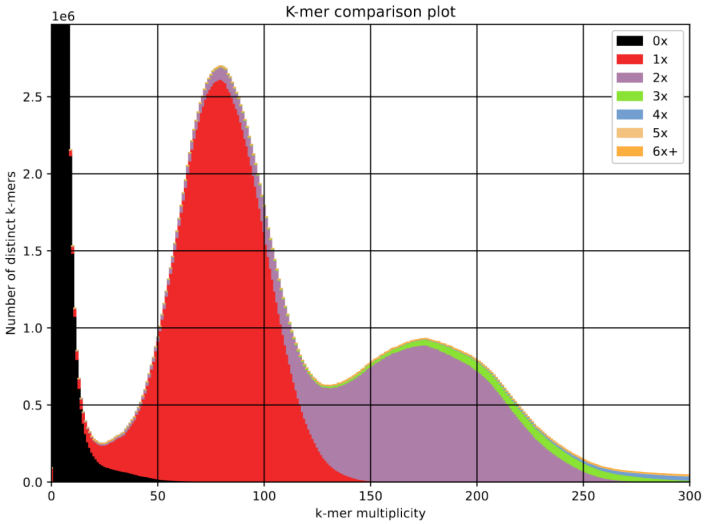


Fig. S1 Hi-C contact map of *Platynothrus peltifer* with a binning of 500 representing 9 chromosome-level scaffolds (chromosome 1 to 9 from left to right, top to bottom). The contact map shows strong intrachromosomal interaction frequencies and no structural errors.



(a) Collapsed assembly Hap0.



(b) Phased assembly.

Fig. S2. *k*-mer comparison of the assemblies of *Platynothrus peltifer* against PacBio HiFi reads. The plots show two peaks, the first one for heterozygous *k*-mers, and the second one for homozygous *k*-mers. In both assemblies, low-multiplicity *k*-mers are not included (black, 0X). **a)** Final chromosome-level collapsed assembly. Homozygous *k*-mers are represented exactly once (red). As the assembly is collapsed, only one version of each heterozygous region can be included in the assembly, thus part of heterozygous *k*-mers are represented once (red) and part are absent from the assembly (black). **b)** Phased assembly including haplotype blocks A and B. Homozygous *k*-mers are represented twice (purple) for both haplotypes and heterozygous *k*-mers are all represented once (red).

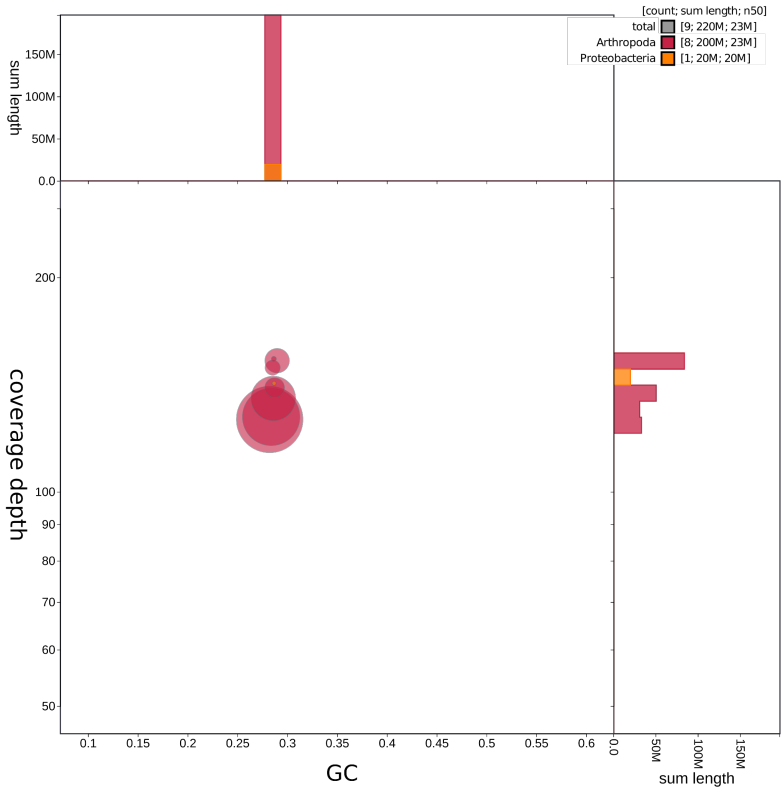
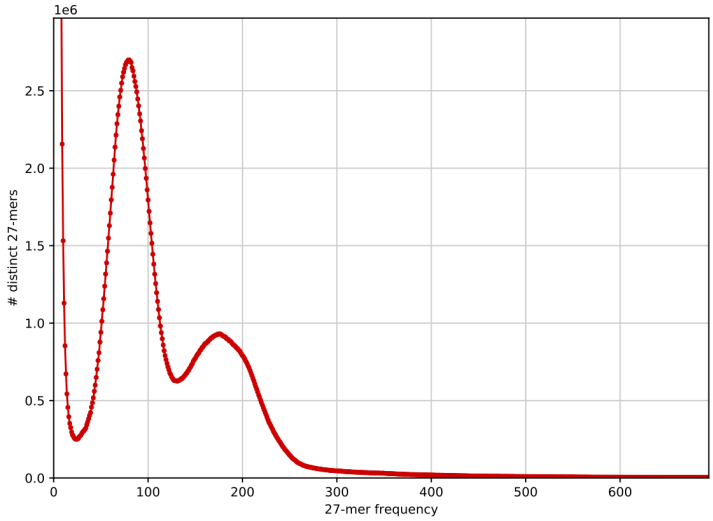
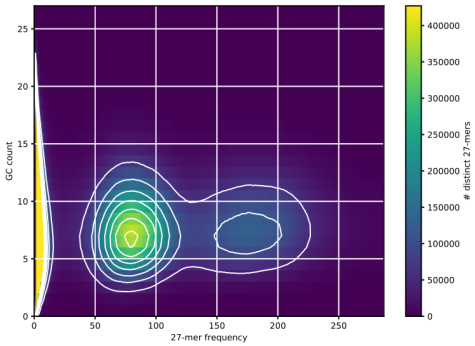


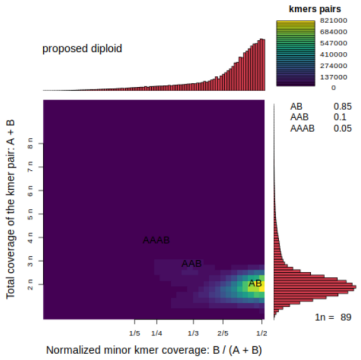
Fig. S3. The final assembly (Hap0 scaffolds) contains no contaminants as seen in the blob plot analysis. The nine chromosome candidates have similar GC contents of 0.28, and their coverage depth (based on mapping of the HiFi reads) ranges from 126X to 153X. Eight of these scaffolds were flagged as Arthropoda, and chr8 was flagged as Proteobacteria due to the integration of *Wolbachia* sequences in the genome.



(a) 27-mer histogram.



(b) 27-mer GC content analysis.



(c) Smudgeplot

Fig. S4. Genome property analyses based on *k*-mers of the PacBio HiFi reads (*k* = 27). **a)** In the *k*-mer histogram, two peaks are visible at 80X and 160X corresponding to the heterozygous *k*-mers and the homozygous *k*-mers respectively, which suggests a diploid, heterozygous genome. **b)** The two peaks identified in histogram (a) have similar GC content, and there is no extra peak at a higher GC content, which indicates that there is no significant contamination in the reads. **c)** The Smudgeplot shows a distinct smudge of *k*-mers with an A/B configuration and the genome is identified as diploid without duplications.

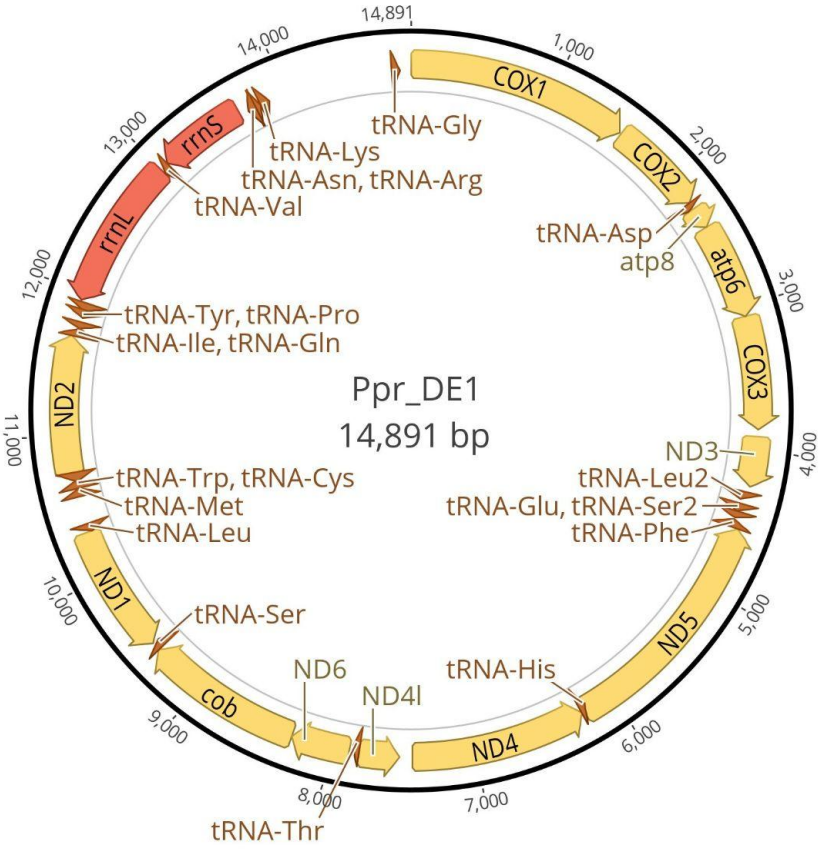


Fig. S5. Annotated mitochondrion of *Platynothrus peltifer*, stemming from the same individual as the reference genome. The size is 14,891 bp and all typical mitochondrial genes could be identified.

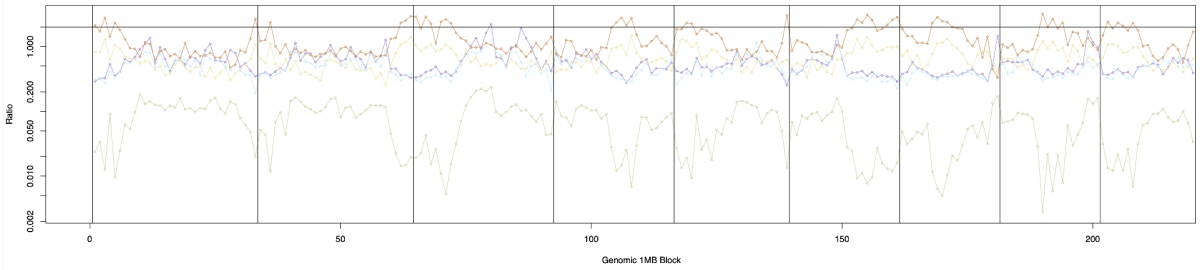


Fig. S6. The ratio of unshared to shared heterozygotic sites from pairwise comparisons, averaged for each population, and separated into 1 Mb genomic blocks. The expected value of this ratio is 2 (black line) in a sexual, random mating population, and trends towards 0 under long-term clonal reproduction. In a sexual population, the expected value of Rx = 2. I.e., for each pair of individuals, the ratio should be approximately 2 Het-Hom, 2 Hom-Het, and 1 Het-Het sites. In a clonal asexual population, because the two haplotypes diverge, any site that is either i) het at the beginning, or ii) fixed in one haplotype but not found in the other, will be shared het sites. Unshared het sites only occur when there are newer mutations (in one haplotype) that are not yet fixed. With time, the number of shared het sites will increase indefinitely, but unshared will stay in an equilibrium (new mutations vs fixation/loss within haplotype). Given enough time, Rx will become smaller and smaller. It cannot reach 0 (because there are always new mutations not fixed yet) but should be much smaller than 2.

The proportion of shared heterozygous sites among individuals of each of the populations is much larger for the Italian population, followed by the German, Russian and Japanese, with the Canadian population sharing only few heterozygous sites (fig. S6). As individual heterozygosity values of the European-Russian populations are comparable, the differences likely reflect population size variation, as in smaller populations the likelihood to sample closely related clones of the same lineage increases. The Canadian and Japanese populations have the lowest individual heterozygosity and also fewer shared heterozygous variants, which might suggest more recent transition events to asexuality, much lower mutation rates and/or substantially larger population sizes compared to European-Russian populations.

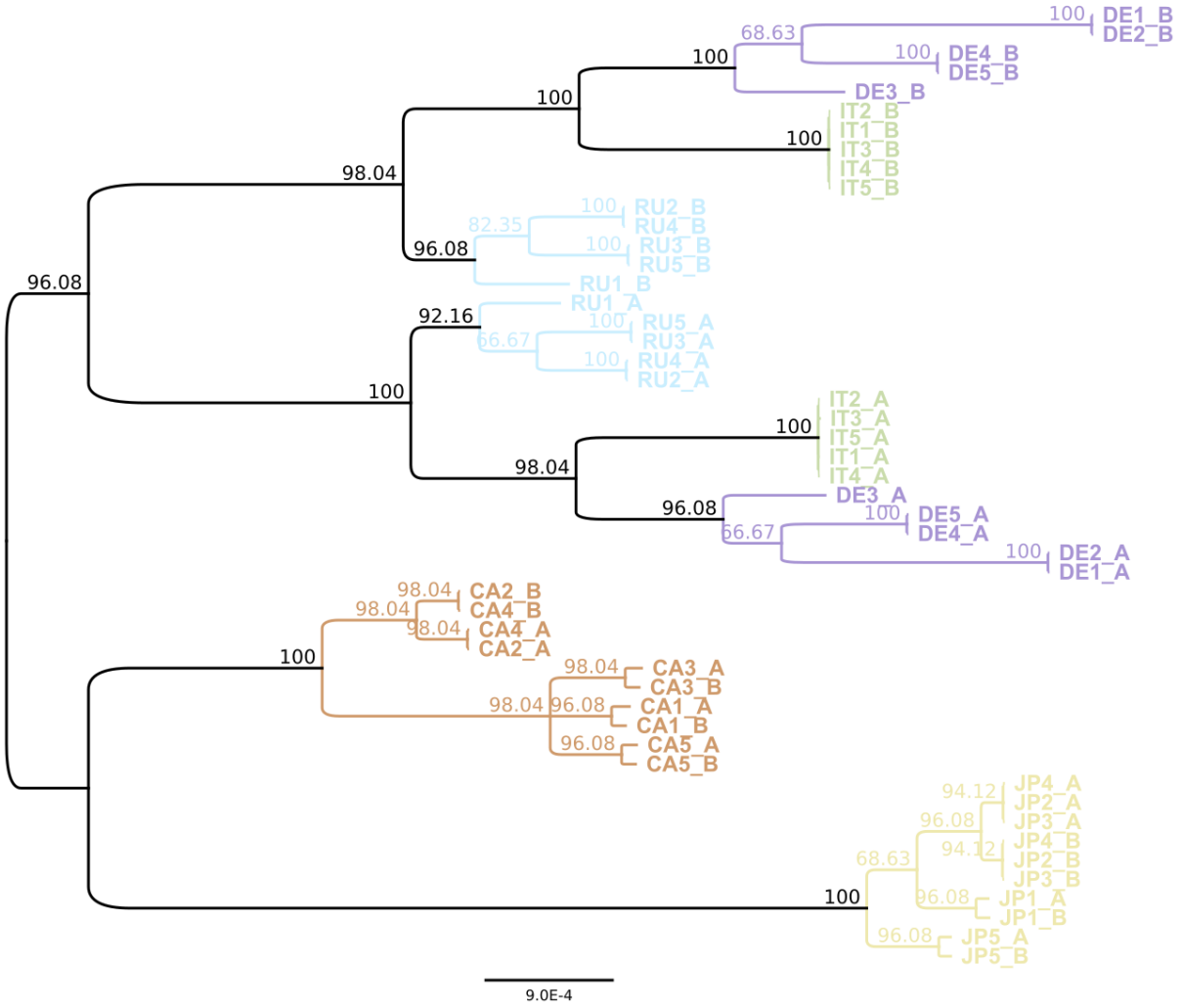


Fig. S7. The haplotype tree illustrates the presence of the Meselson effect in three populations of *Platynothrus peltifer* (Germany - DE, Italy - IT, Russia - RU). In addition, haplotype divergence exceeds divergence among individuals for three individuals from Japan (JA) and two from Canada (CA). The tree is a majority rule consensus tree (i.e. comprises clades that are present in at least 50% of trees) constructed from 51 trees. Each of the 51 trees was reconstructed from a 1 Mb bin derived from the longest alignment block of the in-phase regions per chromosome. For readability the tree was rooted manually at the clade comprising the populations DE, IT, RU. Branch labels represent consensus support values (%; e.g. a value of 96.08 indicates that a clade is present in 96.08% (49) of 51 trees). Branch length indicates divergence (average number of substitutions per site). Colors indicate different populations; _A & _B: haplotypes A and B.

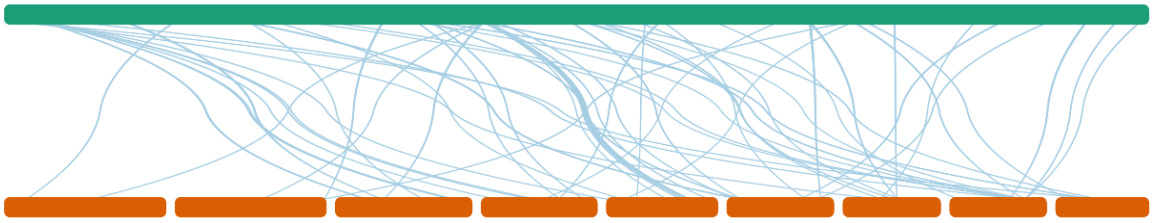


Fig. S8. Integration of *Wolbachia* remnants into the *Platynothrus peltifer* genome. Depending on the stringency of parameters, portions of *Wolbachia* can be detected. Synteny analyses suggest a *Wolbachia* genome (in green) copies throughout the *P. peltifer* genome (in red).

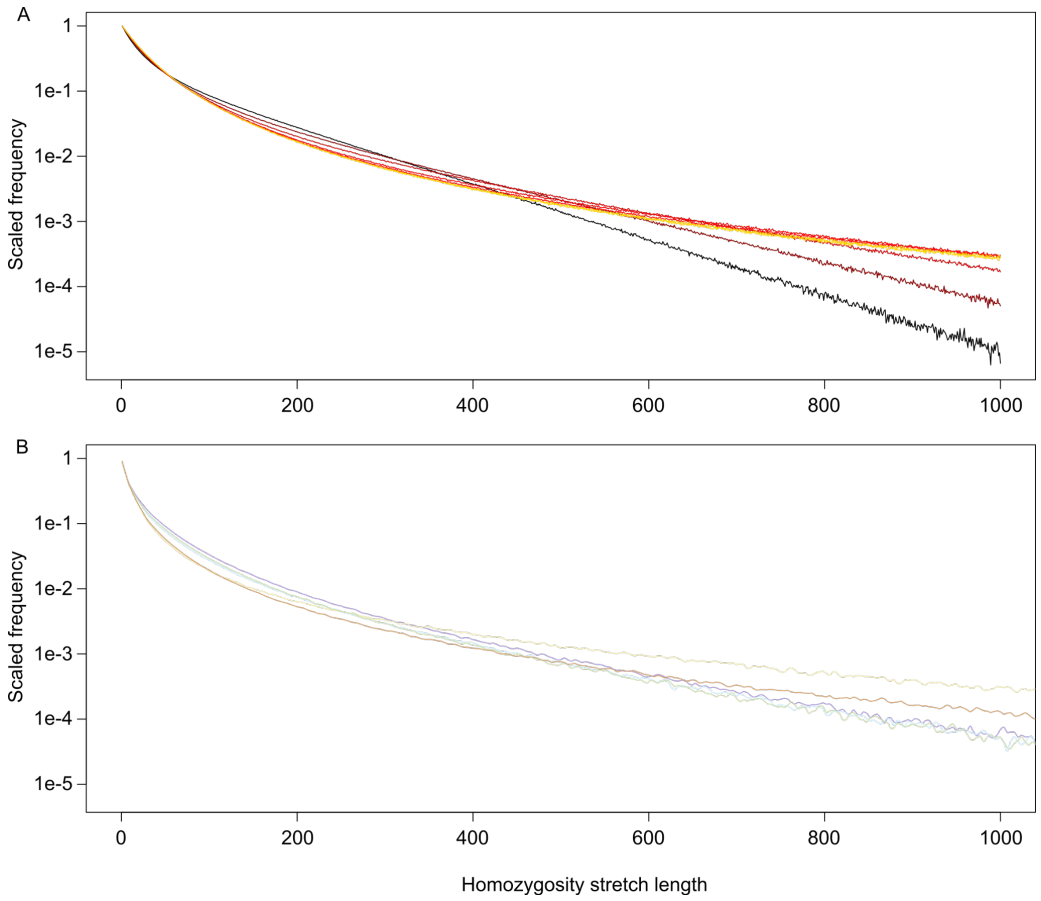


Fig. S9. Gene conversion exists in *Platynothrus peltifer* and the track lengths are at least 500bp. Distribution of homozygosity stretch lengths in **(A)** simulated and **(B)** empirical data. The numbers are normalized by dividing with the counts of stretches of size 0 (consecutive sites are heterozygous). The simulation shows the effects of GC track length on the 'homozygosity stretch', i.e. number of consecutive matching nucleotides between haplotypes. If the process is governed by mutation alone, each site should be independent, and lengths of homozygosity stretches geometrically distributed. GC biases the distribution so that longer stretches are more likely to occur (*76*). Line colors in **(A)** represent different GC track lengths: black = 50bp; dark to light red = 100, 200, 500, 1000bp; dark yellow and yellow = 2000, 5000bp and in **(B)** empirical mean track length estimates of the different populations. When the mean GC track is short (50bp), the distribution resembles the expected geometric; with the increase of track length, it becomes flatter by reducing the number of stretches < 400bp and increasing those > 400bp. However, once the mean GC track is 500bp or longer, increasing track length does not seem to change the distribution anymore. This asymptotic distribution is highly similar to the one observed from the empirical *P. peltifer* data.

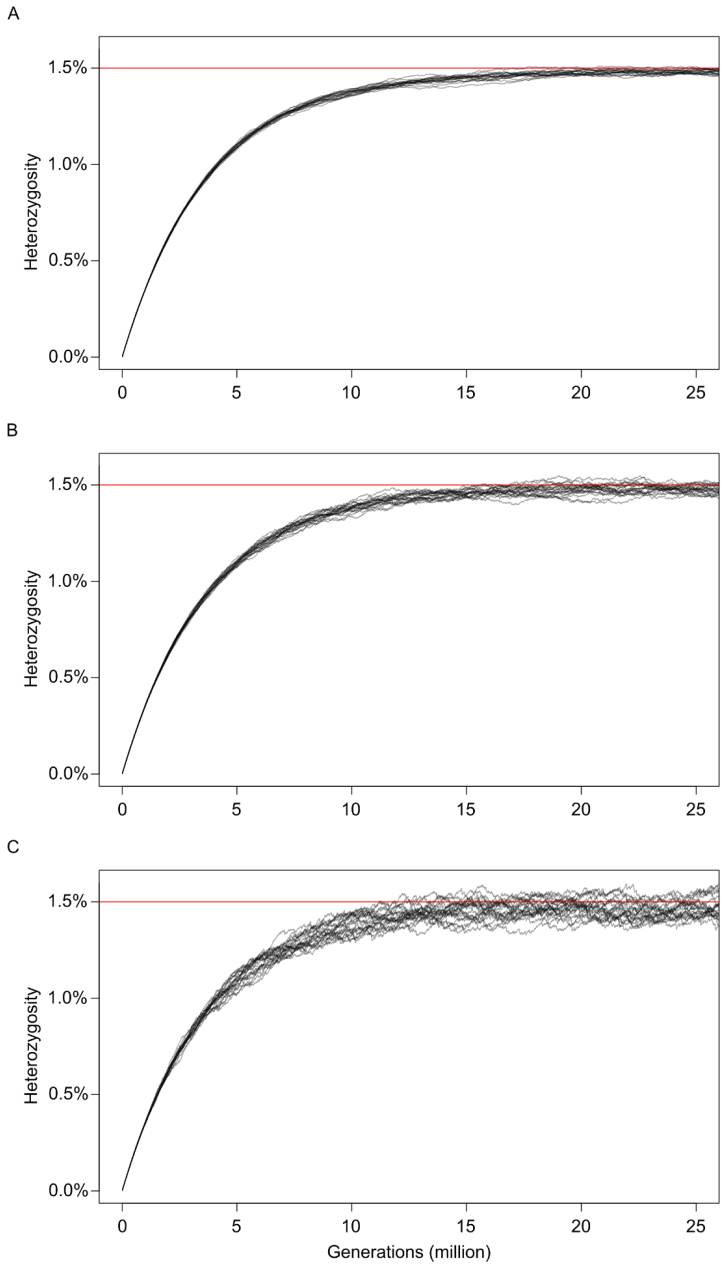


Fig. S10. Simulations of heterozygosity change (i.e., divergence of haplotypes) over time under different gene conversion track length in *Platynothrus peltifer*. The simulation was started from a hypothetical zero-heterozygosity pair of haplotypes, and the mutation rates estimated from *P. peltifer* and a gene conversion rate that gives a theoretical equilibrium heterozygosity of 1.5% (mean genome-wide heterozygosity of five populations) were applied. Average number of nucleotides affected by gene conversion per event were **A:** 200 **B:** 1000 **C:** 5000 bp. While a longer GC track does increase the variation of heterozygosity, it has a small effect compared to the total heterozygosity value.

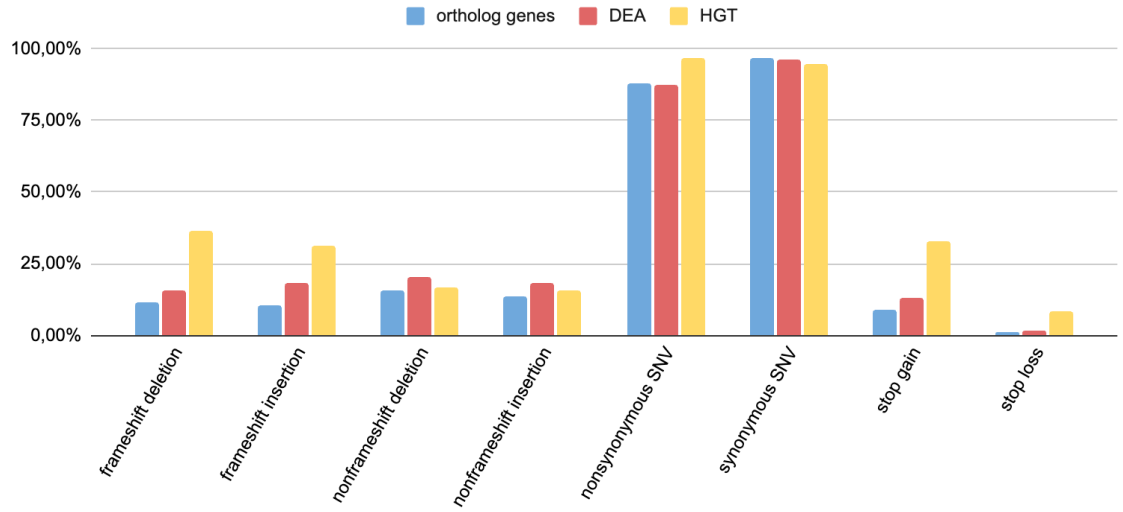


Fig. S11. Categories of different SNP effects and their percentage in orthologous genes (blue), DEAs (red) and HGTs (yellow).

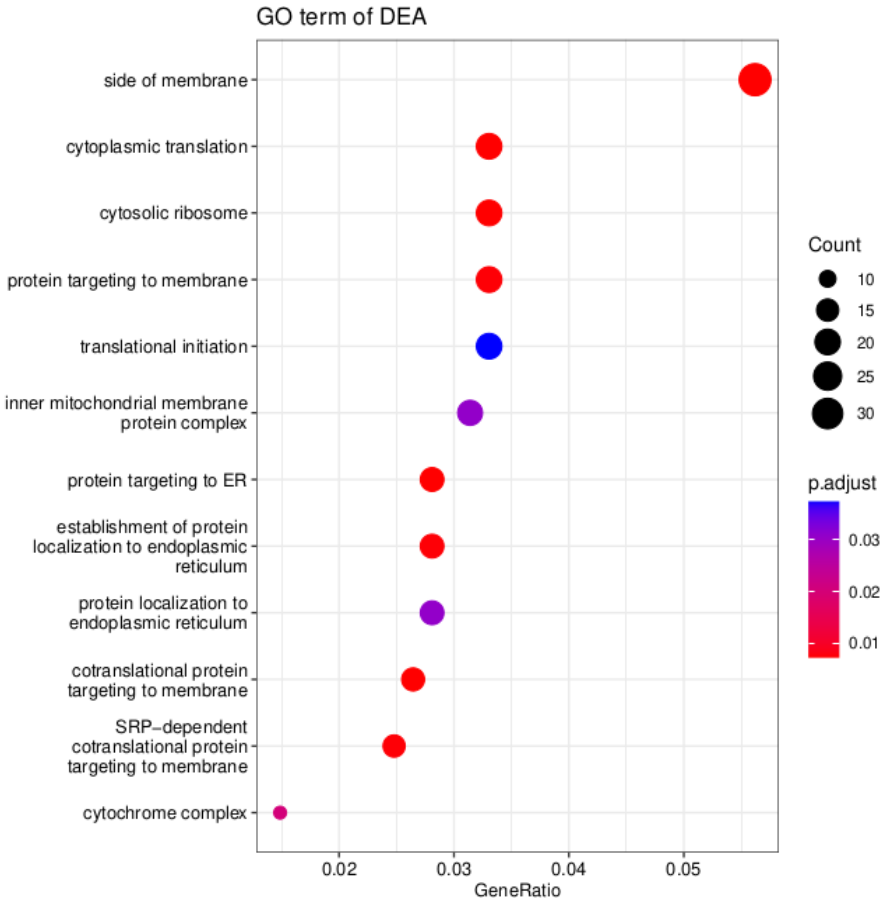


Fig. S12. Enrichment of differentially expressed alleles (DEAs) in GO terms.

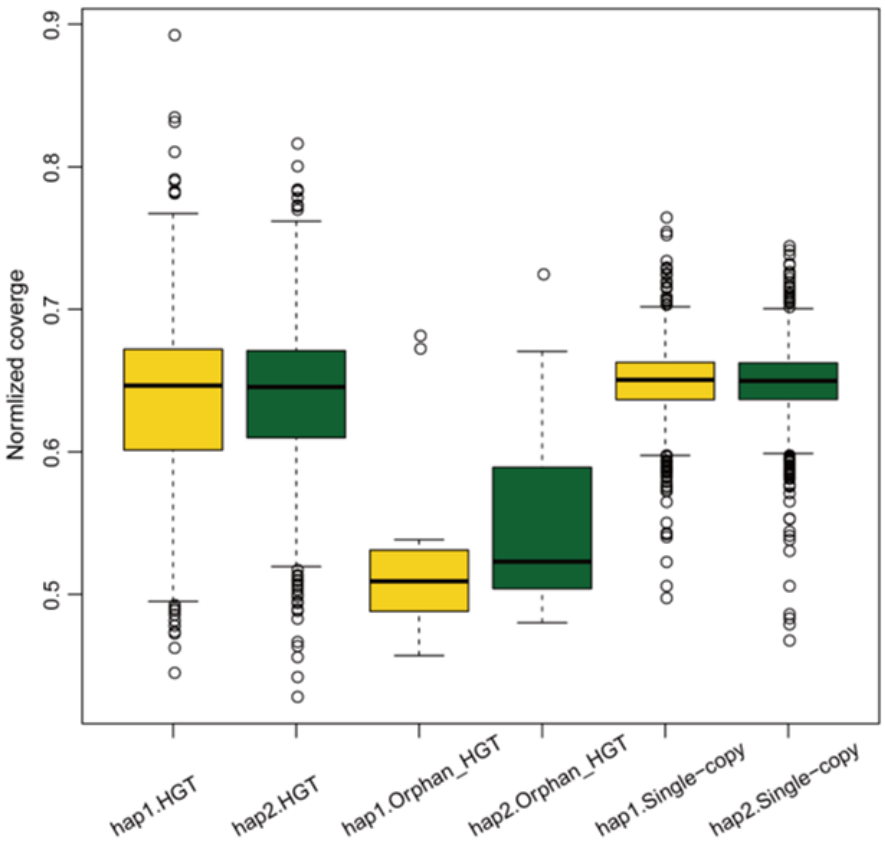


Fig. S13. The orphan HGTs show reduced mapped read coverage compared to HGTs and BUSCO single-copy genes found on both haplotypes. This is consistent with annotations of orphan HGTs being present in only one haplotype and suggests no missing HGT allele in the haplotype assemblies.

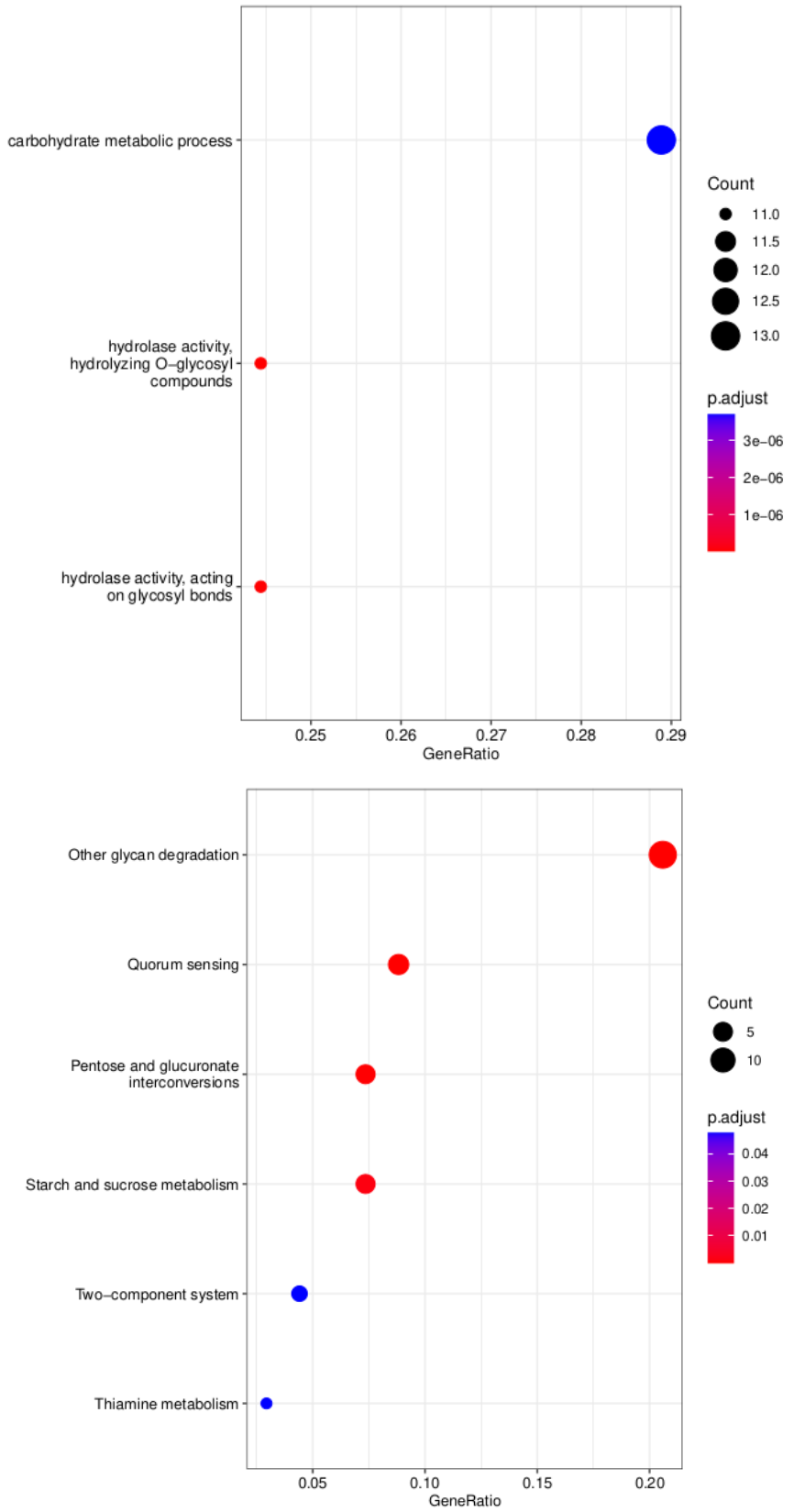


Fig. S14. Enrichment of HGTs in **a)** GO terms and **b)** KEGG pathways.

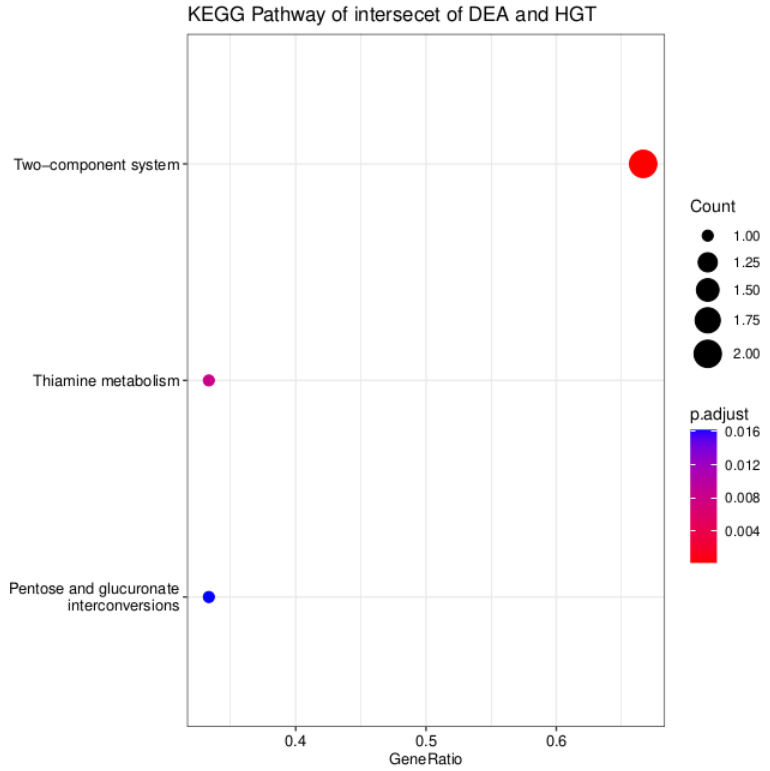


Fig. S15. KEGG pathways enrichment of HGT genes that exhibit differential expression.

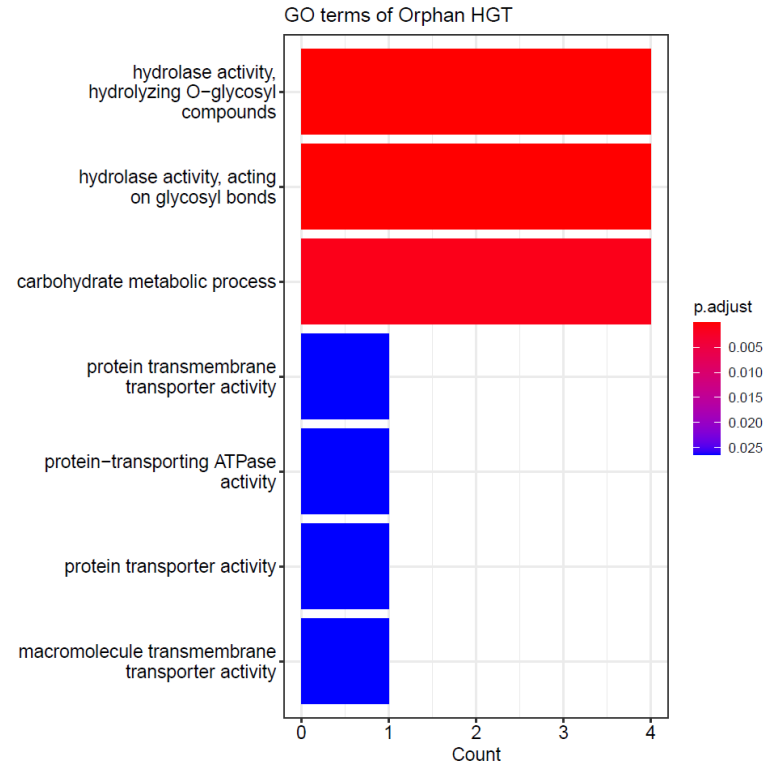


Fig. S16. GO term enrichment for orphan HGTs.

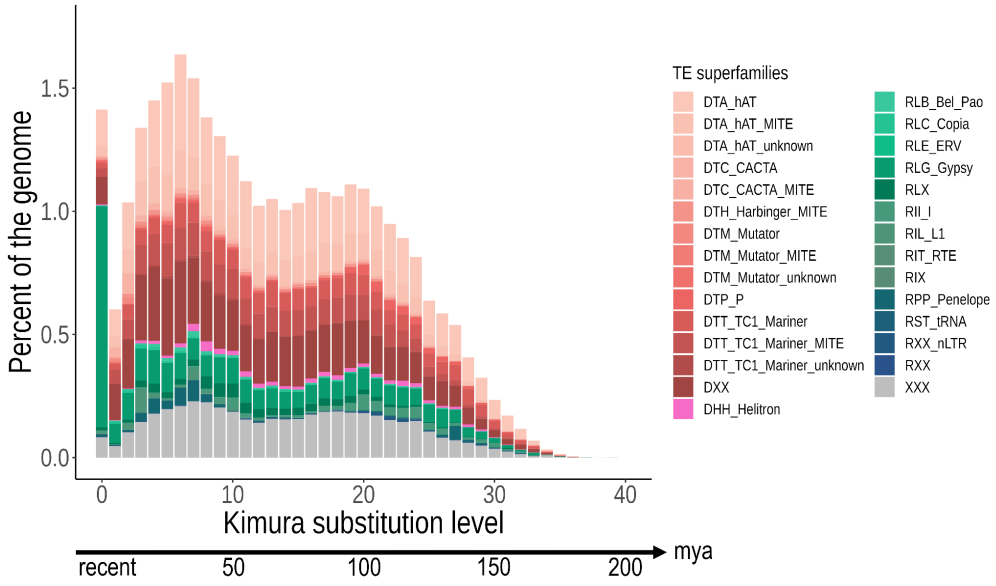


Fig. S17. Transposable element (TE) content and activity in the *Platynothrus peltifer* reference genome. The TE landscape shows historical TE activity and indicates recent activity and a main past expansion event approximately 29-49 mya (6%-10% divergence), coinciding with increase of global temperature dynamics.

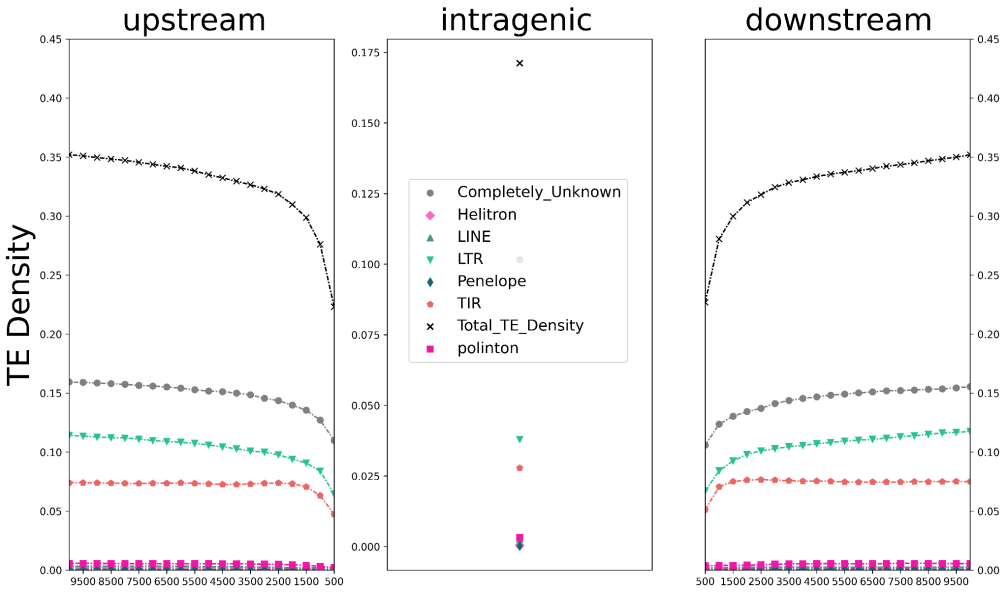


Fig. S18. Transposable elements are selected against in the *Platynothrus peltifer* reference genome. TE density patterns (of chromosome 1) upstream, intragenic and downstream suggest effective selection against TE insertions in the proximity and within genes (note differences in scale).

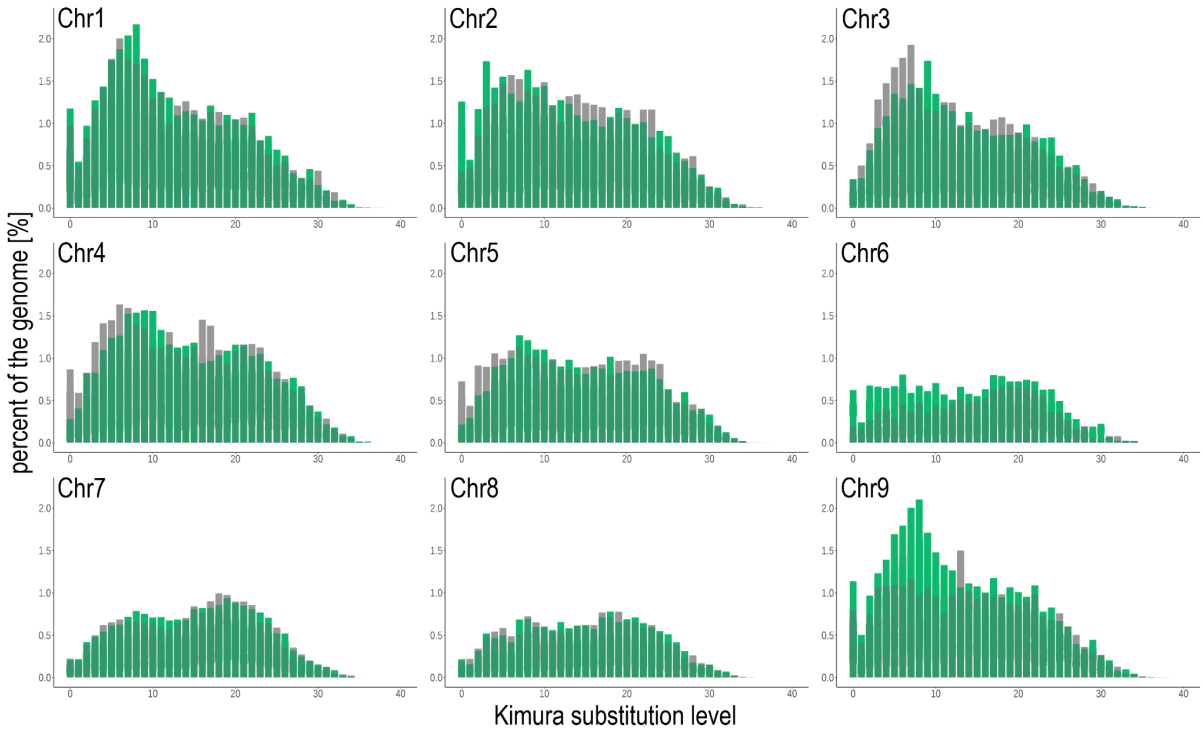


Fig. S19. Transposable element divergence landscapes of the largest haplotypic blocks per chromosome show slight but noticeable haplotype specific historical activity. Green bars represent haplotypic blocks A and gray bars haplotypic blocks B. For example, chr2 and chr5 show haplotype-specific repeat dynamics with a pronounced divergence in activity approximately at 6%, which translates into 29 mya. For chr4, chr6 and chr9 activity differences starting from 12% divergence translate to 59 mya.

Table S1. Assembly and annotation metrics of *Platynothrus peltifer*.

Metrics	Value
Assembly size [Mb]	219
N50 [Mb]	23
Number of scaffolds	9
BUSCO completeness [%] (arthropoda_odb10)	C:96.0 [S:93.9, D:2.1], F:1.0, M:3.0
BUSCO completeness [%] (arachnida_odb10)	C:96.7 [S:93.7, D:3.0], F:1.0, M:2.3
Nr annotated genes	24.9
Nr annotated genes (both haplotypes)	10.24
Repeats [%]	31.9
BUSCO Protein [%]	98

Table S2. Fragment sizes [bp] of haplotypic blocks A and B (in bold the longest scaffold in each chromosome used for Meselson effect and TE analyses)

Haplotypic blocks A		**Haplotypic Blocks B**	
alt1_chr1_1	**14957471**	**alt2_chr1_1_1**	**16050112**
alt1_chr1_2	3366384	alt2_chr1_1_2	30068
alt1_chr1_3	11361060	alt2_chr1_2_1	3329658
		alt2_chr1_3_1	11289522
alt1_chr2_1	4591776	alt2_chr2_1_1	4510490
alt1_chr2_2	5547886	alt2_chr2_1_2	30298
alt1_chr2_3	7322949	alt2_chr2_2_1	5700626
alt1_chr2_4	5292	alt2_chr2_3_1	9480104
alt1_chr2_5	**10752214**	alt2_chr2_3_2	163756
alt1_chr2_6	7574	alt2_chr2_3_3	41987
alt1_chr2_7	46558	**alt2_chr2_5_1**	**8601668**
		alt2_chr2_5_2	47936
alt1_chr3_1	676958	alt2_chr3_10_1	437948
alt1_chr3_2	3879045	alt2_chr3_10_2	34120
alt1_chr3_3	5119837	alt2_chr3_11_1	126106
alt1_chr3_4	38148	alt2_chr3_1_1	657132
alt1_chr3_5	**12829885**	alt2_chr3_2_1	4292554
alt1_chr3_6	16941	alt2_chr3_2_2	26733
alt1_chr3_7	1421540	alt2_chr3_2_3	23797
alt1_chr3_8	22959	alt2_chr3_2_4	15897
alt1_chr3_9	25189	alt2_chr3_3_1	4992518
alt1_chr3_10	375324	alt2_chr3_4_1	35663
alt1_chr3_11	179310	**alt2_chr3_5_1**	**11108905**
alt1_chr3_12	16590	alt2_chr3_5_2	115854
alt1_chr3_13	17503	alt2_chr3_6_1	16643
		alt2_chr3_7_1	966844
		alt2_chr3_7_2	21497
		alt2_chr3_8_1	38849
		alt2_chr3_9_1	47801
alt1_chr4_1	4929078	alt2_chr4_1_1	5114047
alt1_chr4_2	3965152	alt2_chr4_2_1	3867448

alt1_chr4_3	4381018	alt2_chr4_2_2	813
alt1_chr4_4	**9443682**	alt2_chr4_2_3	17851
alt1_chr4_5	6884	alt2_chr4_2_4	32058
		alt2_chr4_3_1	6073
		alt2_chr4_3_2	4246053
		alt2_chr4_4_1	**9773924**
		alt2_chr4_4_2	42082
alt1_chr5_1	**13159702**	**alt2_chr5_1_1**	**12553406**
alt1_chr5_2	3558262	alt2_chr5_1_2	901490
alt1_chr5_3	4240044	alt2_chr5_2_1	2145778
		alt2_chr5_3_1	5744347
alt1_chr6_1	6313945	alt2_chr6_1_1	6077717
alt1_chr6_2	3497893	alt2_chr6_2_1	2259272
alt1_chr6_3	516618	alt2_chr6_3_1	476161
alt1_chr6_4	32118	alt2_chr6_4_1	16509
alt1_chr6_5	**7722251**	**alt2_chr6_5_1**	**9189643**
alt1_chr6_6	2805762	alt2_chr6_5_2	11517
		alt2_chr6_5_3	32350
		alt2_chr6_6_1	2871417
alt1_chr7_1	**15067032**	**alt2_chr7_1_1**	**15422707**
alt1_chr7_2	3480545	alt2_chr7_1_2	18765
		alt2_chr7_1_3	15977
		alt2_chr7_2_1	3041516
alt1_chr8_1	**10309134**	**alt2_chr8_1_1**	**10145487**
alt1_chr8_2	7338071	alt2_chr8_1_2	20432
alt1_chr8_3	321376	alt2_chr8_2_1	8024092
		alt2_chr8_3_1	268508
alt1_chr9_1	**17136559**	**alt2_chr9_1_1**	**16549588**
alt1_chr9_2	1763873	alt2_chr9_1_2	71763
No_1	29782	alt2_chr9_1_3	28075
		alt2_chr9_1_4	27189
		alt2_chr9_2_1	2038438

Table S3. Summary statistics for the longest in-phase region that was used for haplotype tree reconstruction per chromosome. Target start and target end mark the start and end coordinates to extract the longest alignment block

query name	query length	query start	query end	strand direct-ion	target name	target length	target start	target end	# residue matches	alignment block length
alt1_chr 1_1	14957471	3454174	9999998	+	chr1	32608290	3587612	10326678	5228739	6864089
alt1_chr 2_5	10752214	28598	5717847	+	chr2	30492103	89265	5854202	4686177	5862286
alt1_chr 3_3	5119837	32321	5119734	+	chr3	27587832	45418	5170894	4366896	5196015
alt1_chr 4_1	4929078	15814	4928941	+	chr4	23397169	14530458	19553083	4185504	5063857
alt1_chr 5_1	13159702	2582264	6054325	+	chr5	22442908	4686733	8214567	3064880	3552017
alt1_chr 6_5	7722251	1802320	7722233	+	chr6	21557542	13485156	19401970	5149241	6050008
alt1_chr 7_1	15067032	7073018	12787819	+	chr7	19673716	7426596	13193208	4927722	5883881
alt1_chr 8_1	10309134	3280164	7993029	+	chr8	19511964	8588947	13277971	4100966	4785435
alt1_chr 9_1	17136559	7136559	11645221	+	chr9	18765566	7084104	11663639	3571460	4657571

Table S4. Genomic proportion of different transposable element superfamilies detected in the genome.

TE superfamily	Proportion in genome
DHH_Helitron	0.447
DTA_hAT	6.451
DTA_hAT_MITE	2.737
DTA_hAT_unknown	0.071
DTC_CACTA	0.388
DTC_CACTA_MITE	0.075
DTH_Harbinger_MITE	0.243
DTM_Mutator	0.074
DTM_Mutator_MITE	0.097
DTM_Mutator_unknown	0.004
DTP_P	0.481
DTT_TC1_Mariner	2.173
DTT_TC1_Mariner_MITE	2.618
DTT_TC1_Mariner_unknown	0.111
DXX	5.570
RII_I	0.599
RIL_L1	0.214
RIT_RTE	0.026
RIX	0.005
RLB_Bel_Pao	0.152
RLC_Copia	0.139
RLE_ERV	0.033
RLG_Gypsy	3.339
RLX	0.706
RPP_Penelope	0.494
RST_tRNA	0.015
RXX	0.038
RXX_nLTR	0.081
XXX	4.551
Total TEs in genome	31.932

Supplementary Materials Chapter 3

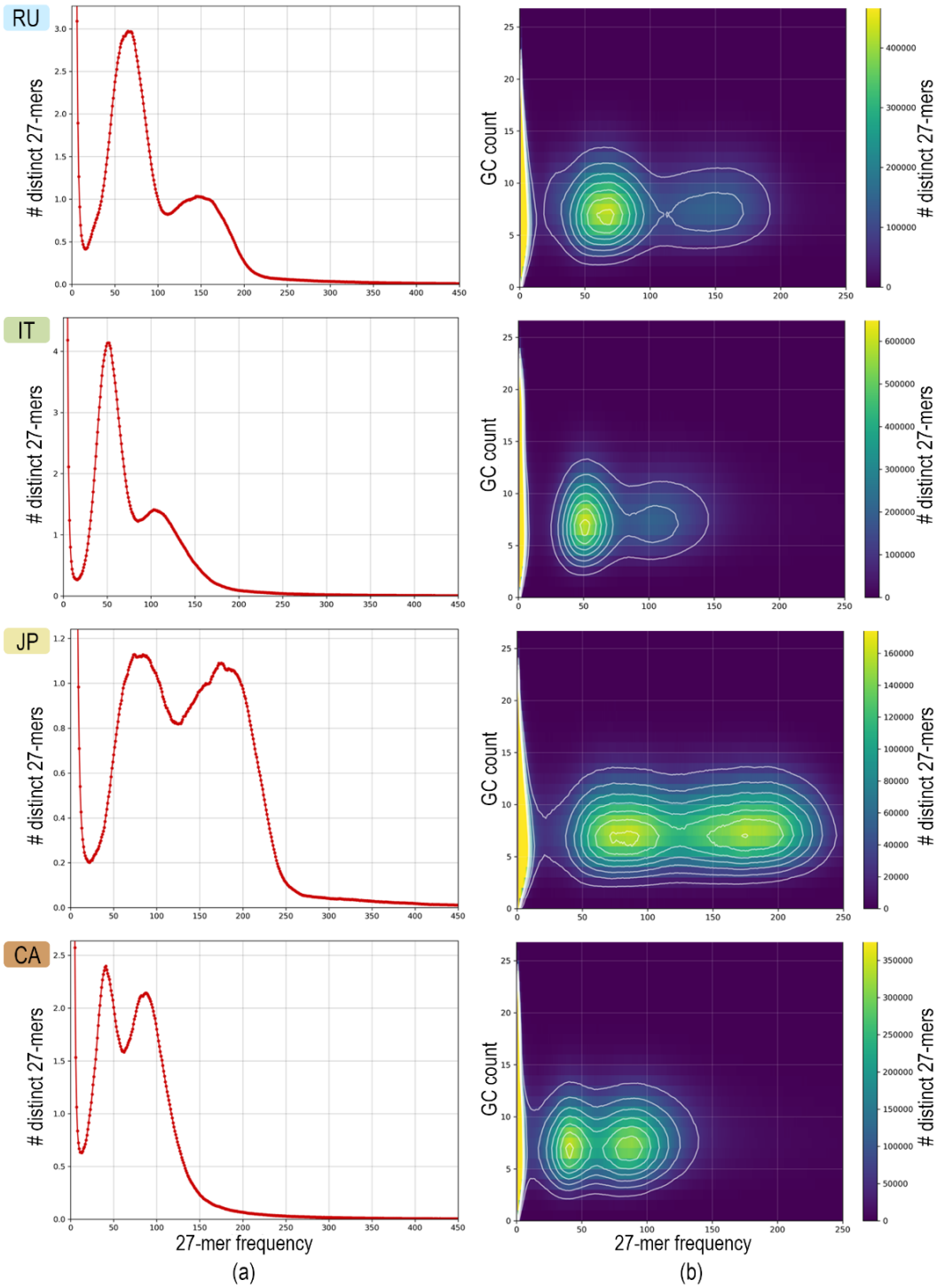


Figure S1: *k*-mer analysis of the PacBio HiFi reads (*k*=27) from different *Platynothrus peltifer* populations. a) Heterozygous and homozygous content is represented in two peaks respectively b) Plots show GC content of these two peaks. The plots further show no contamination, as contamination of the WGS datasets is expected to exhibit differing levels of coverage and/or GC content from the target species, and no further GC peaks can be detected. Abbreviations: RU=Russia, IT=Italy, JP=Japan, CA=Canada

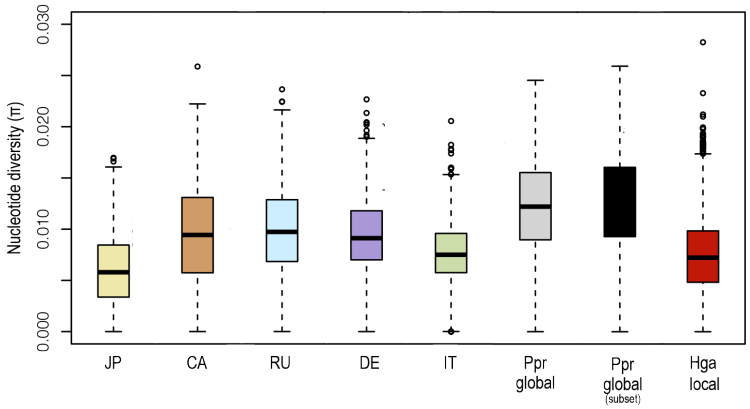


Figure S2: Nucleotide diversity of *Hermannia gibba*, global *Platynothrus peltifer* and population-specific of Japanese (JP), Canadian (CA), Russian (RU), German (DE), and Italian (IT) populations, as well as a subset of one individual of each population as a subset. Mean nucleotide diversity calculated in 1 Mb windows.

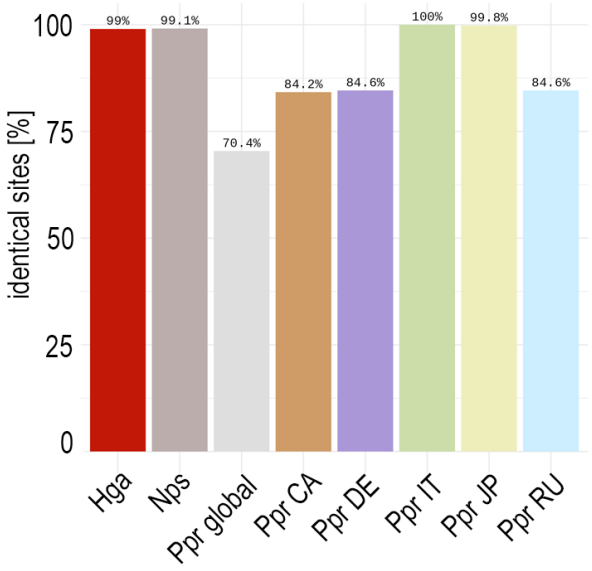


Figure S3: Aligned mitochondrial genomes of *Hermannia gibba*, *Nothrus palurstris* and various populations of *Platynothrus peltifer***.** Percentages show identical sites among individuals of the same population.

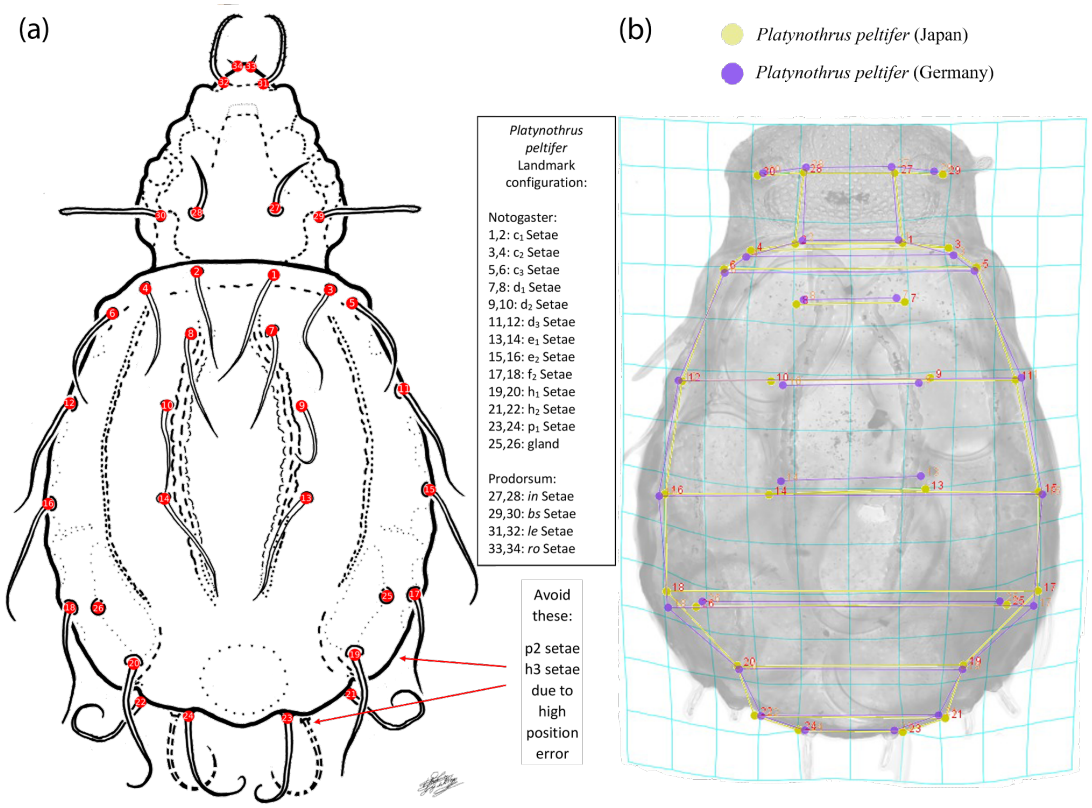


Figure S4: Landmark configuration of *Platynothrus peltifer* set to identify morphological differences through morphometrical analysis (provided by Dr. Fernando Villagomez). **a)** Sketch of *P. peltifer* and the set landmarks (highlighted in red dots) measured for each specimen, including 26 landmarks on the notogaster and 8 on the prodorsum. **b)** Thinplate spline and comparison of the average shape of German (purple) and Japanese (yellow) *P. peltifer*, showing differences in arrangement and distance between landmarks. Note: Significant differences were observed between landmarks: c_1, c_2, c_3, d_1, d_2, e_1 (data not shown).

Table S1: Mitochondrial genome of geographically distinct *Platynothrus peltifer* (Ppr) individuals as well as *Nothrus palustris* (Nps), *Hermannia gibba* (Hga), and *Heminothrus* sp. (Hsp) obtained in this study and their respective *cytochrome oxidase 1* (*cox1*) sequences.

ID	Species	mt genome length [bp]	*cox1* sequence length [bp]
DE 1	Ppr	14891	688
DE 2	Ppr	14891	599
DE 3	Ppr	14914	619
DE 4	Ppr	14914	**121**
DE 5	Ppr	not assembled	648
RU 1	Ppr	14911	523
RU 2	Ppr	14911	563
RU 3	Ppr	14909	648
RU 4	Ppr	14911	645
RU 5	Ppr	14909	644
IT 1	Ppr	14909	-
IT 2	Ppr	not assembled	635
IT 3	Ppr	14909	518
IT 4	Ppr	14909	556
IT 5	Ppr	not assembled	515
JP 1	Ppr	14930	479
JP 2	Ppr	14930	478
JP 3	Ppr	not assembled	487
JP 4	Ppr	14930	478
JP 5	Ppr	14930	468
CA 1	Ppr	14915	660
CA 2	Ppr	not assembled	652
CA 3	Ppr	not assembled	647
CA 4	Ppr	14905	654
CA 5	Ppr	14913	651
Hga 1	Hga	14340	646
Hga 2	Hga	14340	530
Hga 3	Hga	14340	542
Hga 4	Hga	14340	527
Hga 5	Hga	14340	536

Nps 1	Nps	not assembled	617
Nps 2	Nps	14546	641
Nps 3	Nps	14546	626
Nps 4	Nps	14546	628
Nps 5	Nps	14546	643
Hsp_JAP_F7	Hsp	not assembled	574
Hsp_JAP_F8	Hsp	not assembled	569
Hsp_JAP_1	Hsp	not assembled	507
Hsp_JAP_2	Hsp	not assembled	528
Hsp_JAP_3	Hsp	not assembled	511
Hsp_JAP_4	Hsp	not assembled	523
Hsp_JAP_5	Hsp	not assembled	544
Hsp_JAP_8	Hsp	not assembled	651
Hsp_JAP_9	Hsp	not assembled	508
Hsp_JAP_10	Hsp	not assembled	632
Hsp_JAP_11	Hsp	not assembled	642
Hsp_JAP_12	Hsp	not assembled	503
Hsp_JAP_13	Hsp	not assembled	639
Hsp_JAP_14	Hsp	not assembled	482
Hsp_JAP_15	Hsp	not assembled	618
Hsp_JAP_16	Hsp	not assembled	639
Hsp_JAP_17	Hsp	not assembled	609
Hsp_JAP_19	Hsp	not assembled	500

Table S2: Gene order of mitochondrial contigs of *Platynothrus peltifer*, *Nothurs palustris* and *Hermannia gibba*. Consensus sequences are highlighted. Underlined genes are on the reverse strand.

Hga1	cox1	cox2	D	atp8	atp6	cox3	G	nad3	A	L2	S1	E	F	nad5	H	nad4	nad4L	T	nad6	cob	S2	nad1	L1	M	W	C	nad2	I	Q	Y	P	rrnL	V	rrnS	N	K
Nps2	cox1	cox2	D	atp8	atp6	cox3	L2	nad3	L1	S1	G	E	F	nad5	H	nad4	nad4L	T	nad6	cob	S2	nad1	K	R	M	W	C	nad2	I	Q	Y	P	rrnL	V	rrnS	N
Ppr_DE1	cox1	cox2	D	atp8	atp6	cox3	nad3	L2	E	S1	F	nad5	H	nad4	nad4L	T	nad6	cob	S2	nad1	L1	M	W	C	nad2	I	Q	Y	P	rrnL	V	rrnS	N	R	K	G

Table S3: *Cytochrome oxidase I* divergence among clades observed from the phylogenetic analysis.

	number of nucleotides	**number of sequences**	**GC [%]**	**Identical sites [%]**
Platynothrus peltifer	46455	78	43.30	60.70
Clade I	1306	2	42.10	99.50
Clade II	11803	20	43.60	98.00
Clade III	11390	20	43.30	97.20
Clade IV	1958	3	43.00	96.80
Clade V	1306	2	42.10	99.50
Clade VI	3692	6	44.20	99.70
Clade VII	15106	25	43.00	92.40

Table S4: Shared pairwise identity of all *Platynothrus peltifer* individuals in protein alignments of cytochrome oxidase I (*cox1*) and cytochrome b (*cob*).

Gene	**no of sequences**	**pairwise identity [%]**	**sequence length [bp]**
cob	24	96.3	363
cox1	24	99.9	508

Supplementary Materials Chapter 4

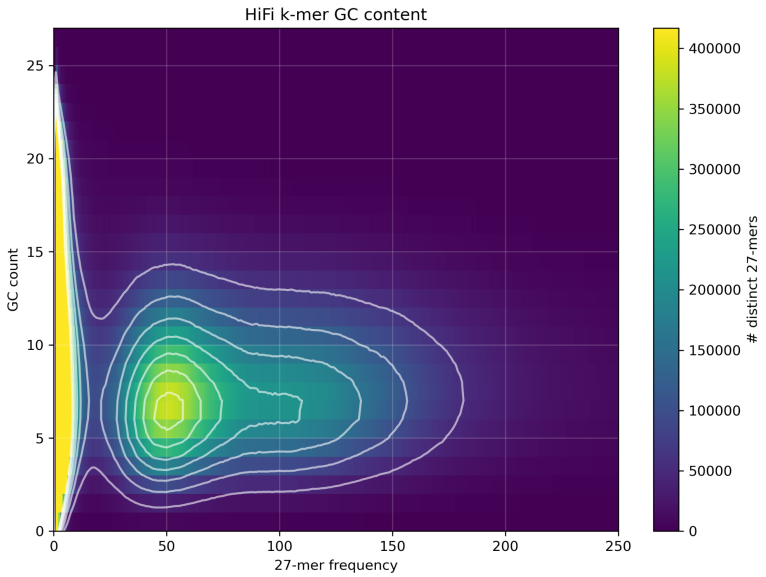


Figure S1 - Analysis of GC content and frequency of 27-mers. There are two peaks at frequencies of 50 and 100 (respectively, heterozygous and homozygous peaks) with a GC content mainly between 4 and 9. This plot shows no bacterial contaminants.

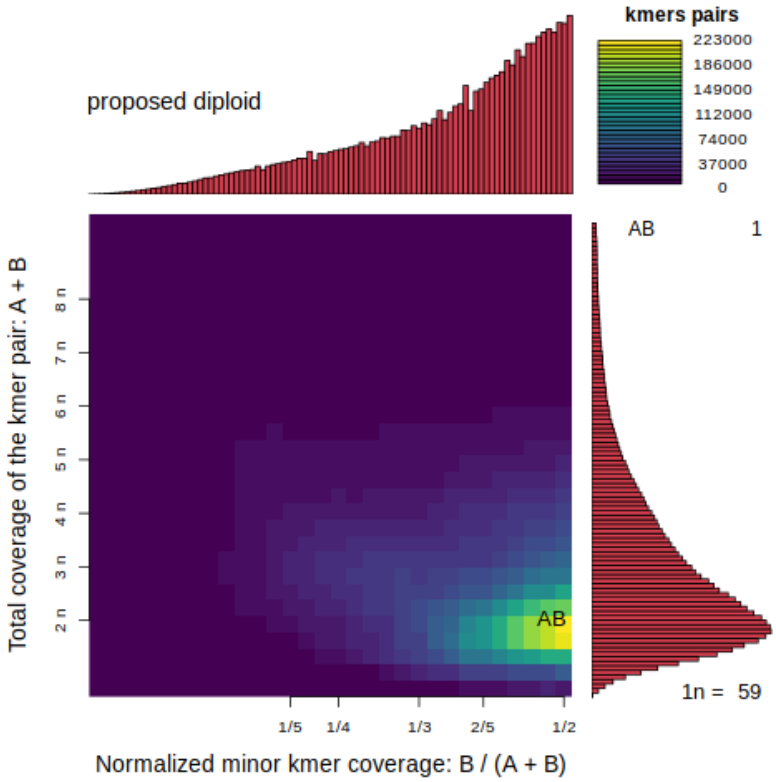


Figure S2 - *k*-mer analysis of ploidy. The strong AB smudge predicts that the species is diploid.

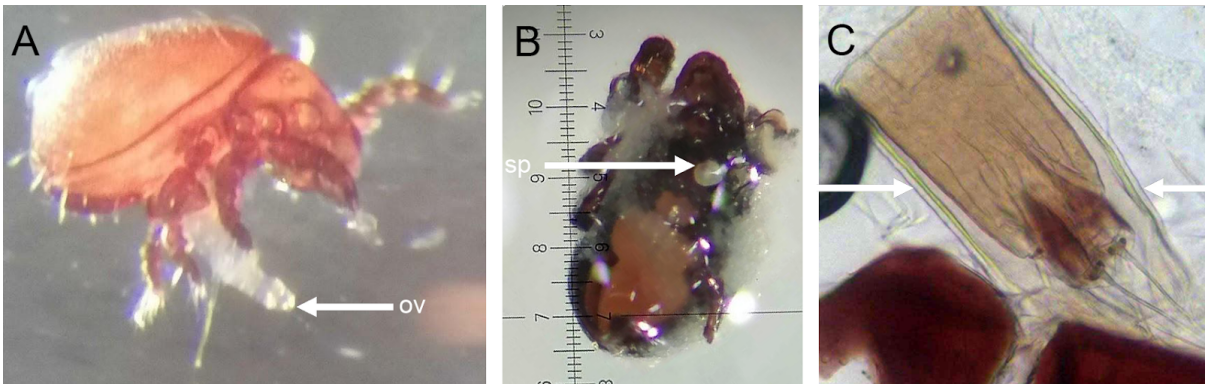


Figure S3 - Overview of morphological sex determination in *Hermannia gibba* **(A)** Arrow points to the ovipositor (ov) of an adult female **(B)** Arrows shows spermatophore (x5 magnification) (sp) of an adult male **(C)** In between arrows dissected ovipositor.

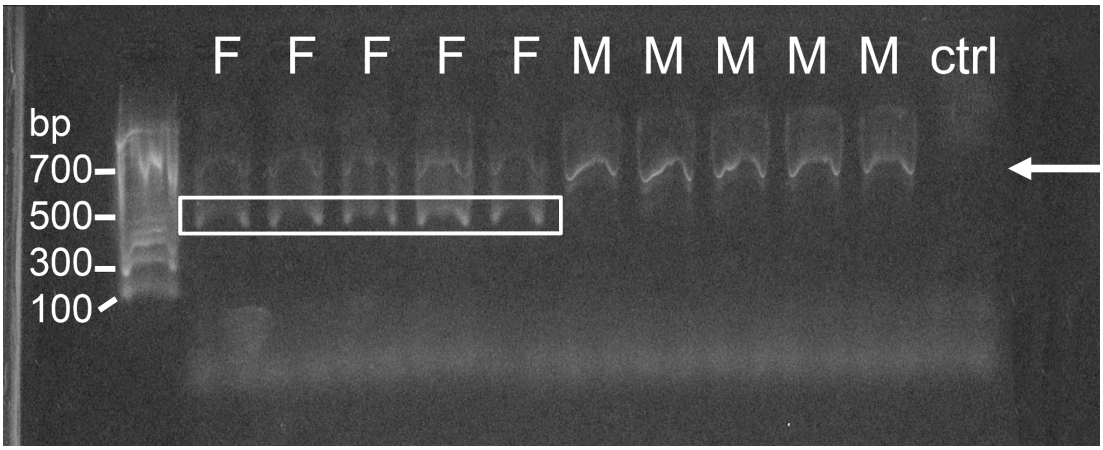


Figure S4 - Restriction enzyme digestion approach was conducted using following individuals with corresponding cDNA concentrations: female ID: 10: 6.20 ng/µl, 111: 2.85 ng/µl, 74: 2.79 ng/µl, 75: 6.06 ng/µl, 77: 3.95 ng/µl; male ID: 65: 5.25ng/µl, 67: 3.56 ng/µl, 68: 4.08 ng/µl, 70: 3.06ng/µl, 81: 5.39 ng/µl). The experiment targeted a female-specific SNP in the Sex Determining Region (SDR) on chromosome 1 using the HindIII restriction enzyme. In both the male and female samples, PCR amplification resulted in a 710-base pair product. After digestion, only female DNA produced two fragments (380 and 330 base pairs), confirming the presence of a female-heterogametic ZW sex-determination system in *H. gibba*.

Restriction enzyme digestion for sex verification

To additionally verify the identification of a ZW system and to be able to sex individuals collected from natural populations without the need for time-consuming morphological analyses, we designed a restriction enzyme digestion assay for *Hermannia gibba*. Therefore, we screened for the presence of a female-specific SNP in the SDR on chromosome 1, which is located within a restriction site of the HindIII restriction enzyme.

PCR amplification of both male and female products has an expected size of 710 base pairs. After restriction enzyme digestion, the PCR product was cleaved as expected into two fragments with sizes of 380 and 330 base pairs, in the female DNA pool only (Figure S4). This confirmed the presence of a female-heterogametic ZW sex-determination system in *H. gibba*.

Table S1 - Morphological sex determination of 30 specimens of both sexes, DNA concentration per individual (ng/µl), pooling amount per individual for equimolar DNA concentrations for each pool (µl)

No.	Sex	Sample-ID	Qubit (ng/µl)	Pooling amount (µl)	No.	Sex	Sample-ID	Qubit (ng/µl)	Pooling amount (µl)
1	Female	1	4,45	7,19	1	Male	65	3,05	10,49
2	Female	2	5,34	5,99	2	Male	66	3,95	8,1
3	Female	3	5,07	6,31	3	Male	67	4,49	7,13
4	Female	4	7,08	4,52	4	Male	68	3,92	8,16
5	Female	5	5,25	6,1	5	Male	70	5,23	6,12
6	Female	6	6,84	4,68	6	Male	71	3,71	8,63
7	Female	7	3,96	8,08	7	Male	81	2,97	10,77
8	Female	8	4,71	6,79	8	Male	86	3,34	9,58
9	Female	9	5,51	5,81	9	Male	87	3,61	8,86
10	Female	10	6,2	5,16	10	Male	88	3,5	9,14
11	Female	74	2,79	11,47	11	Male	89	2,32	13,79
12	Female	75	6,06	5,28	12	Male	90	4,17	7,67
13	Female	77	3,95	8,1	13	Male	92	3,91	8,18
14	Female	78	3,12	10,26	14	Male	93	3,96	8,08
15	Female	105	5,67	5,64	15	Male	94	3,99	8,02
16	Female	110	5,19	6,17	16	Male	95	2,37	13,5
17	Female	111	2,85	11,23	17	Male	100	6,66	4,8

18	Female	114	5,92	5,41	18	Male	107	6,59	4,86
19	Female	125	5,49	5,83	19	Male	112	3,32	9,64
20	Female	126	4,48	7,14	20	Male	113	2,15	14,88
21	Female	127	6,02	5,32	21	Male	115	2,39	13,39
22	Female	128	4,99	6,41	22	Male	117	4,03	7,94
23	Female	129	3,66	8,74	23	Male	118	7,93	4,04
24	Female	137	5,3	6,04	24	Male	119	6,83	4,69
25	Female	138	3,86	8,29	25	Male	120	5,24	6,11
26	Female	139	5,38	5,95	26	Male	121	3,28	9,76
27	Female	140	2,82	11,35	27	Male	122	4,97	6,44
28	Female	142	6,16	5,19	28	Male	123	4,52	7,08
29	Female	143	4,18	7,66	29	Male	124	6,77	4,73
30	Female	144	7,15	4,48	30	Male	130	4,67	6,85

Eidesstaatliche Erklärung

Erklärung zur Dissertation gemäß der Promotionsordnung vom 12. März 2020

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Teilpublikationen

Modified salting out method for high-molecular-weight gDNA extraction (oribatid mites)
Protocols.io (2023) https://doi.org/10.17504/protocols.io.yxmvm3yybl3p/v1

Haplotype independence contributes to evolvability in the long-term absence of sex in a mite
bioRxiv (2023) https://doi.org/10.1101/2023.09.07.556471

A female heterogametic ZW sex-determination system in Acariformes
bioRxiv (2023) https://doi.org/10.1101/2023.10.24.563255

Having babies in soil: is sex really necessary?
Front. Young Minds. 9:611659. (2021) https://doi.org/10.3389/frym.2021.611659

Datum: 30.10.23, Köln

Name und Unterschrift: Hüsna Öztoprak

Hüsna Öztoprak

[redacted]
[redacted] [redacted]
[redacted]
https://orcid.org/0000-0003-0815-1893

Education

- University of Cologne 2019 - now

degree: PhD, title: Evolutionary persistence and speciation under ancient asexuality

degree: M. Sc. Biological Sciences, research focus: ecology and evolution 2017 - 2019

Thesis: Through the magnifying glass: the global diversity of Rhogostomidae and their environmental drivers

degree: B. Sc. Biology 2013 - 2016

Thesis: Polyphyly of the genus *Plagiophrys*

- Rhein-Gymnasium 2004 - 2013

graduation: Abitur

Conferences and presentations

- **CNRS Conference Jaques Monod "Sex uncovered"** Roscoff, France 2023

Poster: Ancient asexual speciation in the oribatid mite *Platynothrus peltifer.*

- **SMBE Annual Meeting** Ferrara, Italy 2023

Poster: Ancient asexual speciation in the oribatid mite *Platynothrus peltifer.*

- **International Symposium on "Reference genomes for biodiversity"** Cologne, Germany 2022

Talk: Genomic signatures of ancient asexuality in the oribatid mite *Platynothrus peltifer*

- **Congress of the European Society for Evolutionary Biology (ESEB)** Prague, Czechia 2022

Talk: Genomic signatures of ancient asexuality in the oribatid mite *Platynothrus peltifer*

- **VIII European Congress of Protistology (ECOP) & International Society of Protistologists (ISOP) joint meeting** Rome, Italy 2019

Talk: Through the magnifying glass – the global diversity of Rhogostomidae and their environmental drivers

- **38th Meeting of the German Society for Protozoology (DGP)** Vienna, Austria 2019

Poster: Through the magnifying glass – the global diversity of Rhogostomidae and their environmental drivers

Research Experience

Workshop on Academic Writing in the Natural and Life Sciences (2021)

Workshop on Stress-Resilience (2021)

Workshop on Scientific Presentation - Getting it Right! (2021)

Workshop on Good Scientific Practice for Empirical Research (2021)

Workshop on Story Telling in Academia - Discover your research narrative; engage your reader (2021)

Internship at the Japan Agency for Marine Science and Technology 09/18 - 10/18

Department of Marine Biodiversity Research
with a focus on the principles and techniques of electron microscopy

Internship at Eric von Elert's lab 03/18 - 04/18

Aquatic Chemical Ecology, University of Cologne
with a focus on bioassay-guided isolation of active compounds from cyanobacteria and fish incubation water: an approach with Daphnia

Internship at Michael Bonkowski's lab

Terrestrial Ecology, University of Cologne
1st with a focus on large subunit ribosomal rRNA (LSU) gene sequence phylogeny 04/16 - 05/16
2nd with a focus on the bioinformatical analysis of throughput-(Illumina) sequencing 08/17 - 09/17

Internship at the Biologische Station Rhein-Kreis Neuss 02/16 - 03/16

with a focus on ongoing field studies in species preservation

Teaching Experience

- **Practical course** part of the Bachelor's program of the University of Cologne:
  Biodiversity of Acari Summer Semester 2022 & Winter Semester 2022/2023
- **1:1 teaching of students**

2021: [redacted]
2022: [redacted]
2023: [redacted]

Professional Memberships

[redacted]

Awards and Honours

- Holz-Conner award, ISOP
- 3rd award for the best poster presentation, DGP

Additional Skills

[redacted]

Employments

[redacted]

Linguistic Proficiency

German (native speaker) | Turkish (native speaker) | English

Publications

Wulsch S, **Öztoprak H**, Guiglielmoni N, Jeffries DL, Bast J (2023). A female heterogametic ZW sex-determination system in Acariformes. *bioRxiv* 2023.10.24.563255.

Öztoprak H, Gao S, Guiglielmoni N, Brandt A, Zheng Y, Becker C, Becker K, Bednarski V, Borgschulte L, Burak KA, Dion-Côté AM, Leonov V, Opherden L, Shimano S, Bast J (2023) Haplotype independence contributes to evolvability in the long-term absence of sex in a mite. *bioRxiv* 2023.09.07.556471.

Öztoprak H, Bast J (2023) Modified salting out method for high molecular weight gDNA extraction (oribatid mites). *protocols.io.*

Dumack K, Ferlian O, Gysi DM, Degrune F, Jauss RT, Walden S, **Öztoprak H**, Wubet T, Bonkowski M & Eisenhauer N (2022) Contrasting protist communities (Cercozoa: Rhizaria) in pristine and earthworm-invaded North American deciduous forests. *Biological Invasions* 24: 1345-1357.

Öztoprak H, Brandt A, Solbach M, Bast J and Schaefer I (2021) Having Babies in Soil: Is Sex Really Necessary? *Front. Young Minds.* 9:611659.

Öztoprak H, Walden S, Heger T, Bonkowski M & Dumack K (2020) What drives the diversity of the most abundant terrestrial Cercozoan family (Rhogostomidae, Cercozoa, Rhizaria)? *Microorganisms* 8(8),1123.

Dumack K, **Öztoprak H**, Rüger L, Bonkowski M (2017) Shedding light on the polyphyletic thecate amoeba genus *Plagiophrys*: Transition of some of its species to *Rhizaspis* (Tectofilosida, Thecofilosea, Cercozoa) and the establishment of *Sacciforma* gen. nov. and *Rhogostomidae* fam. nov. (Cryomonadida, Thecofilosea, Cercozoa). *Protist* 168: 92-10.