

The genomic implications of sex and its absence in old taxa



Dissertation
zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln
vorgelegt von
Shan Gao
aus Shaanxi
Köln, 2025

Begutachtung:

Prof. Dr. Thomas Wiehe

Prof. Dr. Bart Thomma

Vorsitz der Prüfung:

Prof. Dr. Gunther Doehlemann

Beisitzer:

Dr. Jens Bast

Summary

Reproduction is a fundamental characteristic of organisms, and sexual reproduction is nearly universal in eukaryotes. The last eukaryotic common ancestor (LECA) is thought to have been capable of sexual reproduction, as this mode of reproduction is common across many different groups of eukaryotes. Sexual reproduction has three critical processes: meiosis, recombination and cell fusion, all of which are considered ancient aspects of sexual reproduction.

Protists are considered the closest eukaryotes to the LECA and have evolved into many clades. Their abundance and primitive characteristics are crucial for elucidating the origins and evolution of sexual reproduction. However, most of the protists were considered to be asexual or facultative sex because of less sexual behavior or gender observed. This may be due to the protists being too diminutive to directly observe and trace in nature, and examining their entire lifespan when culturing most unicellular protists in laboratory presents a challenge. Detecting potential sexual processes by identifying meiosis-related genes and their expression dynamics may serve as a valuable approach during cell fusion.

Fisculla terrestris (Thecofilosea, Rhizaria, SAR Supergroup) is a unicellular, shell-bearing amoeba living in soils that can be reared in the lab. Starvation triggers cell fusion and cell disintegration when food re-supplied are observed in *F. terrestris*. To check the presence of meiosis-associated genes and the dynamics of their expression during cell separation and aggregation, we sequenced the transcriptomes of the two treatments of starvation and food supply. We found that genes related to recombination are upregulated in the starving treatment compared to the food supply treatment. This might imply the exchange of genetic material and cryptic sex.

Contrasting unicellular protists, most metazoans are known to reproduce sexually. However, there are a few metazoans that have stopped reproducing sexually and reproduce obligately asexually. Although asexuality was regarded as a dead end in evolution due to harmful mutation accumulation, some organisms still survived and evolved in the long-term without sex: "ancient asexual scandals," such as oribatid mites, bdelloid rotifers, and darwinulid ostracods. Studying how these organisms escape

Summary

extinction and adapt to the environment can help us to understand why sexual reproduction is important and how these species evolved.

To elucidate the enigma of asexual evolution following a prolonged absence of sex, we used genome and haplotypic assemblies of *P. peltifer* to detect genomic features as a consequence of asexuality over time. We detected the “Meselson effect” in the *P. peltifer* genome, which indicated long-term asexual reproduction. The haplotypes independently evolved, leading to one allele relaxed and the other allele conserved, which was also reflected on the patterns of differently expressed alleles (DEAs) and horizontal gene transfer (HGT). About 2% of genes showed a HGT origin, and they might provide new traits like cell wall degradation and pesticide resistance, strengthening the mites as decomposers in soil. Selection analysis showed that most of the alleles are under purifying selection, potentially alleviating deleterious mutation accumulation, while certain genes under positive selection may facilitate adaptation. Our results provide insight into long-term asexual genomic features and the potential mechanism for ancient asexual persistence and evolution.

Exploring the sexual reproduction of basal eukaryotic protists and studying the genomic features of obligately ancient asexual mites can help understand the prevalence of sexual reproduction. Additionally, the putative sexual reproduction in protists and genomic evolution after loss of sex in mites can provide a new insight into the implication of sexual reproduction.

Contents

Summary	I
Contents	III
1. Introduction	1
1.1 Sexual Reproduction in the old taxa	1
1.1.1 The origin of meiotic sex	1
1.1.1.1 The first occurrence of sex	1
1.1.1.2 Meiosis originates from mitosis	2
1.1.1.3 Environmental factors likely favored the origin of sex	3
1.1.1.4 Sex and mitochondria	4
1.1.2 The Conservation and Variability of Meiotic Sex	5
1.1.2.1 Conservation of Meiosis	5
1.1.2.2 Variability of Meiosis	7
1.1.3 Sexual reproduction in old taxa “Protista” Kingdom	7
1.2 Why sex? Evolution of asexuality	9
1.2.1 The disadvantages of asexual reproduction	9
1.2.2 Ancient Asexual Scandals: Exception to the Rule?	10
1.2.3 The Evolution of asexuality	12
1.2.3.1 The Meselson Effect in orbatid mites	12
1.2.3.2 Horizontal gene transfer provide evolutionary substrates for asexuality	12
1.2.3.3 Transposable element (TE) activity in asexuals	13
1.3 Aims and Outline of this Study	15
1.4 References of the Introduction	17
2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist <i>Fisculla terrestris</i>	27
3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite <i>Platynothrus peltifer</i>	37
5. General discussion	53
6. Appendix	61
Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist <i>Fisculla terrestris</i>	61

Contents

Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite <i>Platynothrus peltifer</i>	63
7. Declaration	101
8. Acknowledgements	102
9. Curriculum vitae.....	103

1. Introduction

1.1 Sexual Reproduction in the old taxa

1.1.1 The origin of meiotic sex

1.1.1.1 The first occurrence of sex

Sex (including meiosis, recombination and syngamy) is considered an important feature of the last eukaryotic common ancestor (LECA, approximately 1.8 to 2.7 billion years ago), due to its prevalence and conserved attributes among eukaryotes (Vosseberg et al. 2024). The LECA is not a real organism, it is a hypothesized ancestral situation, and its abstract state was inferred by the phylogenetic reconstructions according to all known extant eukaryotes (**Figure 1.**; Krupovic et al. 2023). The exact type of LECA remains unclear, as LECA might have been a single cell or a genetically homogeneous population, or a genomically heterogeneous population, or a consortium (including prokaryotic symbionts) (O'Malley et al. 2019). Besides sexual processes, the LECA might have also contained other features of the crown eukaryotic group, including a nucleus, mitosis, mitochondria, a centriole, a cell wall composed of chitin and/or cellulose, and peroxisomes. These characteristics were thought to have emerged during the evolutionary transition from the first eukaryotic common ancestor (FECA) to the LECA (Fu et al. 2019). From the FECA to the LECA, this period can be defined as the eukaryogenesis and is estimated to have occurred within a broad temporal range of 1.8 to 2.7 billion years ago (Vosseberg et al. 2024). These characteristics might have evolved independently or through mutual influence; it is still an enigma.

The FECA is the first descendant (on the eukaryotic branch) of the last common ancestral node of eukaryotes and an “archaeal” lineage. FECA is the ancestor of all the eukaryotes, whatever they are, whether they are extinct or not. The evolutionary history between FECA and LECA remains a mystery. It is hard to get a satisfactory explanation for how meiosis occurred in this period. But according to the earliest fossil evidence of sexual reproduction, a type of red algae known as *Bangiomorpha pubescens* was discovered in Canada in 1990 and is approximately 1.1 billion years old (Brocks et al. 2023). The

1. Introduction

accurate time when sex and meiosis occurred should be older than the fossil record but later than the FECA (Goodenough and Heitman 2014; Javaux 2023).

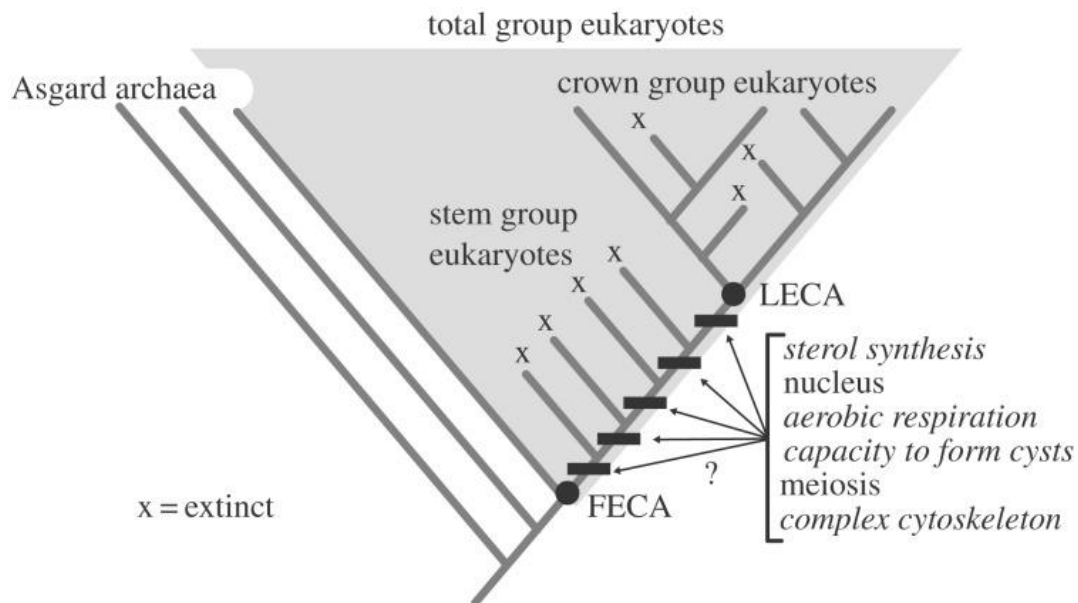


Figure 1. Fossil record and the putative time line of eukaryogenesis (Porter 2020).

1.1.1.2 Meiosis originates from mitosis

Mitosis is a common means of cell proliferation, the purpose of which is to produce new genetically identical daughter cells from one cell, including one DNA replication and one division. Meiosis can be simply depicted as a kind of cell division with “once replicated, twice divided”, but it is not simply twice mitosis after DNA replication. It also includes four innovative steps: the pairing of homologous chromosomes, the occurrence of extensive recombination between non-sister chromatids during pairing, the suppression of sister-chromatid separation during the first meiotic division, and the lack of chromosome replication during the second meiotic division (Wilkins and Holliday 2009). Furthermore, upon the initiation of the two processes, chromosome condensation and movement exhibit remarkably analogous mechanisms within the cell, with meiosis repurposing the essential mitotic machinery for successive divisions and homologous pairing (Marston and Amon 2004). At the genetic level, many mitotic genes also contribute to the meiosis process, and certain meiosis-specific genes have evolved from mitotic genes, such as the meiotic cohesin Rec8 derived from the mitotic cohesin Rad21, and the meiosis-specific recombination protein DMC1 originating from mitotic RecA-like proteins (Ward et al.

2016). These changes to proteins might be key to the basic machinery of mitotic division turning into the more complicated meiotic process. This made it possible for recombination and segregation to happen while keeping the basic mechanisms of chromosome segregation and recombination.

1.1.1.3 Environmental factors likely favored the origin of sex

Understanding the environmental changes of earth can help understand the potential environmental pressure for the origin of eukaryotes and meiosis. It is widely recognized that prokaryotes represent the earliest life forms and promote the occurrence of eukaryotes (Cooper 2000; Vosseberg et al. 2024). The earliest fossil evidence of life found in northeastern Canada has been found to contain carbon traces of life dating back at least 3.95 billion years (Tashiro et al. 2017). It demonstrated that life persists significantly longer than indicated by the fossils.

The earth's early stage (approximately 4.5 billion years ago) was not an ideal environment for the majority of life today. As anticipated, the initial atmosphere was devoid of oxygen and elevated in carbon monoxide, carbon dioxide, methane, water vapor, and ammonia volume (Kasting 1993). About 3.8-3.5 billion years ago, the bacteria and archaea occurred, but they are chemotrophs. According to the analysis of oxygen-utilizing enzymes, oxygen likely became available around 3.1 billion years ago (Jabłońska and Tawfik 2021). Blue-green algae bacteria might have been the first organisms to produce oxygen via photosynthesis; this ability has extremely changed the makeup of the atmosphere (Demoulin et al. 2024). This so-called first Great Oxygenation Event (GOE) provided the foundation for eukaryotes to appear (Lyons et al. 2014). This event occurred approximately 2.4 billion years ago and likely played a crucial role in shaping the evolution of meiosis in early eukaryotes (Holland 2006). As atmospheric oxygen levels rose dramatically, early cells faced a new challenge: oxygen-induced DNA damage. Reactive oxygen species (ROS) and free radicals were made by the highly reactive oxygen molecules. These could damage DNA and break DNA strands in different ways (Commoner et al. 1954; Gerschman et al. 1954). In response to this oxygen-induced stress, cells evolved sophisticated DNA repair mechanisms. These repair processes eventually developed into what we now recognize as meiosis, where homologous recombination allowed cells to use undamaged DNA templates for repair while simultaneously generating genetic diversity (Bernstein and Bernstein 2010). Many

1. Introduction

meiotic proteins come from DNA repair proteins, which is strong evidence for this evolutionary pathway, suggesting that the oxidative environment was a basis for the evolution of sexual reproduction (Villeneuve and Hillers 2001).

1.1.1.4 Sex and mitochondria

Eukaryotes are thought to have evolved from prokaryotes (Woese and Fox 1977; Philippe and Forterre 1999). Asgardarchaeota represents the most closely related archaeal lineage to eukaryotes, characterized by possessing eukaryotic-like proteins (Eme et al. 2017; Zaremba-Niedzwiedzka et al. 2017). Notably, experimental studies have demonstrated that certain Asgardarchaeota proteins can functionally substitute for homologous proteins in yeast, suggesting significant structural and functional conservation across these evolutionary domains (Bonifacino and Glick 2004; Klinger et al. 2016). Promethearchaeoti archaea are generally obligate anaerobes, though Kariarchaeota, Gerdarchaeota and Hodarchaeota may be facultative aerobes (Liu et al. 2021). They are members of the Heimdallarchaeota–Wukongarchaeota branch and exhibit a strong affinity to eukaryotes, suggesting this is the most probable topology (Liu et al. 2021). According to the theory of symbiogenesis, the first common ancestor of eukaryotes was inferred to have been created by an archaea that merged an α -proteobacterium as the mitochondria about 1.5 billion years ago (Martin et al. 2015). This timing aligns with the evolutionary sequence: the oxygen revolution (~2.4 billion years) preceded the evolution of early eukaryotes (~1.6-2.1 billion years) and the acquisition of mitochondria (~1.5 billion years), culminating in the development of meiosis (~1.2 billion years; Speijer et al. 2015).

When the acquisition of mitochondria through endosymbiosis presented early eukaryotic cells, they faced both opportunities and challenges that significantly influenced their evolution (Lane and Martin 2010). Although mitochondria gave cells a lot of energy by making ATP efficiently, which allowed them to develop more complex processes and bigger genomes, they also created a constant source of oxidative stress inside the cells (Lane 2014). The respiratory chain in mitochondria generates reactive oxygen species (ROS) as a byproduct of oxidative phosphorylation, causing ongoing DNA damage (Balaban et al. 2005).

This evolutionary scenario suggests that the acquisition of the mitochondria produces more energy and ROS, which might create a unique cellular inner environment: DNA

damage pressure and efficient energy supply conditions for the cellular process innovation. It might be the drive of the evolution of key eukaryotic traits, including sexual reproduction (meiosis).

Furthermore, eukaryotes exhibit larger cell sizes, multicellularity, and more sophisticated biological processes, whereas prokaryotes possess significantly simpler cell structures.

1.1.2 The Conservation and Variability of Meiotic Sex

For the majority of eukaryotic organisms, meiosis means the occurrence of sexual reproduction. Meiosis is crucial for sexual reproduction, enabling gamete formation and providing the basis for subsequent embryogenesis. The meiotic process keeps the cell's ploidy by cutting in half the diploid set of homologous chromosomes. Gamete fusion then returns the cell to its diploid state. This ensures that the offspring have the same chromosome number as their parents, maintaining the genomic stability of the species. Also, homologous recombination during meiosis is an important part of evolution because it makes it possible for different gene combinations and increases genetic diversity. As the last paragraph showed, meiosis evolved only once and is an important feature of the LECA. This is supported by the phylogenetic analysis of existing eukaryotes. It showed the conservation of the meiosis process; however, the meiotic process also exists with variability.

1.1.2.1 Conservation of Meiosis

Among most eukaryotic lineages, typical meiosis events like homologous chromosome pairing, recombination, and ploidy reduction are very similar. For instance, meiosis exhibits a consistent pattern across various organisms, from unicellular yeast to humans. The first stage of meiosis is prophase I, which includes the duplication of the DNA of progenitor cells. The next stages are meiosis II, which includes prophase II, metaphase II, anaphase II, and telophase II. Homologous recombination is the most important part of meiosis. It happens during prophase I, which is further divided into leptotene, zygotene, pachytene, diplotene, and diakinesis based on how the chromosomes behave. In prophase I, the genetic recombination is highly conserved among the eukaryotes. So, Chi et al. developed the "meiosis toolkit," which is a set of conserved genes that can detect meiosis genetic bases. This toolkit includes 51 genes divided into eleven meiosis-specific and 40 meiosis-related genes, which are also grouped into eight pathways with their

1. Introduction

functions (Chi et al. 2014). Detecting these genes on the genome can potentially ascertain whether an organism can reproduce sexually.

Many of the conserved genes involved in the meiosis toolkit belong to prophase I and serve an important role in chromosome pairing and synaptonemal complex (SC) formation (Chi et al. 2014). *Spo11* causes DNA double-strand breaks (DSBs) on purpose during Leptotene. This is necessary for homologous recombination to start and for chromosome pairing to happen. Additionally, genes such as *REC8* help condense chromosomes and facilitate chromosome pairing. Following breaks produce single-stranded DNA ends that facilitate chromosome pairing via homology recognition. The pairing leads to the formation of the SC, a protein structure that fortifies paired chromosomes and is essential for specific crossover events. After that, the cell goes into Pachytene, a stage where homologous chromosome synapsis and recombination begin. This is when genetic material is exchanged between chromatids that are not sisters. Several important genes that are only found during meiosis are now involved. One of these is *Dmc1*, which is needed for the homolog (nonsister) bias in meiotic recombination. The complex consists of *HOP2* and *MND1*, whose protein products reinforce the binding of *Dmc1* to DNA. *MSH4* and *MSH5* contribute to the stabilization of recombination intermediates; their proteins work as heterodimers (Snowden et al. 2004; Snowden et al. 2008). These genes interact to enable the identification of homologous sequences and enhance strand invasion during recombination, thus ensuring precise alignment and pairing of homologous chromosomes (synapsis).

As the cell progresses to the diplotene stage, the synaptonemal complex dissociates. However, remnants of the complex remain at the chiasmata, the sites of crossing-over, ensuring proper homologous chromosome connections. Additionally, *REC8* plays a crucial role in maintaining cohesion between sister chromatids around the chiasmata, which is essential for accurate chromosome segregation during meiosis. Diakinesis is the final stage of Prophase I, where chromosomes condense further, and the nuclear membrane breaks down as the cell prepares for Metaphase I. During prophase I, genetic recombination occurs, where homologous chromosomes exchange genetic material. During Metaphase I, the nuclear envelope breaks down, and the spindle apparatus captures homologous chromosomes, aligning them at the metaphase plate. In Anaphase I, the previously formed chiasmata dissolve, allowing homologous chromosomes to segregate to opposite poles. Telophase I signifies the end of meiosis I, during which the

1. Introduction

cell divides into two and homologous chromosomes separate, while sister chromatids remain joined until meiosis II. Meiosis II is a faster process with four stages similar to Meiosis I, but it lacks recombination and primarily serves to separate sister chromatids, ultimately producing four gametes.

1.1.2.2 Variability of Meiosis

The genes of meiosis are conserved across eukaryotes, indicating that meiosis originated once in the LECA, but there is also some variability. For example, the regulation and timing of the meiotic process between sexes in mammals is variable (Gao et al. 2024). For instance, males exhibit a comparatively prolonged prophase I and subsequently undergo two rapid divisions, while females experience a shorter prophase I and finalize the first meiotic division during ovulation, with timing contingent upon the onset of sexual maturity.

In addition, the synaptonemal complex (SC) exhibits similar morphology yet demonstrates significant structural variations among organisms and consists of unique proteins across different eukaryotic phyla (Zwettler et al. 2020). For instance, *Zip1* is a structural component of the SC in yeast, promoting synapsis between homologs (Storlazzi et al. 1996). In mammals, *SYCP1* serves a similar function and is critical for SC assembly, which stabilizes paired chromosomes and assists in crossover formation (Vries et al. 2005). However, in flies, the protein called C(3)G replaced the similar function (Page and Hawley 2001).

There are two recombinational pathways in yeast, but only one of them is shared with *C. elegans*, which do not need *MSH4* (Ross-Macdonald and Roeder 1994; Zalevsky et al. 1999; Housworth and Stahl 2003). Besides the variable genes, some ciliate species have lost the SC, like *Tetrahymena thermophila* and *Paramecium tetraurelia*, but they still exhibit crossover without the SC structure (Chi et al. 2014).

1.1.3 Sexual reproduction in old taxa “Protista” Kingdom

Reproduction is an essential biological process of organisms, facilitating the life cycle and producing subsequent generations (Evans 2013; Jalvingh et al. 2016; Minelli and Fusco 2019). While sexual reproduction is the primary and conserved mechanism employed by eukaryotes, emerging as a nearly universal characteristic, its role in protists remains unclear (Goodenough and Heitman 2014). Unlike animals and plants, where sex and

1. Introduction

reproduction are closely linked, most protozoa were considered asexual or facultatively sexual (Hofstatter et al. 2018; Yadav et al.). This bears resemblance to prokaryotes, which primarily facilitate population growth via asexual reproduction.

One major challenge in studying sexual reproduction in protists is that it is often facultative and occurs only under specific environmental conditions (Kooij and Matuschewski 2007; Venugopal et al. 2020). Due to their microscopic size and difficulty in cultivation, direct observations of their complete life cycle, especially sexual stages are rare. The absence of sexual reproduction reports in many protists does not necessarily indicate that they have abandoned it entirely; rather, their sexual life cycle may be triggered only under particular stress conditions, making it difficult to detect.

Some protists exhibit unique adaptations in their reproductive strategies. For example, the malaria parasite undergoes asexual reproduction in both vertebrate and invertebrate hosts but completes its sexual reproduction cycle only in invertebrates (Sinden et al. 1985; Sinden and Hartley 1985; Sinden 1991). Additionally, sexual reproduction in certain protists is influenced by external signals. The EroS protein from *Vibrio fischeri* has been found to induce sexual reproduction in the choanoflagellate *Salpingoeca rosetta*, one of the closest living relatives of animals (Woznica et al. 2017).

Some protists have also evolved highly specialized meiotic mechanisms. The social amoeba *Dictyostelium* has developed a meiosis-like process that functions without the SPO11 protein, which is typically required for initiating double-strand DNA breaks and homologous recombination during prophase I (Bloomfield 2018). These findings indicate that meiotic pathways in protists may have diverged significantly over evolutionary time. Despite these challenges, genomic studies suggest that sexual reproduction in protists may be more widespread than previously assumed (da Silva and Machado 2022; Weedall and Hall 2015). Most eukaryotes share highly conserved meiotic genes, supporting the hypothesis that meiosis is an ancestral trait (Ramesh et al. 2005; Malik et al. 2008; Lenormand et al. 2016; Fu et al. 2019). Advances in sequencing technology have made it possible to assess the presence of meiosis-associated genes in protist genomes, providing the convenience to infer the genetic basis of their potential sexual reproduction. To investigate protist meiosis further, we observed cell fusion under a microscope under starvation conditions and conducted RNA-Seq analysis on samples with and without nutrient deprivation. As ancient representatives of eukaryotes, studying

protist meiosis not only sheds light on the evolution of sex but also offers insights into the broader origins of eukaryotic life.

1.2 Why sex? Evolution of asexuality

Most metazoans reproduce sexually, although various reproductive strategies exist. Compared to the apparent simplicity and efficiency of asexual reproduction (Maynard-Smith 1971; Williams 1975; Smith and Maynard-Smith 1978), sexual reproduction remains dominant despite its disadvantages, such as the production of males, genome dilution and the disruption of advantageous gene combinations (Lively 2011; Lehtonen et al. 2012).

The evolutionary advantages that sustain the dominance of sexual reproduction remain unclear (Hörandl et al. 2020). Despite decades of research, the widespread existence of sexual reproduction, often called the "queen of problems" in evolutionary biology, remains an unresolved mystery (White 1954; Bell 1982; Sharp and Otto 2016; Neiman et al. 2017). Studying rare, long-term asexual species may provide insights into the persistence of sexual reproduction (Birky 1996; Judson and Normark 1996).

1.2.1 The disadvantages of asexual reproduction

Asexual reproduction theoretically results in the permanent accumulation of harmful mutations and diminished adaptability due to the absence of genetic recombination, a phenomenon referred to as Muller's Ratchet in evolutionary biology (Muller 1964; Felsenstein 1974). Over time, this process increases the mutational burden in asexual populations and might cause lethal mutations, making the asexual lineage more vulnerable to extinction. Additionally, the Hill-Robertson (HR) effect further highlights the disadvantages of asexuality. In finite populations, the linkage between sites under selection will reduce the overall effectiveness of selection (Hill and Robertson 1966). In asexual populations, absence of recombination (effectively the whole chromosome is linked) results in increased interference among beneficial and deleterious mutations, thereby leading to a diminished rate of adaptation (Comeron et al. 2008). In contrast, sexual reproduction enhances adaptation by uncoupling linked loci, which allows the removal of deleterious mutations and advantageous mutations to spread more efficiently in the population. Sex thereby increases the evolutionary potential of populations (Otto

2019). This provides a strong explanation for why sexual reproduction remains prevalent in nature, despite its associated costs. Although there are a lot of explanations of advantages of sex, there is still a lack of clear empirical evidence.

Besides mutation accumulation, another popular potential explanation for the predominance of sexual reproduction is the Red Queen Hypothesis, first proposed by Van Valen (Van Valen 2014). This theory suggests that species must continuously evolve to survive in response to coevolving parasites, predators, and competitors. The name comes from Lewis Carroll's *Through the Looking Glass*, in which the Red Queen tells Alice, "It takes all the running you can do to stay in the same place." This metaphor describes how species must constantly adapt just to maintain their existence in a dynamic environment (Van Valen 2014).

Sexual reproduction plays a vital role in this evolutionary arms race by generating genetic diversity, which enables host populations to escape rapidly adapting parasites through recombination (Lively 2010). In contrast, asexual lineages produce genetically identical offspring, making them highly susceptible to parasite attacks. Once a parasite adapts to a dominant asexual genotype, it can quickly infect the entire population, leading to a higher risk of extinction.

Although sex and biological antagonism can also be driven by other mechanisms (Haafke et al. 2016), the connection of the Red Queen and sex seems to have more empirical studies supported (Neiman et al. 2018). Research has shown that parasites exert strong selective pressure on hosts, influencing the prevalence of sexual reproduction. Field studies indicate that asexual reproduction is more common in environments with low parasite diversity, whereas sexual reproduction dominates in regions where parasites rapidly adapt to their hosts (Barbuti et al. 2012; Decaestecker and Bulteel 2019). These findings suggest that parasites can play a crucial role in maintaining sexual reproduction, as recombination provides a key evolutionary advantage in the ongoing battle between hosts and pathogens.

1.2.2 Ancient Asexual Scandals: Exception to the Rule?

Despite these theoretical disadvantages for asexuality, some ancient asexual species have persisted for millions of years, contradicting expectations of inevitable extinction. These organisms, often referred to as "ancient asexual scandals", include bdelloid rotifers, oribatid mites, and darwinulid ostracods (Judson & Normark 1996). Unlike most

1. Introduction

short-term asexual lineages, these species have managed to avoid the detrimental effects on fitness while diversifying and spreading across various ecosystems.

Bdelloid rotifers represent one of the most intriguing cases of putative ancient asexual reproduction, with claims of asexuality spanning over 60 million years (Mark Welch and Meselson 2000). The initial evidence for their long-term asexuality came from the discovery of deeply diverged homologs without recombination, which were attributed to ancient hybridization events (Mark Welch et al. 2008). This understanding was further supported by genomic studies revealing extensive genome rearrangements and translocations that would prevent conventional meiosis (Flot et al. 2013). However, the story became more complex when evidence of genetic exchange emerged through the detection of allele sharing between individuals (Signorovitch et al. 2015). More recent studies have revealed that bdelloid rotifers can engage in non-canonical forms of genetic exchange, though this process differs significantly from traditional meiotic sex (Nowell et al. 2018; Eyres et al. 2022).

Oribatid mites (Acari: Oribatida) represent another remarkable example of ancient asexual reproduction, challenging traditional theories about the evolutionary necessity of sex (Heethoff et al. 2009). Dating back to the Silurian period, these soil-dwelling arthropods have evolved into approximately 10,000 described species, with about 10% being obligately asexual (Norton and Palmer, 1991; Norton et al. 1993; Magilton et al. 2019). Their ecological success is evidenced by their remarkable abundance, reaching up to 350,000 individuals per square meter in soil environments (Maraun and Scheu 2000). Unlike many asexual organisms that recently diverged from sexual ancestors, some oribatid mite lineages belong to entirely asexual clades estimated to have persisted without conventional sex putatively for over 100 million years (Heethoff et al., 2007). This ancient asexuality is particularly intriguing as these mites occasionally produce males with distorted spermatophores, suggesting a long history of asexual reproductive strategies (van der Kooi and Schwander 2014). In addition, the asexual mites (parthenogenesis) are very abundant, like ten times more than in other metazoan taxa (Norton et al. 1993). The evolutionary replicates of parthenogenesis make the comparison of asexuality and their closely related species possible. The long-term survival and diversification of asexual oribatid mites, coupled with their large population sizes, present a significant challenge to conventional theories about the evolutionary

disadvantages of asexual reproduction, making them one of the most compelling "ancient asexual scandals" in the animal kingdom.

1.2.3 The Evolution of asexuality

1.2.3.1 The Meselson Effect in oribatid mites

The extraordinary evolutionary success of ancient asexual oribatid mites challenges our fundamental understanding of the necessity of sexual reproduction. The Meselson Effect was firstly proposed by Birly, and Meselson examined it in rotifers firstly, but rotifers were found to have cryptic sex (Signorovitch et al. 2015). It showed a consequence of the genome for their long-term survival without recombination (Birky 1996; Mark Welch and Meselson 2000).

This effect describes allelic gene copies accumulating mutations independently over time in the absence of recombination, leading to increased genetic divergence between haplotypes in individuals in asexual parallel populations. Recently, a study comparing the population of asexual mites to their close sexual relatives proved the Meselson effect in the asexual oribatid mite *Oppiella nova* which indicated they are long-term asexuals (Brandt et al. 2021).

1.2.3.2 Horizontal gene transfer provide evolutionary substrates for asexuality

Asexual organisms primarily rely on *de novo* mutations as a fundamental source of genetic variation (Barrick and Lenski 2013). However, the frequency and mechanisms of *de novo* mutations remain uncertain, and their occurrence is relatively slow, limiting the adaptive potential of asexual lineages. In contrast, horizontal gene transfer (HGT) presents another source for genetic innovation in asexual organisms (Flot et al. 2013). Unlike vertical inheritance, HGT allows for the direct acquisition of genetic material from other organisms, rapidly introducing new genes and functions into asexual genomes. This process has been particularly beneficial for asexual species that live in close association with diverse microbial communities, facilitating gene uptake (Boto 2014). For example, in bdelloid rotifers, approximately 10% of genes are derived from HGT, contributing to their remarkable adaptability (Gladyshev et al. 2008). Evidence also suggests that HGT provides novel functional traits. A previous study demonstrated that HGT introduced genes for lignocellulose degradation in the asexual

springtail *Folsomia candida* and the root-knot nematode *Meloidogyne incognita* (Abad et al. 2008; Faddeeva-Vakhrusheva et al. 2017). Additionally, recent research has confirmed that horizontally acquired genes can influence host mating behaviors, such as male courtship in lepidopterans (Li et al. 2022). Despite its evolutionary significance, the exact mechanisms by which HGT-derived genes integrate and function within asexual genomes remain poorly understood. Nonetheless, the growing body of evidence suggests that HGT plays a crucial role in overcoming the limitations of asexual reproduction by supplementing genetic diversity and driving evolutionary innovation.

1.2.3.3 Transposable element (TE) activity in asexuals

Transposable elements (TEs), often termed "jumping genes," are mobile genetic sequences present across nearly all organisms, both prokaryotic and eukaryotic, and typically exist in substantial numbers. In eukaryotes, the TEs are also major determinants of genome size and composition, with some species genomes reaching as high as 90% (Schnable et al. 2009; Wang et al. 2014; Meyer et al. 2021). The impact of TEs largely depends on their genomic location; for instance, insertions into intronic regions can alter gene splicing patterns, influencing phenotypes such as male courtship behavior in fruit flies (Ding et al. 2016). Additionally, TEs have been co-opted for essential cellular functions, such as telomere maintenance in *Drosophila*, highlighting their potential function in vital genomic processes (González et al. 2008).

In asexual organisms, TEs were traditionally thought to accumulate due to the absence of meiotic recombination and effective purifying selection, potentially leading to genomic instability. However, recent studies challenge this assumption. Investigations into asexual arthropods, including water fleas, wasps, and oribatid mites, revealed no significant increase in TE load compared to their sexual counterparts (Bast et al. 2016). These findings suggest that asexual organisms may have evolved mechanisms to regulate TE proliferation, possibly through enhanced genome defense systems or TE-mediated DNA repair processes (Gladyshev and Arkhipova 2010; Bast et al. 2019).

Interestingly, TEs may also play a role in the transition to asexuality. In *Daphnia pulex*, a transposon insertion in the gene encoding the meiotic cohesin *Rec8* is associated with obligate asexual reproduction, indicating that TEs can influence reproductive modes by disrupting genes essential for meiosis (Eads et al. 2012). In asexual eukaryotes, such as springtails and nematodes, TE-mediated HGT also plays a crucial role in acquiring

1. Introduction

adaptive traits, such as lignocellulose degradation enzymes (Abad et al. 2008; Faddeeva-Vakhrusheva et al. 2017). In bdelloid rotifers, approximately 10% of their genes are acquired through HGT, many of which are associated with TE activity (Gladyshev et al. 2008). Recently, a study on a kind of TE from fungi called "Starship" showed the TE can transfer large DNA segments between species and make the host acquire new adaptations, including traits like metal resistance and pathogenicity (Urquhart et al. 2024). While TEs have the potential to drive genomic innovation, asexual organisms appear to possess strategies to control TE proliferation, maintaining genomic stability despite the absence of recombination. At the same time, the interplay between TEs and HGT provides an additional source of genetic diversity, offering an evolutionary advantage to asexual lineages. This dual role of TEs, both as potential threats to genome integrity and as facilitators of genetic innovation, highlights their complex evolutionary significance in asexual organisms.

1.3 Aims and Outline of this Study

In this thesis, I provide evidence that sexual reproduction of the protist (*Fisculla terrestris*) can be triggered by starvation and contributes to a better understanding of this cryptic protist and their sexual reproduction within old taxa. Furthermore, sexual reproduction may be more prevalent among eukaryotes, and the reported absence of sexual reproduction in protists could account for their hard and less observation. Additionally, we explored the genomic features (haplotypic independence, one allele conserved and the other relaxed) and potential evolution (genes under positive selection, DEA and HGTs) of ancient asexual oribatid mites (*Platynothrus peltifer*) provide an example of ancient asexual and new insight into the evolution of ancient asexual species. The following three hypotheses will be tested:

- (1) *Fisculla terrestris* exhibits sexual reproduction triggered by starvation.
- (2) *Platynothrus peltifer* is a long-term asexual and the Meselson effect can be detected in the asexual individuals among parallel populations.
- (3) The DEAs and HGT can provide new substrates for the evolution of asexual mites (*Platynothrus peltifer*).

To study the meiosis genes and the transcriptome dynamics of cryptic sex in protist (*Fisculla terrestris*), I constructed a reference transcriptome and detected the meiosis-associated genes using the “meiosis toolkits”. Moreover, I compared the expression change between cells fusion induced by starvation and the cells supplied with food. I discovered that the genes associated with the “recombination” process are up-regulated during cell fusion and combined with microscopic observations, indicating the potential occurrence of cryptic sex. The findings of this study may serve as evidence that sexual reproduction is prevalent among eukaryotes and that it is a characteristic of the last common ancestor of eukaryotes (Chapter 1).

To examine the impact on the genome under long-term asexuality, I assembled the reference genome for an oribatid mite (*Platynothrus peltifer*). The Meselson effect was detected in different populations, which indicated *Platynothrus peltifer* is an ancient

1. Introduction

asexual mite that persisted and evolved over time. The haplotypic divergence and selection analysis showed that one allele is conserved and the other is relaxed, which might protect the genome from the accumulation of harmful mutations and provide evolutionary substrates for the long-term asexual mites. In addition, most DEA and HGT are under relaxed selection, but few are under positive selection, which freely evolved and might provide new traits for the evolution of asexual mites. Our findings provide an example of ancient asexual organisms, and the genomic characteristic may elucidate their evolutionary mechanism. Research on asexual organisms also offers new insights into the implications of sexual reproduction (Chapter 2).

1.4 References of the Introduction

- Abad P, Gouzy J, Aury J-M, Castagnone-Sereno P, Danchin EG, Deleury E, Perfus-Barbeoch L, Anthouard V, Artiguenave F, Blok VC. 2008. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat. Biotechnol.* 26:909–915.
- Balaban RS, Nemoto S, Finkel T. 2005. Mitochondria, Oxidants, and Aging. *Cell* 120:483–495.
- Barbuti R, Mautner S, Carnevale G, Milazzo P, Rama A, Sturmbauer C. 2012. Population dynamics with a mixed type of sexual and asexual reproduction in a fluctuating environment. *BMC Evol. Biol.* 12:49.
- Bast J, Jaron KS, Schuseil D, Roze D, Schwander T. 2019. Asexual reproduction reduces transposable element load in experimental yeast populations. Coop G, Tautz D, Coop G, Charlesworth B, editors. *eLife* 8:e48548.
- Barrick JE, Lenski RE. 2013. Genome dynamics during experimental evolution. *Nat. Rev. Genet.* 14:827–839.
- Bast J, Parker DJ, Dumas Z, Jalvingh KM, Tran Van P, Jaron KS, Figuet E, Brandt A, Galtier N, Schwander T. 2018. Consequences of Asexuality in Natural Populations: Insights from Stick Insects. Wright S, editor. *Mol. Biol. Evol.* 35:1668–1677.
- Bast J, Schaefer I, Schwander T, Maraun M, Scheu S, Kraaijeveld K. 2016. No Accumulation of Transposable Elements in Asexual Arthropods. *Mol. Biol. Evol.* 33:697–706.
- Bell G. 1982. The masterpiece of nature: the evolution and genetics of sexuality. *Routledge*
- Bernstein H, Bernstein C. 2010. Evolutionary Origin of Recombination during Meiosis. *BioScience* 60:498–505.
- Betts HC, Puttick MN, Clark JW, Williams TA, Donoghue PCJ, Pisani D. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* 2:1556–1562.
- Birky CW. 1996. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* 144:427–437.

1. Introduction

- Bloomfield G. 2018. Spo11-Independent Meiosis in Social Amoebae. *Annu. Rev. Microbiol.* 72:293–307.
- Bonifacino JS, Glick BS. 2004. The mechanisms of vesicle budding and fusion. *Cell* 116:153–166.
- Boto L. 2014. Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc. R. Soc. B Biol. Sci.* 281:20132450.
- Brandt A, Schaefer I, Glanz J, Schwander T, Maraun M, Scheu S, Bast J. 2017. Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* 8:873.
- Brandt A, Van PT, 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.*
- Brocks JJ, Nettersheim BJ, Adam P, Schaeffer P, Jarrett AJM, Güneli N, Liyanage T, van Maldegem LM, Hallmann C, Hope JM. 2023. Lost world of complex life and the late rise of the eukaryotic crown. *Nature* 618:767–773.
- Butterfield NJ, Knoll AH, Swett K. 1990. A Bangiophyte Red Alga from the Proterozoic of Arctic Canada. *Science* 250:104–107.
- Chen J-M, Cooper DN, Chuzhanova N, Férec C, Patrinos GP. 2007. Gene conversion: mechanisms, evolution and human disease. *Nat. Rev. Genet.* 8:762–775.
- Chi J, Mahé F, Loidl J, Logsdon J, Dunthorn M. 2014. Meiosis Gene Inventory of Four Ciliates Reveals the Prevalence of a Synaptonemal Complex-Independent Crossover Pathway. *Mol. Biol. Evol.* 31:660–672.
- Comeron JM, Williford A, Kliman RM. 2008. The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* 100:19–31.
- Commoner B, Townsend J, Pake GE. 1954. Free Radicals in Biological Materials. *Nature* 174:689–691.
- Cooper GM. 2000. The Origin and Evolution of Cells. In: Sinauer Associates.
- Da Silva VS, Machado CR. 2022. Sex in protists: A new perspective on the reproduction mechanisms of trypanosomatids. *Genet. Mol. Biol.* 45:e20220065.
- Decaestecker E, Bulteel L. 2019. Parasites opt for the best of both worlds. *eLife* 8:e49615.
- Demoulin CF, Lara YJ, Lambion A, Javaux EJ. 2024. Oldest thylakoids in fossil cells directly evidence oxygenic photosynthesis. *Nature* 625:529–534.

1. Introduction

- Ding Y, Berrocal A, Morita T, Longden KD, Stern DL. 2016. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536:329–332.
- Eads BD, Tsuchiya D, Andrews J, Lynch M, Zolan ME. 2012. The spread of a transposon insertion in Rec8 is associated with obligate asexuality in *Daphnia*. *Proc. Natl. Acad. Sci.* 109:858–863.
- Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG. 2017. Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* 15:711–723.
- Evans TJ. 2013. Chapter 16 - Reproductive Toxicology of Male and Female Companion Animals. In: Peterson ME, Talcott PA, editors. *Small Animal Toxicology* (Third Edition). Saint Louis: W.B. Saunders. p. 167–202.
- Faddeeva-Vakhrusheva A, Kraaijeveld K, Derks MFL, Anvar SY, Agamennone V, Suring W, Kampfraath AA, Eilers J, Le Ngoc G, van Gestel CAM, et al. 2017. Coping with living in the soil: the genome of the parthenogenetic springtail *Folsomia candida*. *BMC Genomics* 18:493.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737–756.
- Flot J-F, Hespeels B, Li X, Noel B, Arkhipova I, Danchin EGJ, Hejnowicz A, Henrissat B, Koszul R, Aury J-M, et al. 2013. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* 500:453–457.
- Fu C, Coelho MA, David-Palma M, Priest SJ, Heitman J. 2019. Genetic and genomic evolution of sexual reproduction: echoes from LECA to the fungal kingdom. *Curr. Opin. Genet. Dev.* 58–59:70–75.
- Gao J, Qin Y, Schimenti JC. 2024. Gene regulation during meiosis. *Trends Genet.*
- Gerschman R, Gilbert DL, Nye SW, Dwyer P, Fenn WO. 1954. Oxygen poisoning and x-irradiation: a mechanism in common. *Science* 119:623–626.
- Gladyshev EA, Arkhipova IR. 2010. Genome Structure of Bdelloid Rotifers: Shaped by Asexuality or Desiccation? *J. Hered.* 101:S85–S93.
- Gladyshev EA, Meselson M, Arkhipova IR. 2008. Massive Horizontal Gene Transfer in Bdelloid Rotifers. *Science* 320:1210–1213.
- González J, Karasov TL, Messer PW, Petrov DA. 2010. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. *PLoS Genet.* 6:e1000905.

1. Introduction

- González J, Lenkov K, Lipatov M, Macpherson JM, Petrov DA. 2008. High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biol.* 6:e251.
- González J, Petrov DA. 2009. The adaptive role of transposable elements in the *Drosophila* genome. *Gene* 448:124–133.
- Goodenough U, Heitman J. 2014. Origins of Eukaryotic Sexual Reproduction. *Cold Spring Harb. Perspect. Biol.* 6:a016154.
- Haafke J, Abou Chakra M, Becks L. 2016. Eco-evolutionary feedback promotes Red Queen dynamics and selects for sex in predator populations. *Evol. Int. J. Org. Evol.* 70:641–652.
- Heethoff M, Norton RA, Scheu S, Maraun M. 2009. Parthenogenesis in Oribatid Mites (Acari, Oribatida): Evolution Without Sex. In: Schön I, Martens K, Dijk P, editors. *Lost Sex. Dordrecht: Springer Netherlands.* p. 241–257.
- Hespeels B, Penninckx S, Cornet V, Bruneau L, Bopp C, Baumlé V, Redivo B, Heuskin A-C, Moeller R, Fujimori A, et al. 2020. Iron Ladies – How Desiccated Asexual Rotifer *Adineta vaga* Deal With X-Rays and Heavy Ions? *Front. Microbiol.*
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8:269–294.
- Hofstatter P, Lahr D. 2025. Ancient asexuality: No scandals found with novel data. *BioEssays* 47:2400227.
- Hofstatter PG, Brown MW, Lahr DJG. 2018. Comparative Genomics Supports Sex and Meiosis in Diverse Amoebozoa. Martin B, editor. *Genome Biol. Evol.* 10:3118–3128.
- Holland HD. 2006. The oxygenation of the atmosphere and oceans. *Philos. Trans. R. Soc. B Biol. Sci.* 361:903–915.
- Housworth EA, Stahl FW. 2003. Crossover Interference in *Humans*. *Am. J. Hum. Genet.* 73:188–197.
- Hörandl E, Bast J, Brandt A, Scheu S, Bleidorn C, Cordellier M, Nowrousian M, Begerow D, Sturm A, Verhoeven K, et al. 2020. Genome Evolution of Asexual Organisms and the Paradox of Sex in Eukaryotes. In: Pontarotti P, editor. *Evolutionary Biology—A Transdisciplinary Approach.* Cham: Springer International Publishing. p. 133–167.
- Javaux EJ. 2023. Eukaryotes, Appearance and Early Evolution of. In: Gargaud M, Irvine

1. Introduction

- WM, Amils R, Claeys P, Cleaves HJ, Gerin M, Rouan D, Spohn T, Tirard S, Viso M, editors. *Encyclopedia of Astrobiology*. Berlin, Heidelberg: Springer. p. 934–939.
- Jabłońska J, Tawfik DS. 2021. The evolution of oxygen-utilizing enzymes suggests early biosphere oxygenation. *Nat. Ecol. Evol.* 5:442–448.
- Jalvingh K, Bast J, Schwander T. 2016. Sex, Evolution and Maintenance of. In: *Encyclopedia of Evolutionary Biology*. Elsevier. p. 89–97.
- Judson OP, Normark BB. 1996. Ancient asexual scandals. *Trends Ecol. Evol.* 11:41–46.
- Kassir Y, Granot D, Simchen G. 1988. IME1, a positive regulator gene of meiosis in *S. cerevisiae*. *Cell* 52:853–862.
- Kasting JF. 1993. Earth's Early Atmosphere. *Science* 259:920–926.
- Kearney MR, Jasper ME, White VL, Aitkenhead IJ, Blacket MJ, Kong JD, Chown SL, Hoffmann AA. 2022. Parthenogenesis without costs in a grasshopper with hybrid origins. *Science* 376:1110–1114.
- Khakhlova O, Bock R. 2006. Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J. Cell Mol. Biol.* 46:85–94.
- Klinger CM, Spang A, Dacks JB, Ettema TJG. 2016. Tracing the Archaeal Origins of Eukaryotic Membrane-Trafficking System Building Blocks. *Mol. Biol. Evol.* 33:1528–1541.
- Kooij TW, Matuschewski K. 2007. Triggers and tricks of Plasmodium sexual development. *Curr. Opin. Microbiol.* 10:547–553.
- Krisko A, Leroy M, Radman M, Meselson M. 2012. Extreme anti-oxidant protection against ionizing radiation in bdelloid rotifers. *Proc. Natl. Acad. Sci.* 109:2354–2357.
- Krupovic M, Dolja VV, Koonin EV. 2023. The virome of the last eukaryotic common ancestor and eukaryogenesis. *Nat. Microbiol.* 8:1008–1017.
- Lane N. 2014. Bioenergetic Constraints on the Evolution of Complex Life. *Cold Spring Harb. Perspect. Biol.* 6:a015982.
- Lane N, Martin W. 2010. The energetics of genome complexity. *Nature* 467:929–934.
- Lehtonen J, Jennions MD, Kokko H. 2012. The many costs of sex. *Trends Ecol. Evol.* 27:172–178.
- Lenormand T, Engelstädter J, Johnston SE, Wijnker E, Haag CR. 2016. Evolutionary mysteries in meiosis. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20160001.
- Li Y, Liu Z, Liu C, Shi Z, Pang L, Chen C, Chen Y, Pan R, Zhou W, Chen X, et al. 2022.

1. Introduction

- HGT is widespread in insects and contributes to male courtship in lepidopterans. *Cell* 185:2975-2987.e10.
- Liu Y, Makarova KS, Huang W-C, Wolf YI, Nikolskaya AN, Zhang X, Cai M, Zhang C-J, Xu W, Luo Z, et al. 2021. Expanded diversity of Asgard archaea and their relationships with eukaryotes. *Nature* 593:553–557.
- Lively CM. 2010. Antagonistic Coevolution and Sex. *Evol. Educ. Outreach* 3:19–25.
- Lively CM. 2011. The cost of males in non-equilibrium populations. *Evol. Ecol. Res.* 13:105–111.
- Lyons TW, Reinhard CT, Planavsky NJ. 2014. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506:307–315.
- Maciver SK, Koutsogiannis Z, de Obeso Fernández del Valle A. 2019. 'Meiotic genes' are constitutively expressed in an asexual amoeba and are not necessarily involved in sexual reproduction. *Biol. Lett.* 15:20180871.
- Magilton M, Maraun M, Emmerson M, Caruso T. 2019. Oribatid mites reveal that competition for resources and trophic structure combine to regulate the assembly of diverse soil animal communities. *Ecol. Evol.* 9:8320–8330.
- Mandel S, Robzyk K, Kassir Y. 1994. IME1 gene encodes a transcription factor which is required to induce meiosis in *Saccharomyces cerevisiae*. *Dev. Genet.* 15:139–147.
- Mark Welch DB, Mark Welch JL, Meselson M. 2008. Evidence for degenerate tetraploidy in bdelloid rotifers. *Proc. Natl. Acad. Sci.* 105:5145–5149.
- Mark Welch DB, Meselson M. 2000. Evidence for the Evolution of Bdelloid Rotifers Without Sexual Reproduction or Genetic Exchange. *Science* 288:1211–1215.
- Marston AL, Amon A. 2004. Meiosis: cell-cycle controls shuffle and deal. *Nat. Rev. Mol. Cell Biol.* 5:983–997.
- Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140330.
- Maynard-Smith J. 1971. The origin and maintenance of sex In Williams GC (Ed.), Group selection (pp. 163–175). Chicago, IL: Aldine Atherton.
- Malik S-B, Pightling AW, Stefaniak LM, Schurko AM, Jr JML. 2008. An Expanded Inventory of Conserved Meiotic Genes Provides Evidence for Sex in *Trichomonas vaginalis*. *PLOS ONE* 3:e2879.
- Melone G, Fontaneto D. 2005. Trophic Structure in Bdelloid Rotifers. *Hydrobiologia*

546:197–202.

- Meyer A, Schloissnig S, Franchini P, Du K, Woltering JM, Irisarri I, Wong WY, Nowoshilow S, Kneitz S, Kawaguchi A, et al. 2021. Giant lungfish genome elucidates the conquest of land by vertebrates. *Nature* 590:284–289.
- Mitchell AP, Herskowitz I. 1986. Activation of meiosis and sporulation by repression of the RME1 product in yeast. *Nature* 319:738–742.
- Minelli A, Fusco G eds. 2019. Introductory Concepts. In: *The Biology of Reproduction*. Cambridge: Cambridge University Press. p. 6–46.
- Moris VC, Bruneau L, Berthe J, Heuskin A-C, Penninckx S, Ritter S, Weber U, Durante M, Danchin EGJ, Hespeels B, et al. 2024. Ionizing radiation responses appear incidental to desiccation responses in the bdelloid rotifer *Adineta vaga*. *BMC Biol.* 22:11.
- Muller HJ. 1964. The relation of recombination to mutational advance. *Mutat. Res. Mol. Mech. Mutagen.* 1:2–9.
- Neiman M, Lively CM, Meirmans S. 2017. Why sex? A pluralist approach revisited. *Trends Ecol. Evol.* 32:589–600.
- Neiman M, Meirmans PG, Schwander T, Meirmans S. 2018. Sex in the wild: How and why field-based studies contribute to solving the problem of sex*: PERSPECTIVE. *Evolution* 72:1194–1203.
- Norton RA, Palmer SC. 1991. The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites. In: Schuster R, Murphy PW, editors. *The Acari: Reproduction, development and life-history strategies*. Dordrecht: Springer Netherlands. p. 107–136.
- Norton RA, Kethley JB, Johnston DE, O'Connor BM, Wrensch DL, Ebbert MA. 1993. Phylogenetic perspectives on genetic systems and reproductive modes of mites.
- O'Malley MA, Leger MM, Wideman JG, Ruiz-Trillo I. 2019. Concepts of the last eukaryotic common ancestor. *Nat. Ecol. Evol.* 3:338–344.
- Otto SP. 2019. Evolutionary potential for genomic islands of sexual divergence on recombining sex chromosomes. *New Phytol.* 224:1241–1251.
- Page SL, Hawley RS. 2001. c(3)G encodes a Drosophila synaptonemal complex protein. *Genes Dev.* 15:3130–3143.
- Philippe H, Forterre P. 1999. The rooting of the universal tree of life is not reliable. *J. Mol. Evol.* 49:509–523.

1. Introduction

- Porter SM. 2020. Insights into eukaryogenesis from the fossil record. *Interface Focus* 10:20190105.
- Ramesh MA, Malik S-B, Logsdon JM. 2005. A Phylogenomic Inventory of Meiotic Genes: Evidence for Sex in *Giardia* and an Early Eukaryotic Origin of Meiosis. *Curr. Biol.* 15:185–191.
- Ricci C. 1998. Anhydrobiotic capabilities of bdelloid rotifers. *Hydrobiologia* 387:321–326.
- Ricci C, Caprioli M, Boschetti C, Santo N. 2005. Macrotrachela quadricornifera featured in a space experiment. *Hydrobiologia* 534:239–244.
- Ross-Macdonald P, Roeder GS. 1994. Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* 79:1069–1080.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112–1115.
- Sharp NP, Otto SP. 2016. Evolution of sex: Using experimental genomics to select among competing theories. *Bioessays* 38:751–757.
- Signorovitch A, Hur J, Gladyshev E, Meselson M. 2015. Allele Sharing and Evidence for Sexuality in a Mitochondrial Clade of Bdelloid Rotifers. *Genetics* 200:581–590.
- Simion P, Narayan J, Houtain A, Derzelle A, Baudry L, Nicolas E, Arora R, Cariou M, Cruaud C, Gaudray FR, et al. 2021. Chromosome-level genome assembly reveals homologous chromosomes and recombination in asexual rotifer *Adineta vaga*. *Sci. Adv.* 7:eabg4216.
- Sinden RE. 1991. Mitosis and meiosis in malarial parasites. *Acta Leiden.* 60:19–27.
- Sinden RE, Hartley RH. 1985. Identification of the meiotic division of malarial parasites. *J. Protozool.* 32:742–744.
- Sinden RE, Hartley RH, Winger L. 1985. The development of Plasmodium ookinetes in vitro: an ultrastructural study including a description of meiotic division. *Parasitology* 91 (Pt 2):227–244.
- Smith JM, Maynard-Smith J. 1978. The evolution of sex. *Cambridge University Press Cambridge*
- Snowden T, Acharya S, Butz C, Berardini M, Fishel R. 2004. hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Mol. Cell* 15:437–451.
- Snowden T, Shim K-S, Schmutte C, Acharya S, Fishel R. 2008. hMSH4-hMSH5

1. Introduction

- Adenosine Nucleotide Processing and Interactions with Homologous Recombination Machinery*. *J. Biol. Chem.* 283:145–154.
- Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* 16:472–482.
- Speijer D. 2016. What can we infer about the origin of sex in early eukaryotes? *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150530.
- Speijer D, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci.* 112:8827–8834.
- Storlazzi A, Xu L, Schwacha A, Kleckner N. 1996. Synaptonemal complex (SC) component Zip1 plays a role in meiotic recombination independent of SC polymerization along the chromosomes. *Proc. Natl. Acad. Sci.* 93:9043–9048.
- Tashiro T, Ishida A, Hori M, Igisu M, Koike M, Méjean P, Takahata N, Sano Y, Komiya T. 2017. Early trace of life from 3.95 Ga sedimentary rocks in Labrador, Canada. *Nature* 549:516–518.
- Urquhart AS, Gluck-Thaler E, Vogan AA. 2024. Gene acquisition by giant transposons primes eukaryotes for rapid evolution via horizontal gene transfer. *Sci. Adv.* 10:eadp8738.
- van der Kooi CJ, Schwander T. 2014. On the fate of sexual traits under asexuality. *Biol. Rev.* 89:805–819.
- Van Valen L. 2014. 19. A New Evolutionary Law (1973). In: *Foundations of Macroecology*. University of Chicago Press. p. 284–314.
- Venugopal K, Hentzschel F, Valkiūnas G, Marti M. 2020. Plasmodium asexual growth and sexual development in the haematopoietic niche of the host. *Nat. Rev. Microbiol.* 18:177–189.
- Villeneuve AM, Hillers KJ. 2001. Whence Meiosis? *Cell* 106:647–650.
- Vosseberg J, van Hooff JJE, Köstlbacher S, Panagiotou K, Tamarit D, Ettema TJG. 2024. The emerging view on the origin and early evolution of eukaryotic cells. *Nature* 633:295–305.
- Vries FAT de, Boer E de, Bosch M van den, Baarends WM, Ooms M, Yuan L, Liu J-G, Zeeland AA van, Heyting C, Pastink A. 2005. Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes Dev.* 19:1376–1389.
- Wang Xianhui, Fang X, Yang P, Jiang X, Jiang F, Zhao D, Li B, Cui F, Wei J, Ma C, et al.

2014. The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* 5:2957.
- Ward A, Hopkins J, McKay M, Murray S, Jordan PW. 2016. Genetic Interactions Between the Meiosis-Specific Cohesin Components, *STAG3*, *REC8*, and *RAD21L*. *G3* Bethesda Md 6:1713–1724.
- Weedall GD, Hall N. Sexual reproduction and genetic exchange in parasitic protists | *Parasitology*. Camb. Core
- White MJD. 1954. Animal cytology & evolution. *Cambridge university press*
- Wilkins AS, Holliday R. 2009. The Evolution of Meiosis From Mitosis. *Genetics* 181:3–12.
- Williams GC. 1975. Sex and Evolution. (MPB-8), Volume 8. *Princeton University Press*
- Woese CR, Fox GE. 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci.* 74:5088–5090.
- Woznica A, Gerdt JP, Hulett RE, Clardy J, King N. 2017. Mating in the Closest Living Relatives of Animals Is Induced by a Bacterial Chondroitinase. *Cell* 170:1175-1183.e11.
- Yadav V, Sun S, Heitman J. On the evolution of variation in sexual reproduction through the prism of eukaryotic microbes. *Proc. Natl. Acad. Sci. U. S. A.* 120:e2219120120.
- Zalevsky J, MacQueen AJ, Duffy JB, Kempfues KJ, Villeneuve AM. 1999. Crossing Over During *Caenorhabditis elegans* Meiosis Requires a Conserved MutS-Based Pathway That Is Partially Dispensable in Budding Yeast. *Genetics* 153:1271–1283.
- Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, et al. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541:353–358.
- Zwettler FU, Spindler M-C, Reinhard S, Klein T, Kurz A, Benavente R, Sauer M. 2020. Tracking down the molecular architecture of the synaptonemal complex by expansion microscopy. *Nat. Commun.* 11:3222.

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

The following chapter consists of a manuscript: “Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*”.

Manuscript was published on ***BMC Biology***

The CRediT (Contributor Roles Taxonomy) is employed to clarify the contribution of the authors.

Authors:

Shan Gao, Marcel Dominik Solbach, Jens Bast, Kenneth Dumack

First author

Author contributions of this thesis:

Formal analysis: Performed all of the bioinformatics analysis.

Writing - original draft: Wrote the original draft and improved the manuscript with K.D. and J.B..

Writing - review & editing: Respond to review comments and revise the article with all authors.

Visualization: All figures in the main text and supplementary, except Figure 1 in the main text.

Contribution of other authors

J.B. and K.D. conceptualized and supervised the study; K.D. and J.B. provided funding;

K.D. and M.D.S. did wet lab work.

RESEARCH

Open Access



Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

Shan Gao¹, Marcel Dominik Solbach¹, Jens Bast^{1†} and Kenneth Dumack^{1†*}

Abstract

Background Unicellular eukaryotes were widely considered to reproduce without sex. However, recent findings suggest that meiosis, and by extension (sometimes cryptic) sexual reproduction, might be present in almost all eukaryotic lineages.

Results Here, we investigate the transcriptomic response underlying starvation-induced fusion in the Rhizaria protist *Fisculla terrestris*. Investigations of differentially expressed genes (DEGs) with a particular focus on the expression of meiosis-associated genes suggest that some form of meiosis and recombination might occur in these Rhizaria.

Conclusions We showed that starvation triggered changes in gene expression of meiosis-associated genes in *F. terrestris*. However, if these processes are coupled with sexual reproduction remains to be investigated.

Keywords Meiosis, Protist, Rhizaria, Differentially expressed genes (DEGs), Meiosis toolkits, Cell fusion, Recombination

Background

Meiotic sex, the reciprocal genetic exchange between individuals, including the fusion of nuclei, recombination, and segregation of chromosomes, was commonly assumed to be restricted almost entirely to animals, fungi, and plants [1]. The vast majority of unicellular eukaryotes, i.e., protists, were considered to be asexual; however, recent findings suggest the existence of some form of sex. Animals, fungi (both Amorphea), and plants (Archaeplastida) are unrelated lineages nesting in the large diversity of eukaryotes (Fig. 1) [2]. The apparent polyphyly of sexual macroorganisms and the conserved manner of eukaryotic sex (via meiosis) suggest an early evolutionary origin of sexual reproduction within

eukaryotes. Intriguing is the conservancy and homology of eukaryotic meiosis-associated (including “Meiosis-specific genes” and “Meiosis-related genes” from the meiosis toolkit) enzymes to recombination enzyme machinery of prokaryotes [3–5] and the finding of archaeal cells that undergo fusion coupled with recombination, a process being very similar to eukaryotic sex [6]. It is therefore commonly suggested that sex evolved with the eukaryotic common ancestor (that derived in some form from an archaeal ancestor) [1, 7]. Nonetheless, it is important to note that there is considerable ascertainment bias in our understanding of sex across eukaryotic lineages, as many protistan groups remain understudied.

While sexual reproduction has been well-documented in plants, animals, and fungi, there is growing evidence for its occurrence across a wide range of eukaryotic lineages [9]. This evidence comprises observations of the fusion of cells, the fusion of nuclei (karyogamy), certain chromosomal behaviors that were interpreted as meiosis, and subsequent fission [10–12]. Although fusion in protists is widely observed, the entire sequence of putative

[†]Jens Bast and Kenneth Dumack contributed equally to this work.

*Correspondence:

Kenneth Dumack

kenneth.dumack@uni-koeln.de

¹ Institute for Zoology, University of Cologne, Zulpicher Str. 47b, Cologne 50674, Germany



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

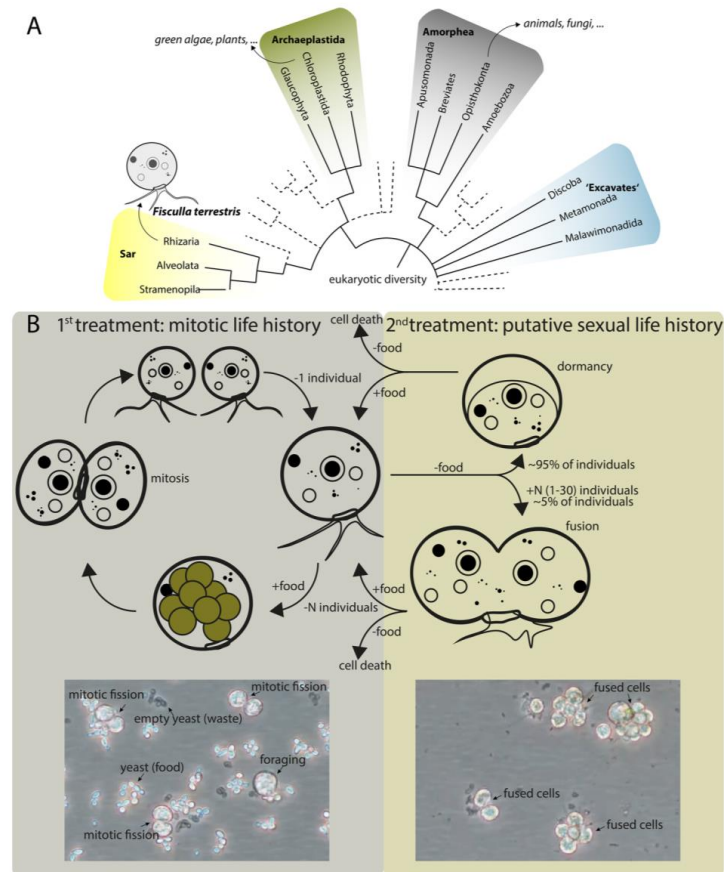


Fig. 1 Evolutionary relationship and life history of the used organism. **A** *F. terrestris* is a representative of the Rhizaria, being only distantly related to the well-known eukaryotes like plants (Archaeplastida) and animals or fungi (Amorphea). The phylogenetic tree is based on Burki et al. [8] and only the most important eukaryotic groups are indicated by names; other branches are reduced. **B** The life history of *F. terrestris* is divided into the two treatments used for the differential genes expression (DGE) experiment, i.e., predominantly mitotically growing, non-starved individuals (left, i.e., the non-starved treatment) and predominantly starved individuals comprising dormant and fused individuals (right, i.e., the starved treatment)

sexual behaviors could so far only be described in very few taxa. Accordingly, it proved difficult to gather direct evidence for sex in protists [13] and attention of the research field shifted to the accumulation of indirect evidence. Currently, the most common approach to investigate sex in protists is thus screening for genes that are considered to be involved in meiosis [5, 14–17]. With this, diverse groups of protists have been shown to possess meiosis-associated genes [4, 16–18]. Since these genes seem to be conserved in almost all eukaryotic lineages, it

was concluded that sex must be widely distributed, cryptic, and not necessarily coupled with offspring production [19–21]. However, recently, it was questioned to rely solely on the detection of these meiosis-associated genes as evidence for sex, as these genes possibly exert pleiotropic functions [22, 23].

The Thecofilosea (Rhizaria, protists) are only very distantly related to the well-known sexual and macroscopic eukaryotes (Fig. 1A). Thecofilosea species can undergo a complete life history under laboratory conditions,

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

including fusion, karyogamy (fusion of nuclei), and fission [11, 24]. *Fisculla terrestris* (Thecofilosea) is a unicellular, shell-bearing amoeba living in soils where it feeds predominantly on fungi (yeasts and spores) and algae. In *Fisculla terrestris*, the depletion of prey reliably triggers the fusion of cells and karyogamy, a subsequent addition of prey (or its supernatant) induces the separation of fused individuals (Fig. 1B). This process seems reminiscent of “cyclical asexuality,” i.e., the ability to switch to sexual life cycles under unfavorable conditions, while otherwise reproducing asexually [25, 26]. To generate insights into potentially occurring meiosis and sex in Thecofilosea, we identify the expression response to starvation-induced fusion in *Fisculla terrestris* with a specific focus on meiosis-associated genes.

Results

Experimental setup and reference transcriptome assembly

In many eukaryotes, starvation triggers a switch from an asexual to a sexual reproductive cycle [27]. In the cultures used here, *F. terrestris* fed, digested, and underwent typical mitotic fission, when food (*Saccharomyces cerevisiae*) was plentiful (i.e., in the “non-starved treatment”; Fig. 1B, left). The majority of *F. terrestris* individuals entered dormancy, i.e., a state of immobility and low activity, when food was depleted (i.e., “starved treatment”). However, under starvation, about 5% of individuals fused into large aggregates with a huge, very active pseudopodial mass (Fig. 1B, right).

To identify the molecular signatures associated with fusion (and potential meiosis-facilitated sexual reproduction), we extracted RNA from five replicated cultures under each treatment, i.e., non-starved and starved. Next, we combined all data and assembled a de novo reference transcriptome. In total, 60 GB of raw data were obtained, resulting in 43,531 assembled transcripts in the reference transcriptome, of which 26,525 remained after quality filtering, i.e., selecting for longest isoform length, duplicate, and redundancy removal. Then, we mapped the reads of the respective samples back to the reference transcriptome. Transcription expression patterns were highly similar among replicates, whereas patterns were clearly different between treatments (Fig. 2A).

Differentially expressed enrichments

To estimate functional changes in expression patterns between treatments, we determined 16,473 differentially expressed genes (DEGs), of which 15,327 were found to be protein-coding. Protein-coding DEGs were used to identify specific pathways associated with either treatment. For this, protein-coding DEGs were annotated and subjected to pathway over-representation, gene set enrichment analyses, and gene ontology (GO) term

enrichment analysis. As annotation of protistan genomic data is challenging, these DEGs were annotated with both, the human KOBAS database and a universal eukaryote eggNOG-mapper database, obtaining 3497 and 3577 annotated genes respectively of which 3093 were shared [28–30]. The pathway over-representation analysis revealed DEGs assigned to numerous pathways (Fig. 2B), including pathways like “Cell cycle,” “Oocyte meiosis,” and “DNA replication.” The gene set enrichment analysis showed similar results but further indicated enrichment in sulfur-related metabolism, including “Glutathione” and “Cysteine and methionine metabolism” pathways. To identify specifically upregulated DEGs between the two treatments, we then analyzed the GO term enrichment. A total of 8345 genes (of which 1929 could be annotated) were significantly upregulated in the non-starved treatment and 8128 genes (of which 1648 could be annotated) in the starved treatment. In total, 300 GO terms and 23 Kyoto Encyclopedia of Genes and Genomes (KEGG) terms with differentially expressed genes were obtained for the non-starved treatment, and 112 GO terms and 7 KEGG terms with differentially expressed genes were obtained for the starved treatment. Subsequently, to elucidate the biological significance represented by numerous GO terms, the GO terms were simplified. Fourteen and nine simplified groups were obtained for the non-starved (Fig. 2C) and starved treatment (Fig. 2D). In the non-starved treatment, upregulated simplified groups included terms such as “Mitotic,” “Cell division,” “Lamellipodium,” and “Metabolic,” indicating mitotic division, rapid growth, and pseudopodia-facilitated predation, as expected for the treatment and consistent with the observed behaviors of cells (Fig. 2C). In the starved treatment, we observed an upregulation of terms such as “Cell adhesion,” “RNA,” “Response,” and “Macropinocytosis,” supporting the expected transcriptomic response to fusion, starvation stress, and an increased communication. The upregulation of macropinocytosis-related genes might indicate a non-selective uptake of remaining environmental nutrients as a stress response [31–33].

Differential expression of meiosis-associated transcriptions

Lastly, genes that are typically associated with meiosis in species with sexual reproduction were identified in *F. terrestris* and analyzed for changes in their expression. A previously published list of meiosis-associated genes comprised 11 meiosis-specific and 40 meiosis-related genes [7]. Here, a total of 41 out of 51 meiosis-associated genes were detected, including eight meiosis-specific genes and 33 meiosis-related genes (Fig. 3; Additional file 1: Table S1). The ten missing transcriptions in *F. terrestris* included *MSH4*, *ZIP1*, *RED1*, *LIF1*, *REC114*, *BRCA1*, *BRCA2*, *YEN1*, *MEC1*, and *RAD52*.

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

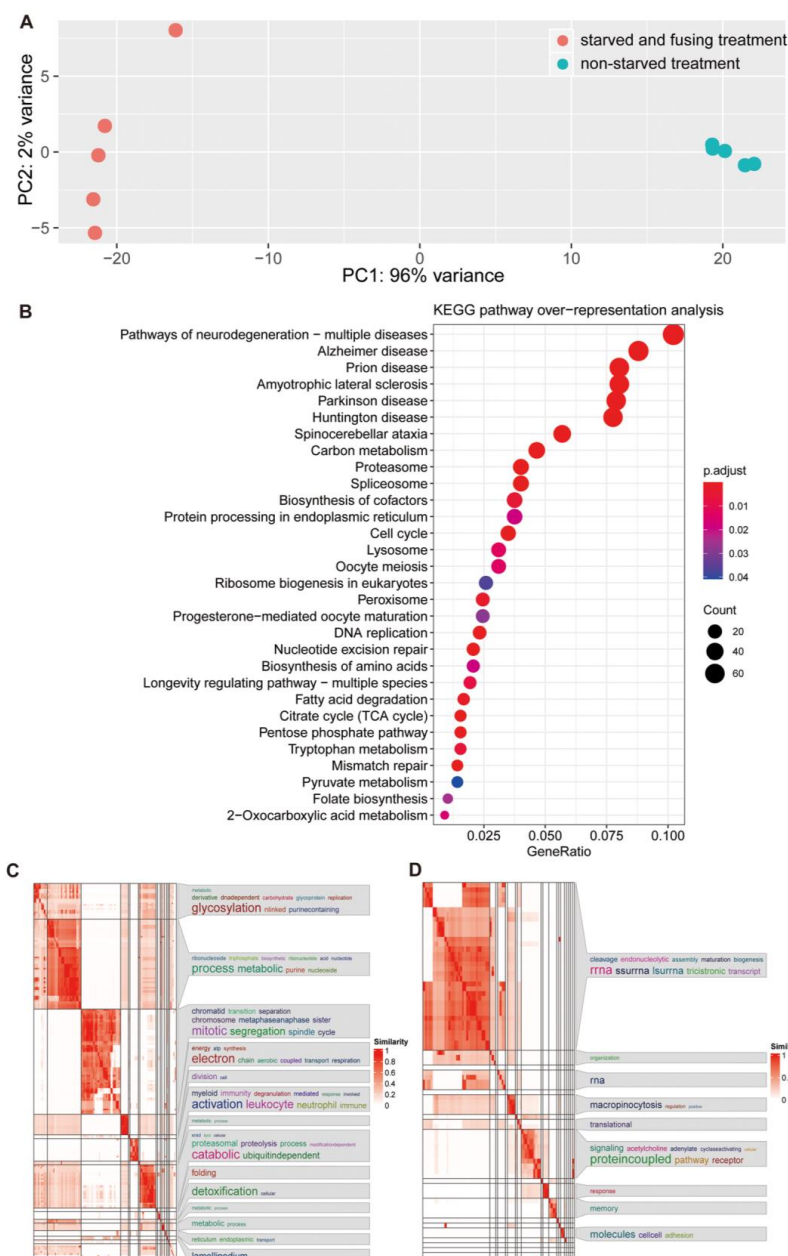


Fig. 2 Overview of the differential gene expression data. **A** Principal component analysis (PCA) based on the expression level of all transcripts for each replicate included in the experiment. Both investigated life history stages exhibit clearly distinct expression patterns and variation within treatments was low. **B** Dot plot of KEGG pathway over-representation analysis of total DEGs. Some meiosis-associated pathways were enriched. The changed cellular functions (GO terms) in each treatment (**C** non-starved treatment, **D** starved treatment)

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

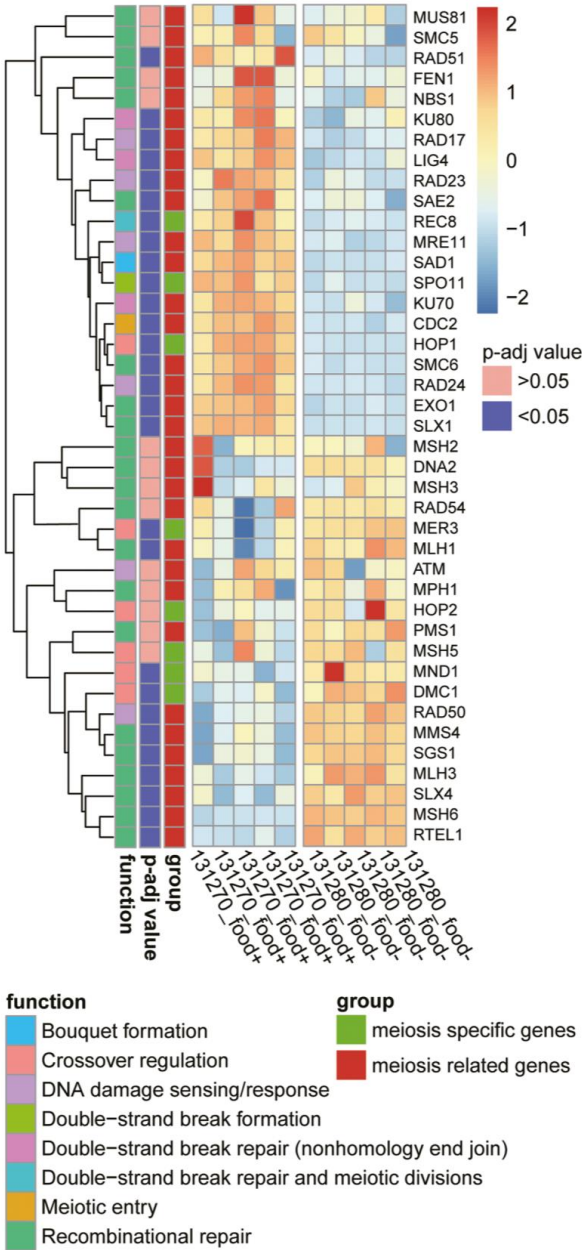


Fig. 3 Overview of the differential expression of 41 meiosis-associated genes, their functions, assignment, and expression. Numerous genes are up- or downregulated between treatments. The left dendrogram of gene expression clustering shows the separation of the ten samples and two treatments

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

Among the eight meiosis-specific genes, three (*DMC1*, *MND1*, and *MER3*) were upregulated and three (*SPO11*, *HOP1*, and *REC8*) were downregulated during starvation. The three upregulated genes function in crossover regulation. The downregulated genes, *REC8* forms the bouquet, *HOP1* regulates crossover formation, and *SPO11* forms the double-strand break during meiosis. The meiosis-specific genes *HOP2* and *MSH5* did not exhibit significant expression changes between treatments [5].

Additionally, we identified 22 meiosis-related DEGs, of which 8 were upregulated and 14 were downregulated during starvation. All upregulated genes contribute to recombinational repair, except for *RAD50*, which functions in DNA damage sensing/response. The downregulated DEGs fall into five distinct functional groups: five genes (*EXO1*, *SLX1*, *SMC6*, *RAD51*, and *SAE2*) fulfill functions in recombinational repair; four genes (*RAD17*, *RAD24*, *RAD23*, and *MRE11*) are involved in DNA damage sensing/response; three genes (*LIG4*, *KU70*, and *KU80*) are assigned to double-strand break repair (non-homologous end joining); *CDC2* indicates meiotic entry; and *SAD1* is involved in bouquet formation. Eleven meiosis-related genes did not change significantly in expression intensity between treatments [5].

Discussion

F. terrestris expresses meiosis-associated genes during starvation

We were able to determine 41 of the initially surveyed 51 meiosis genes. Twenty-eight of the 41 detected meiosis-associated genes were shown to be differentially expressed. Seventeen of these genes were downregulated during starvation. Downregulated genes include predominantly genes that are thought to be involved in five distinct meiosis-related processes: (1) double-strand break and its repair (*SPO11*, *REC8*, and the three genes *LIG4*, *KU70*, and *KU80*, both being involved in nonhomology end joining integration of DNA); (2) five genes that are reported to function in recombinational repair (*RAD51*, *SAE2*, *SLX1*, *EXO1*, and *SMC6*); (3) four genes belong to DNA damage sensing/response function, including *RAD17*, *RAD23*, *RAD24*, and *MRE11*; (4) *CDC2* indicating “Meiotic entry”; and (5) *HOP1* regulates crossover. Specifically, *SPO11* produces DNA double-strand breaks (DSBs), which are necessary for meiotic recombination [34]. According to studies in yeast, the expression of *SPO11* peaks at meiotic entry and then gradually declines [35]. Starving yeast was shown to induce genome-wide meiotic *SPO11*-dependent double-strand breaks, and then return to mitotic growth when food was resupplied indicating a temporal separation and termination of initiated meiosis, a process known as “Return To Growth” (RTG) [36, 37]. Considering that fused *F. terrestris* can

separate when food is resupplied (Fig. 1B, right), a temporal separation of meiosis and an RTG-like process in protists like *F. terrestris* is thinkable, potentially explaining the mixed response of meiosis-associated genes in our analysis.

Most upregulated meiosis-associated genes during fusion of *F. terrestris* are known to be involved in crossover regulation (*DMC1*, *MER3*, and *MND1*) and recombinational repair (*MLH1*, *MMS4*, *MSH6*, *MLH3*, *SGS1*, *SLX4*, and *RTEL*). Upregulated *RAD50* is related to DNA damage sensing/response. Among the upregulated genes, including the equally expressed *HOP2*, *DMC1* is expected to be crucial for entering meiosis as it promotes the formation of strand-invasion products (D-loops) between homologous molecules [27]. *MND1* and *HOP2* stabilize the binding of *DMC1* and DNA [38–40]. Thus, their simultaneous activity suggests active recombination between homologues. The last upregulated meiosis-specific gene, the *MER3* helicase is a member of the ZMM protein group (also known as the synapsis initiation complex = SIC), facilitating the formation of the majority of crossovers during meiosis [41, 42]. Based on the differential expression of said genes, we hypothesize that *F. terrestris* undergoes double-strand breaks and homologous recombination during fusion.

We were not able to detect the expression of three meiosis-specific transcriptions (*MSH4*, *ZIP1*, and *RED1*) and seven meiosis-related transcriptions (*LIF1*, *REC114*, *BRCA1*, *BRCA2*, *YEN1*, *MEC1*, and *RAD52*), indicating their absence. These transcriptions may either not have been expressed during our experiment or may have been lost in *F. terrestris*. Many members of the SAR clade seem to miss some or all of those genes, such as *MSH4* [5, 43]. *ZIP1* and *RED1* are both involved in the synaptonemal complex, which is not strictly needed for meiosis in many taxa [43]. The common absence of these genes from transcriptomic or genomic data in the SAR clade might indicate a loss of those genes in its common ancestor (consisting of the Stramenopiles, Alveolata, and Rhizaria, see Fig. 1A). However, the absence of meiosis-associated genes does not necessarily confirm an absence of meiosis. For example, *Caenorhabditis elegans* and *Drosophila melanogaster* are missing a whole series of meiosis-related genes and evidently undergo meiosis nonetheless as these taxa exclusively using crossover pathway 1 [44, 45]. The same applies to the protistan Alveolata (in SAR group) for which there is plenty of evidence for sex, but an absence of numerous meiosis-associated genes [5].

Sulfur availability may trigger fusion and recombination in *F. terrestris*

Under starvation, the amoeba *Dictyostelium discoideum* (Amoebozoa, Fig. 1A) shifts from a unicellular

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

to a multicellular life history stage and forms sexually derived spores [46, 47]. During extensive research on this model organism, it was found that harmful reactive oxygen species (ROS) form under starvation. As a response, the amino acid cysteine is sequestered in the antioxidant glutathione. Altogether, this links sulfur availability to fusion and sex in this protist [48]. Although only distantly related, we found the down-regulation of the pathways “Cysteine and methionine metabolism” and “Glutathione metabolism” indicating this to be a conserved eukaryotic response. Even in bacteria, sulfur metabolism regulates cell cycle progression via shifts in glutathione levels [49].

Conclusions

The detection of sexual processes in protists is generally quite challenging because of their small size, restricted culture conditions, and limited access to genetic material. A number of studies interpreted the presence of conserved meiosis genes as evidence for sex [50]. As discussed, such conserved genes are often pleiotropic and meiosis can be present in asexual organisms [21–23, 51]. Consequently, the presence of meiosis genes alone is of limited reliability for detecting meiosis and sex. Here, to alleviate some of the mentioned challenges, we induced fusion in *F. terrestris* and analyzed the transcriptional response with a specific focus on meiosis-associated genes and processes. Based on the findings here, the genes that likely regulate crossover and repair recombination are upregulated, so meiosis might be active, but temporal dynamics and RTG-like processes could confound meiosis-associated expression dynamic patterns which cannot be ruled out. Ultimately, to clearly elucidate the occurrence and role of meiosis, genetic recombination, and sexual processes in Rhizaria, further studies are needed analyzing, e.g., fine-scale cytological processes during fusion and/or population genetic outcomes of sex such as allele sharing among individuals. It is important to note that while meiosis, genetic recombination, and sexual reproduction are often associated, they are not universally linked across all eukaryotic groups. For instance, in ciliates, meiosis and genetic exchange occur without being directly tied to reproduction. This highlights the need for careful distinction during the investigation of these processes in diverse eukaryotic lineages, including Rhizaria. *Fisculla terrestris* is a well-suited candidate to further investigate putative sexual reproduction in distant relatives of animals and plants. Cultures of *F. terrestris* can be easily established and maintained, fusion can be easily induced, and its slow and transparent cells facilitate microscopical investigation.

Methods

Experimental setup and transcriptome sequencing

F. terrestris was cultured in Waris-H+Si (McFadden & Melkonian) with *Saccharomyces cerevisiae* as food in 80 ml batch cultures. Under these culture conditions, the highest density of *F. terrestris* cells was expected after 5 days. We, therefore, harvested five replicates for transcriptome sequencing after 4 days of culture growth when food was declining in abundance, and another set of five replicates was harvested after an additional 3 days of growth, when food was depleted and a large number of cells were fused or dormant. RNA was extracted with the Qiagen RNeasy plant mini kit. For this, the culture flasks were vigorously shaken to detach amoeba cells from the plastic surface, and approx. 80 ml of each culture was filtered with an 8- μ m pore size filter to maximize the density of *F. terrestris* while minimizing contamination by environmental bacteria and *S. cerevisiae*. The filter was given into a 1.5-ml Eppendorf tube with 1 ml of ice-cold Soerensen buffer and vortexed vigorously to detach the cells from the filter. Subsequently, the Eppendorf tube was centrifuged at 1000 rpm for 5 min at 4 °C and the filter was removed without disturbing the pellet. The Eppendorf tube was centrifuged at 1000 rpm for 2 min and 4 °C to firm the pellet. The Soerensen buffer was discarded and replaced by 1 ml of clean new buffer. The tube was centrifuged at 1000 rpm for 5 min and 4 °C to firm the pellet. The Soerensen buffer was discarded and replaced by 170 μ l ice-cold RLN buffer. The tube was vortexed, and 450 μ l RLT buffer of the Qiagen RNeasy plant mini kit was added and mixed by vortexing. The tubes were kept in liquid nitrogen until following the rest of the manual's instructions for the Qiagen RNeasy plant mini kit. Poly-A selection, cDNA synthesis, paired-end library preparation, and subsequent Illumina NovaSeq 6000 sequencing were conducted in the West German Genome Center.

Reference transcription assembly

Sequence reads were trimmed with TrimGalore (version v0.6.5, <https://github.com/FelixKrueger/TrimGalore>) before assembly with Trinity (version v2.1.1) [52]. We used the default parameter for assembly and then used Trinity's “get longest isoform seq per trinity gene.pl” to select for the longest transcripts. The CD-HIT (version v4.8.1) [53] was employed to select the unique transcripts according to a similarity criterion of 90% and reads with a length of <300 bp were removed by Seqkit (version 0.14.0) [54]. Translation of transcripts was performed by Transdecoder (<https://github.com/TransDecoder/TransDecoder/wiki>; last accessed November 14, 2018) to establish the protein data set used for further analyses.

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

Contamination removal and quality assessment

To identify and remove putative contamination by the food organism *Saccharomyces cerevisiae*, we downloaded its genome and related protein sequences (ASM308665v1). We used Minimap2 (version v2.1) [55] and Blastp [56] to screen for the presence of potential yeast sequences. In addition, we used the BlobTools2 kit (version v2.3.3) [57] to check the results made by Busco (version 5.0.0) [58] and Blastp. Most of the transcripts (29M/30M) were no-hit in the nt database and the Busco results showed a completeness of 77.3% against the eukaryota_obd10 database.

Differentially expressed genes detection

To identify and quantify the DEGs, we used RSEM (version v1.3.3) [59] to map the filtered reads to the reference transcriptome with parameters “--est_method RSEM --aln_method bowtie2 --trinity_mode --prep_reference.” DEGs were determined with DeSeq2 (version v. 1.10.1) [60]. Transcripts with <50 reads among all samples were discarded and we only kept DEGs with an adjusted *p* value of less than 0.05.

Annotation and enrichment of DEGs

The longest resulting protein sequences were used for annotation. DEGs were annotated with KOBAS (version v. 3.0) [29] and its human database and eggNOG-mapper (version v2.1.6) [30]. Subsequently, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and gene ontology (GO) enrichment analysis and gene set enrichment analysis (GSEA) were carried out using clusterProfiler (version v. 3.15) [61]. The protein sequences derived from DEGs were submitted to the human database, and hypergeometric test/Fisher's exact test were used to calculate the corrected *p* values.

Meiosis-associated genes Blast

Lastly, we compared our genes to 11 meiosis-specific and 40 meiosis-related genes published before [5, 50]. For this, we aligned these sequences to our protein database. Aligned sequences with an *e* value <10⁻⁴ were kept and then manually curated.

Abbreviations

DEGs	Differentially expressed genes
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
DSBs	Double-strand breaks
RTG	Return to growth
SIC	Synapsis initiation complex
ROS	Reactive oxygen species
GSEA	Gene set enrichment analysis
PCA	Principal component analysis

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-025-02246-3>.

Additional file 1: Table S1 DESeq2 expression level of the 41 DEGs showed the raw expression.

Acknowledgements

This work was funded by the core funding of J.B.

Authors' contributions

Conceptualization: K.D.; investigation: K.D. and M.D.S.; formal analysis: S.G.; writing – original draft: S.G., J.B. and K.D.; writing – review & editing: all authors; visualization: S.G. and K.D.; funding acquisition: J.B.

Funding

Open Access funding enabled and organized by Projekt DEAL. This work was funded by the core funding of J.B.

Data availability

The raw data is available at NCBI Bioproject PRJNA1108686 [62].

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 24 May 2024 Accepted: 13 May 2025
Published online: 22 May 2025

References

1. Garg SG, Martin WF. Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biol Evol*. 2016;8:1950–70.
2. Adl SM, Bass D, Lane CE, Lukeš J, Schoch CL, Smirnov A, et al. Revisions to the classification, nomenclature, and diversity of eukaryotes. *J Eukaryot Microbiol*. 2019;66:4–119.
3. Camerini-Otero RD, Hsieh P. Homologous recombination proteins in prokaryotes and eukaryotes. *Annu Rev Genet*. 1995;29:509–52.
4. Ramesh MA, Malik S-B, Logsdon JM. A Phylogenomic inventory of meiotic genes: evidence for sex in giardia and an early eukaryotic origin of meiosis. *Curr Biol*. 2005;15:185–91.
5. Chi J, Mahé F, Loidl J, Logsdon J, Dunthorn M. Meiosis gene inventory of four ciliates reveals the prevalence of a synaptonemal complex-independent crossover pathway. *Mol Biol Evol*. 2014;31:660–72.
6. Bernstein H, Bernstein C. Sexual communication in archaea, the precursor to eukaryotic meiosis. In: Witzany G, editor. *Biocommunication of archaea*. Cham: Springer International Publishing; 2017. p. 103–17.
7. López-García P, Moreira D. The syntrophy hypothesis for the origin of eukaryotes revisited. *Nat Microbiol*. 2020;5:655–67.
8. Burki F, Roger AJ, Brown MW, Simpson AGB. The new tree of eukaryotes. *Trends Ecol Evol*. 2019;35:43–55.
9. Hörandl E, Bast J, Brandt A, Scheu S, Bleidorn C, Cordellier M, et al. Genome evolution of asexual organisms and the paradox of sex in eukaryotes. In: Pontarotti P, editor, et al. *Evolutionary biology—a trans-disciplinary approach*. Cham: Springer International Publishing; 2020. p. 133–67.
10. Bell G. Sex and death in protozoa: the history of obsession. Cambridge: Cambridge University Press; 1988.

2. Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

11. Bélař K. Protozoenstudien. III. Arch Protistenk. 1921;43:431–62.
12. Valkanov A. *Paracadra madarica* n. sp. (Rhizopoda-Testacea) und ihre Kopulation. Dokl Bolg Akad Nauk Rep Bulg Acad Sci. 1962;15:423–6.
13. Lahr DJG, Parfrey LW, Mitchell EAD, Katz LA, Lara E. The chastity of amoebae: re-evaluating evidence for sex in amoeboid organisms. *Proc R Soc B Biol Sci*. 2011;278:2081–90.
14. Normark BB, Judson OP, Moran NA. Genomic signatures of ancient asexual lineages. *Biol J Linn Soc*. 2003;79:69–84.
15. Patil S, Moeyers S, von Dassow P, Huysman MJJ, Mapleson D, De Veylder L, et al. Identification of the meiotic toolkit in diatoms and exploration of meiosis-specific SPO11 and RAD51 homologs in the sexual species *Pseudonitzschia multistriata* and *Seminavis robusta*. *BMC Genomics*. 2015;16:930.
16. Schurko AM, Logsdon JM. Using a meiosis detection toolkit to investigate ancient asexual "scandals" and the evolution of sex. *BioEssays*. 2008;30:579–89.
17. Hofstatter PG, Brown MW, Lahr DJG. Comparative genomics supports sex and meiosis in diverse Amoebozoa. *Genome Biol Evol*. 2018;10:3118–28.
18. Malik SB, Pightling AW, Stefaniak LM, Schurko AM, Logsdon JM Jr. An expanded inventory of conserved meiotic genes provides evidence for sex in *Trichomonas vaginalis*. *PLOS ONE*. 2008;3:e2879.
19. Dunthorn M, Katz LA. Secretive ciliates and putative asexuality in microbial eukaryotes. *Trends Microbiol*. 2010;18:183–8.
20. Dacks J, Roger AJ. The first sexual lineage and the relevance of facultative sex. *J Mol Evol*. 1999;48:779–83.
21. Dunthorn M, Zufall RA, Chi J, Paszkiewicz K, Moore K, Mahé F. Meiotic genes in colpodan ciliates support secretive sexuality. *Genome Biol Evol*. 2017;9:1781–7.
22. Maciver SK. Asexual eukaryotes reproduced asexually, facilitated by polyploidy: a hypothesis. *BioEssays*. 2019;41:1900152.
23. Maciver SK. Asexual amoebae escape Muller's ratchet through polyploidy. *Trends Parasitol*. 2016;32:855–62.
24. Breuer R. Fortpflanzung und biologische Erscheinungen einer Chlamydomorphform auf Agarkulturen. *Arch Protistenk*. 1917;37:65–92.
25. Decaestecker E, De Meester L, Mergey J. Cyclical parthenogenesis in daphnia: sexual versus asexual reproduction. In: Schön I, Martens K, Dijk P, editors. *Lost sex: the evolutionary biology of parthenogenesis*. Dordrecht: Springer Netherlands; 2009. p. 295–316.
26. Lucchesi P, Santangelo G. How often does conjugation in ciliates occur? Clues from a seven-year study on marine sandy shores. *Aquat Microb Ecol*. 2004;36:195–200.
27. Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO, et al. The transcriptional program of sporulation in budding yeast. *Science*. 1998;282:699–705.
28. Bu D, Luo H, Huo P, Wang Z, Zhang S, He Z, et al. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res*. 2021;49:W317–25.
29. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*. 2011;39 suppl_2:W316–22.
30. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021. <https://doi.org/10.1093/molbev/msab293>.
31. Commisso C, Davidson SM, Soydaner-Azeloglu RG, Parker SJ, Kamphorst JJ, Hackett S, et al. Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. *Nature*. 2013;497:633–7.
32. Kamphorst JJ, Nofal M, Commisso C, Hackett SR, Lu W, Grabocka E, et al. Human pancreatic cancer tumors are nutrient poor and tumor cells actively scavenge extracellular protein. *Cancer Res*. 2015;75:544–53.
33. Xiao F, Li J, Huang K, Li X, Xiong Y, Wu M, et al. Macropinocytosis: mechanism and targeted therapy in cancers. *Am J Cancer Res*. 2021;11:14–30.
34. Cao L, Alani E, Kleckner N. A pathway for generation and processing of double-strand breaks during meiotic recombination in *S. cerevisiae*. *Cell*. 1990;61:1089–101.
35. Primig M, Williams RM, Winzler EA, Tevzadze GG, Conway AR, Hwang SY, et al. The core meiotic transcriptome in budding yeasts. *Nat Genet*. 2000;26:415–23.
36. Mozzachiodi S, Tattini L, Llored A, Irizar A, Školjanc N, D'Angiolo M, et al. Aborting meiosis allows recombination in sterile diploid yeast hybrids. *Nat Commun*. 2021;12:6564.
37. Mozzachiodi S, Liti G. Evolution of yeast hybrids by aborted meiosis. *Curr Opin Genet Dev*. 2022;77:101980.
38. Petukhova GV, Pezza RJ, Vanevski F, Ploquin M, Masson J-Y, Camerini-Otero RD. The Hop2 and Mnd1 proteins act in concert with Rad51 and Dmc1 in meiotic recombination. *Nat Struct Mol Biol*. 2005;12:449–53.
39. Pezza RJ, Petukhova GV, Ghirlando R, Camerini-Otero RD. Molecular activities of meiosis-specific proteins Hop2, Mnd1, and the Hop2-Mnd1 complex *. *J Biol Chem*. 2006;281:18426–34.
40. Bugreev DV, Pezza RJ, Mazina OM, Voloshin ON, Camerini-Otero RD, Mazin AV. The resistance of DMC1 D-loops to dissociation may account for the DMC1 requirement in meiosis. *Nat Struct Mol Biol*. 2011;18:56–60.
41. Nakagawa T, Ogawa H. The *Saccharomyces cerevisiae* MER3 gene, encoding a novel helicase-like protein, is required for crossover control in meiosis. *EMBO J*. 1999;18:5714–23.
42. Nakagawa T, Kolodner RD. *Saccharomyces cerevisiae* Mer3 is a DNA helicase involved in meiotic crossing over. *Mol Cell Biol*. 2002;22:3281–91.
43. Eichinger CS, Jentsch S. Synaptonemal complex formation and meiotic checkpoint signaling are linked to the lateral element protein Red1. *Proc Natl Acad Sci*. 2010;107:11370–5.
44. Masson J-Y, West SC. The Rad51 and Dmc1 recombinases: a non-identical twin relationship. *Trends Biochem Sci*. 2001;26:131–6.
45. Pezza RJ, Voloshin ON, Vanevski F, Camerini-Otero RD. Hop2/Mnd1 acts on two critical steps in Dmc1-promoted homologous pairing. *Genes Dev*. 2007;21:1758–66.
46. Bonner JT, Savage LJ. Evidence for the formation of cell aggregates by chemotaxis in the development of the slime mold *Dictyostelium discoideum*. *J Exp Zool*. 1947;106:1–26.
47. Katoh M, Chen G, Roberge E, Shaulsky G, Kuspa A. Developmental commitment in *Dictyostelium discoideum*. *Eukaryot Cell*. 2007;6:2038–45.
48. Kelly B, Carrizo GE, Edwards-Hicks J, Sanin DE, Stanczak MA, Priesnitz C, et al. Sulfur sequestration promotes multicellularity during nutrient limitation. *Nature*. 2021;591:471–6.
49. Hartl J, Kiefer P, Kaczmarczyk A, Mittelviehhaus M, Meyer F, Vonderach T, et al. Untargeted metabolomics links glutathione to bacterial cell cycle progression. *Nat Metab*. 2020;2:153–66.
50. Kraus D, Chi J, Boenigk J, Beisser D, Graupner N, Dunthorn M. Putatively asexual chrysophytes have meiotic genes: evidence from transcriptomic data. *PeerJ*. 2019;6:e5894.
51. Jaron KS, Bast J, Nowell RW, Ranallo-Benavidez TR, Robinson-Rechavi M, Schwander T. Genomic features of parthenogenetic animals. *J Hered*. 2021;112:19.
52. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
53. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*. 2012;28:3150–2.
54. Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS ONE*. 2016;11:e0163962.
55. Li H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics*. 2021;37:4572–4.
56. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
57. Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit – interactive quality assessment of genome assemblies. *G3 GenesGenomesGenetics*. 2020;10:1361–74.
58. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38:4647–54.
59. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
60. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
61. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2:100141.
62. Meiosis-related expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*. Raw sequence reads. NCBI Bioproject [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1108686/%20\(2024\)](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1108686/%20(2024)).

Publisher's Note

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

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

The following chapter consists of a manuscript: “Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*”

Manuscript published in *Science Advances*.

The CRediT (Contributor Roles Taxonomy) is employed to clarify the contribution of the authors.

Authors:

Hüsna Öztoprak, **Shan Gao**, Nadège Guiglielmoni, Alexander Brandt, Yichen Zheng, Mohammed Errbi, Viktoria Bednarski, Christian Becker, Kerstin Becker, Lea Borgschulte, Katharina Atsuko Burak, Anne-Marie Dion-Côté, Vladislav Leonov, Linda Opherden, Satoshi Shimano, Jens Bast

Shared Corresponding author

Author contributions of this thesis:

Formal analysis: Assembled and annotated the reference genome and annotated haplotype assemblies, differentially expressed alleles (DEAs) detection, horizontal gene transfer (HGTs) detection, repeat region annotation, supervised students on identifying *de novo* mutations, *Wolbachia* detection, orthology and *Ka/Ks* analyses between the two haplotypes.

Methodology: Detection of DEAs, identifying *de novo* mutations, identifying orphan HGTs, decontamination of the raw data.

Visualization: Fig. 3. C, D, E in the main text and Fig. S10., Fig. S13., Fig. S14., Fig. S15., Fig. S19., Fig. S20., Fig. S21., Fig. S22. and Fig. S23. in supplementary.

Supervision: Training students in author list (K.A.B. and L.O.)

Data curation: Upload the raw data and assemblies to NCBI and Zenodo, manage the code in GitHub and Zenodo with N.G..

Writing of methods for analyses parts and contributions of writing results parts of the main article and contributions of writing and added analyses as response to reviewer comments.

“Substantial contributions, specifically through the analysis of the German reference genome, focusing on haplotypic differences. Both corresponding authors were additionally contributing as stated in the above-listed roles”, as stated in the published paper “Author Contributions” part.

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*



GENETICS

Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

Hüsna Öztoprak^{1*}, Shan Gao^{1*}, Nadège Guiglielmoni¹, Alexander Brandt², Yichen Zheng^{1,3}, Mohammed Errbi¹, Viktoria Bednarski¹, Christian Becker⁴, Kerstin Becker⁵, Lea Borgschulte¹, Katharina Atsuko Burak¹, Anne-Marie Dion-Côté⁶, Vladislav Leonov⁷, Linda Opherden¹, Satoshi Shimano⁸, Jens Bast¹

copyright © 2025 the Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. no claim to original U.S. Government Works. distributed under a creative commons Attribution noncommercial license 4.0 (cc BY-NC).

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 on the basis of 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 million years ago in European lineages, contrasting Japanese and Canadian lineages. Multiple lines of evidence indicate conservation of one haplotype copy and relaxed selection in the other for the ancient asexual lineages. These findings highlight the evolutionary genomic singularities of ancient asexual oribatid mites that may have contributed to escaping the early demise typically associated with asexuality.

INTRODUCTION

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 to 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 and vice versa (9). To date, however, the genomic signatures of successful persistence without sex, i.e., how peculiar genome dynamics of long-term asexuals affect genome evolution and differ from other, younger asexuals, remain largely unknown (1, 3, 10). One of the most iconic old asexual oribatid mite species is the diploid *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, 12). *P. peltifer* has a generation time of 1 year and produces only female offspring in lab rearings, and extremely rare spanandric males are infertile (13, 14).

In this study, we first generated a phased reference genome of *P. peltifer* to identify the genomic properties that might be associated with ancient asexuality. Second, we tested for independent haplotype evolution (aka “the Meselson effect”), which is theoretically expected under the given type of cytological mechanism and long-term asexuality, using worldwide *P. peltifer* populations. Last, we explored haplotype-specific genomic properties and evolutionary dynamics under independent haplotype evolution to identify a potential impact on the evolution under a long-term absence of sex.

RESULTS AND DISCUSSION

Phased genome assembly to identify signatures of ancient asexuality

To reveal the genomic properties and signatures of haplotype evolution associated with ancient asexuality, 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; next, phased haplotypic blocks were anchored to each other and the chromosome-scale assembly (Fig. 1B and 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, figs. S2 and S3, and 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 and table S2). The genome size estimate via flow cytometry closely matches the assembly size (6% difference: 232 Mb versus 219 Mb) and, together with *k*-nucleotide oligomer spectrum ploidy analyses and synteny analyses between nonhomologous chromosomes, suggests no historical whole-genome duplication (fig. S4 and

¹Institute of Zoology, University of Cologne, Cologne, Germany. ²Department of ecology and evolution, University of Lausanne, Lausanne, Switzerland. ³Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut, Neustadt am Ruenberg, Germany. ⁴Cologne Center for Genomics (CCG), Medical Faculty, University of Cologne, Cologne, Germany. ⁵Genomics & transcriptomics Laboratory, Biologisch-Medizinisches Forschungszentrum, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany. ⁶Département de Biologie, Université de Moncton, Moncton, Canada. ⁷A. N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia. ⁸Science Research Center, Hosei University, Tokyo, Japan. *Corresponding author. email: hoeztoprak@uni-koeln.de (H.Ö.); sgao3@uni-koeln.de (S.G.)

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

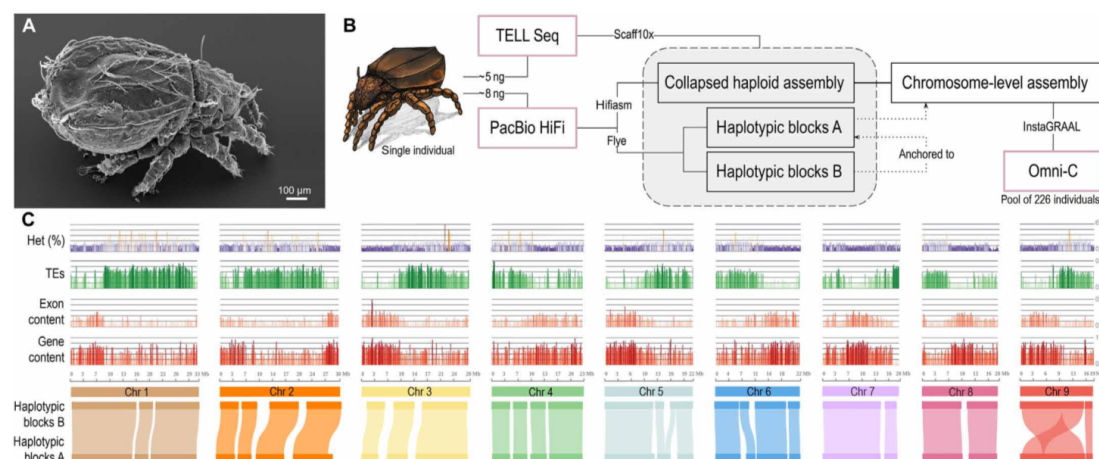


Fig. 1. Assembly strategy and characteristics of the *P. peltifer* genome. (A) *P. peltifer* electron microscopy picture. Scale bar, 100 μm. (B) Schematic workflow of single-individual sequencing and assembly. (C) Proportion of both aligned and collinear regions for haplotypic blocks A (98% of the total length of haplotype A) and B (93% of the total length of haplotype B). Genomic properties of the collapsed assembly with nine pseudochromosomes are depicted in densities of 100-kb blocks: gene content, exon content, tes, and heterozygosity [het (%); percentage of sites differing between haplotypes].

Supplementary Text). In addition, we assembled the complete mitochondrial genome from the same individual (fig. S5).

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”) (15). 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 and by haplotypes (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). Regions of reduced heterozygosity at chromosome extremities are likely the consequence of recombination between haplotypes that is restricted toward telomeres in *P. peltifer*.

Larger regions of high and low heterozygosity are more likely driven by differences in purifying selection among the various chromosomal regions, given that gene content negatively correlates with heterozygosity ($r = -0.55$) and heterozygosity fluctuations are 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. The mean heterozygosity of individuals from Canadian (1.4%) and Japanese (1.0%) populations is consistently lower compared to that from the other populations but shows notably similar heterozygosity dynamics along chromosomes (Fig. 2C). As expected under asexuality, individuals within each population share heterozygous sites; the largest proportions are shared by European individuals and lower proportions by Japanese and Canadian individuals. Notably, a large excess of heterozygous genotypes is shared among all individuals from European populations, but shared variants remain private to individuals of Canadian or Japanese populations (Fig. 2E and figs. S7 and S8). These patterns are consistent with independent haplotype divergence after a transition to asexuality and subsequent population separation for European *P. peltifer* lineages. Together, the very similar heterozygosity dynamics along chromosomes are thus not the consequence of large distinct stretches of loss of heterozygosity that can occur under some nonclonal forms of asexuality (16). 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 and the rate of gene conversion removing heterozygosity, and/or suggest independent 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 whether the transition to asexuality occurred considerable amounts

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

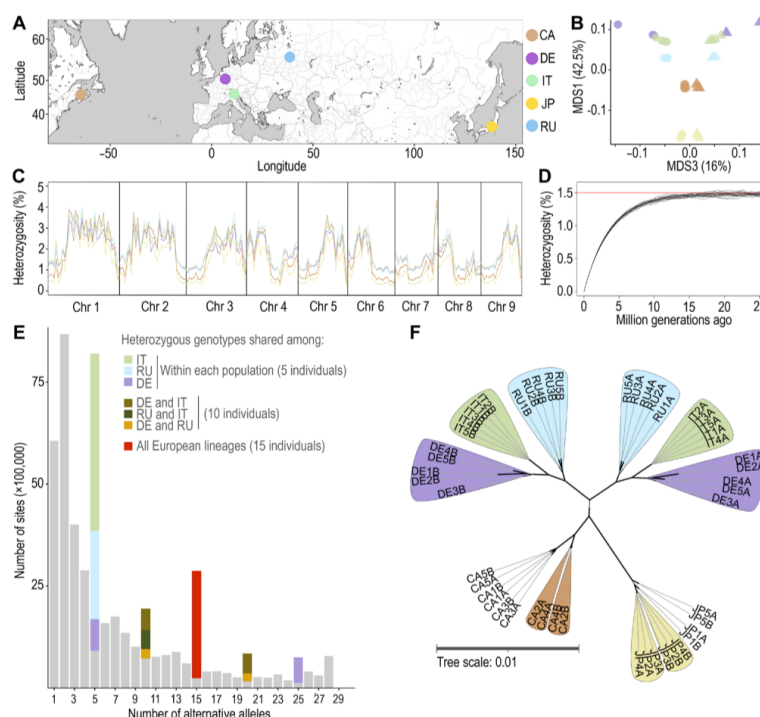


Fig. 2. Haplotype dynamics of worldwide *P. peltifer* populations suggest independent haplotype evolution for at least 20 million years under asexuality of the DE-IT-RU lineages. (A) Sampling locations worldwide, with water bodies shaded in gray. Abbreviations: CA, Canada; DE, Germany; IT, Italy; JP, Japan; RU, Russia (color code kept throughout the figure). (B) Multidimensional scaling (MDS) of genomic population from five individuals per population, accounting for 58.5% of the total variability. Triangles and circles represent haplotypes A and B, respectively. (C) Mean heterozygosity between haplotypes within individuals for each population across nine chromosomes in 1-Mb genomic blocks. (D) Simulations of increased divergence between haplotypes over time with a gene conversion track length of 1000 bp, plateauing after 20 million generations. (E) The site frequency spectrum for European populations is consistent with the accumulation of shared heterozygous variants after loss of sex, followed by lineage separation and subsequent independent accumulation of heterozygosity (see Fig. S7 for details). (F) Unrooted consensus haplotype tree showing the Meselson effect (complete separation of haplotypes A and B with mirror topologies), among European populations from Germany, Russia, and Italy. Bold branches indicate consensus supports >96%.

of time before the separation into different populations (Fig. 2F, fig. S9, and 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. 2F). 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. 2F). Together, heterozygosity patterns are corroborated by haplotype topologies and suggest an ancient transition to asexuality for the ancestor lineage of the European populations and an independent loss of sex for both the Japanese and Canadian populations (Fig. 2, C, E, and F; and figs. S7 to S9). 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 cannot generate the shared haplotype differences among the European populations. Moreover, hybridization would entail substantial genomic changes and thus cannot explain the similar heterozygosity patterns for all populations (17). As remnants of the *Wolbachia* endosymbiont can be detected in the *P. peltifer* genome, the transition to asexuality might have happened via reproductive manipulation of this endosymbiont (fig. S10). This type of transition often results in fully homozygous lineages (18). Together, and given that haplotypic divergence of the Canadian and Japanese populations are considerably lower, transitions to asexuality in *P. peltifer* might have occurred via such a mechanism that substantially removes heterozygosity.

Next, we estimated the age of asexuality (Fig. 2D). 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

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

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 $\mu = 2.05 \times 10^{-9}$ (Supplementary Text). Second, we estimated gene conversion to occur with a minimum track length of 500 base pairs (bp) in the German *P. peltifer* (fig. S11). 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. 2D and fig. S12). Using these biological parameters of *P. peltifer* and assuming that the initial heterozygosity is close to zero, about 20×10^6 generations are necessary to attain the mean 1.5% heterozygosity equilibrium for *P. peltifer* (Fig. 2D). 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.

Haplotype-specific genome evolution in the long-term absence of sex

Next, we explored haplotype-specific genome dynamics associated with the exceptional independent evolution of haplotypes. These are exemplified in the ancient asexual German *P. peltifer* lineage (Fig. 3) encompassing allelic diversification, evolutionary rate differences, gene expression, horizontally transferred genes, and transposable element (TE) activity.

We first analyzed how haplotypes and haplotypic alleles differ in their genetic composition, variation, and rates of evolution. Haplotypes show

a mean heterozygosity of 2.4% (Fig. 1C). Among 10,816 allelic, single-copy 1:1 orthologs, the vast majority differ from each other by nonsynonymous (88.0%) and/or synonymous (96.8%) variants. Approximately 10% of allelic orthologs gained a stop codon, and 1.4% lost a stop codon. About 11% of alleles showed frameshift and nearly 15% nonframeshift insertions and/or deletions (figs. S13 and S14). Nonsynonymous-to-synonymous divergence (K_a/K_s) analysis suggests an allele-specific direction of selection with 4.6% (470 of 10,238) of alleles showing signs of strong selection ($K_a/K_s > 1$) (fig. S15). Similarly, a number of haplotypic blocks exhibit significantly higher levels of genetic diversity (π) compared to their allelic counterparts (>44% of haplotypic blocks; Fig. 3A). Moreover, 9% (650 of 7404) of alleles with population- and haplotype-specific variation show strongly differential rates of evolution (ω) (Fig. 3B and table S4). In contrast to the German lineage, the Japanese lineage shows overall lower K_a/K_s values for allelic orthologs with fewer under strong selection (3.6%; fig. S15). Moreover, there is no evidence for different levels of genetic diversity between haplotypes (figs. S16 to S18) and almost no differential rates of evolution between alleles (fig. S15 and table S4) for the Japanese population. Together, these analyses indicate different evolutionary dynamics acting on the haplotypes for European lineages, consistent with conservation of one haplotypic copy (or allele) and relaxed selection in the other under independent haplotype evolution. Such differences are not present in the Japanese (and Canadian) lineages, where haplotypes are not evolving independently.

In (nearly) clonal species, under haplotypic independence, changes in expression patterns between the two alleles are theoretically expected as the consequence of progressive haploidization of expression via an "enhancer divergence process" (19). In the predictions, the overexpressed gene locus is purged from deleterious alleles, whereas the lowly expressed locus is under relaxed selection.

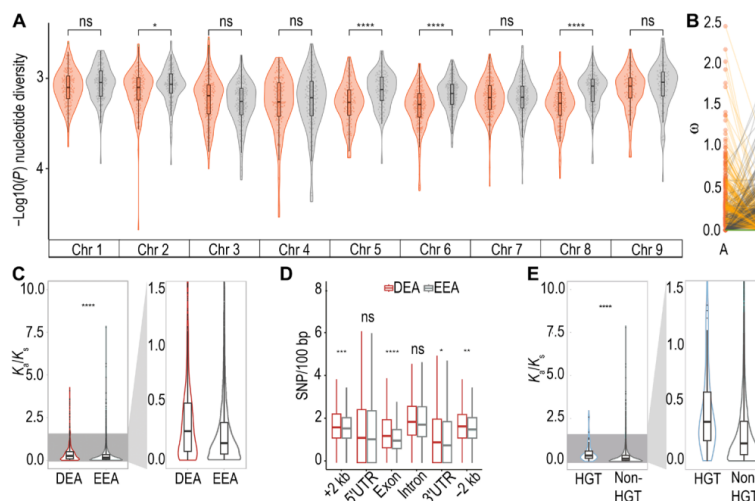


Fig. 3. Haplotype-specific genome evolution. (A) nucleotide diversity (π) for the largest haplotypic blocks in German *P. peltifer*. All estimates were calculated in 25,000 site windows sliding by 10,000 sites. (B) Allele-specific evolutionary rates (ω ; see table S4 for count data). (C) violin plots show levels of K_a/K_s of deAs and eeAs, highlighting $K_a/K_s < 1.5$ (effect size, 0.0820; small). (D) Boxplots calculated for SNPs/100 bp for genic and neighboring regions, including the 5'UTRs and 3'UTRs, exons, introns, and up/downstream regions (± 2 kb, respectively) for deA and eeA (see table S5 for effect sizes). (E) violin plots show levels of K_a/K_s of hGTs and non-hGTs, highlighting $K_a/K_s < 1.5$ (effect size, 0.0468; small). Statistics for (A) to (E) were calculated using the Wilcoxon rank-sum test between deAs ($n = 879$) and eeAs ($n = 9359$) and hGTs ($n = 62$) and non-hGTs ($n = 10,176$) (ns, not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

Hence, we identified and analyzed differentially expressed alleles (DEAs) under haplotypic independence. About 8.7% (936 of 10,816) of alleles are differentially expressed and show elevated K_a/K_s values compared to equivalently expressed alleles (EEAs), albeit with small effect size (Fig. 3C). Specifically, upstream and exonic regions of DEAs showed elevated variant densities compared to EEAs, unlike introns and 5' untranslated regions (5'UTRs), suggesting selection on and diversification of transcription factor, promoter regions, and gene copy, as predicted (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. S19).

One mechanism providing novel traits to organisms is horizontal transfer of preexisting genes (20). Horizontal gene transfers (HGTs) contributed substantially to the evolution of novel traits in the ancient asexual rotifer *Adineta vaga* (21). In *P. peltifer*, such HGTs are distributed throughout the genome without apparent hotspots and represent 2.0% (504) of genes, which are within the typical range of asexual animals (18). These HGTs stem mainly from bacteria but also fungi, plants, and viruses. 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 single-copy HGTs that show signs of diversification, 25.2% are differentially expressed, again suggesting diverging HGT alleles under differential selection (Fig. 3E). These HGTs might have arrived before the transition to asexuality, but divergent haplotypes likely are subject to different selection trajectories, 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. S20). 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, 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). Furthermore, HGT genes of the UGT [uridine diphosphate (UDP)-glycosyltransferases] family may contribute to pesticide resistance, suggesting a contribution to the mite's ecology as soil-living decomposers (figs. S14 and S21 to S23 and data S1).

Another important genomic component contributing to genome evolution is TEs. They proliferate throughout genomes independently of the host cell cycle, and their activity is often deleterious but occasionally can be beneficial (22, 23). Obligate asexual reproduction couples the fate of TEs and their host; thus, TEs likely evolve to be more benign in asexual genomes compared to sexuals (4, 22). We identified TEs and their haplotype-specific activity in the *P. peltifer* reference genome. TEs comprise 27% of the genome assembly, and TE density distributions suggest effective selection against TE insertion within genes, in concordance with previous results (fig. S24 and table S6) (4). We detected fluctuating historical activity as emphasized by Kimura substitution levels (fig. S25). Moreover, we detect noticeable differences in haplotype-specific historical TE activity between some of the largest phased haplotypic blocks of each chromosome (fig. S26). Very recent activity (0% Kimura substitution) is rather low or largely restricted to one haplotype. Similar to haplotypic alleles, these findings suggest divergence of one haplotype and more conservation of the other in TE activity and content.

Overall, all analyses on haplotype-specific genome evolution indicate conservation of one allelic copy (or haplotype) and relaxed

selection in the other in the ancient asexual lineages. This suggests that the decoupling of haplotypes after the transition to asexuality in *P. peltifer* mimics duplications of the whole genome with genes representing quasihomologs. Evolutionary trajectories of duplicated genes often result in a conserved copy and a copy that is "free to evolve." While most gene copies under such relaxed selection will deteriorate via mutation accumulation, a small fraction can potentially provide novel functions (24, 25). Similarly, independent haplotype dynamics might have helped ancient asexual oribatid mites to escape the dead-end fate typically associated with asexuality by maintaining effective purifying selection in one haplotype (3) and harboring the potential for an increased substrate for evolvability in the other in the long-term absence of sex. The impact of independent haplotype evolution on the adaptive potential of ancient asexual oribatid mites remains to be investigated in future, functional studies.

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 leads to conservation of one haplotype and relaxed selection in the other. 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 contribution of haplotypic independence in masking deleterious mutations and serving as a possible source of novelty that allows persistence in the absence of sex over time.

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 were 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 (26). *P. peltifer* was identified morphologically using (27) and molecularly confirmed by cytochrome oxidase I sequencing.

Sample preparation

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

High-molecular-weight gDNA extraction (for ultralow input)

For high-molecular-weight (HMW) single-individual DNA extraction, we established a modified salting-out protocol (28, 29). In short, a single individual was submerged in TNES buffer and flash frozen in liquid nitrogen. The sample was then homogenized using

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

a sterile pestle. After adding proteinase K, the sample was incubated for at least 1 hour. 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 a Qubit Fluorometer version 4 with the Qubit dsDNA HS Assay kit (Thermo Fisher Scientific, Waltham, MA).

Cytochrome oxidase I sequencing

The cytochrome oxidase subunit 1 (~700 bp) region was amplified using the universal primers LC01490F/HC02198R (30). The polymerase chain reaction (PCR) was run with 1 ng of DNA, 2× Thermo Fisher Scientific DreamTaq Green PCR Master Mix, 1 μM forward and 1 μM reverse primers, and double-distilled H₂O added to bring the total reaction volume to 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 exonuclease (1 U/ml) and FastAP (0.3 U/ml) to 8 μl of PCR product and 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

The HMW genomic DNA (gDNA) sample was assessed at the Genomics & Transcriptomics Laboratory (Düsseldorf, Germany) using the Agilent Femto Pulse system. HMW gDNA with an OD_{260/280} 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 WGS library) both with ultralow DNA input of at least 5 ng. Library preparation and sequencing of the SMRTbell templates were conducted by the Genomics & Transcriptomics Laboratory. 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. A total of 226 whole adult specimens were flash frozen with liquid N₂, crushed, and vortexed, and the Omni-C Proximity Ligation Assay protocol for insects and marine invertebrates was followed. The library was sequenced on Illumina NovaSeq 6000 by Novogene (Cambridge, United Kingdom), which generated 92.7 × 10⁶ 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 data S1.

RNA sequencing

Total RNA was extracted from 10 adult individuals from the German population using TRIzol reagent treated with deoxyribonuclease I

within a Direct-zol RNA MicroPrep kit (Zymo Research, Irvine, CA). An RNA sequencing (RNA-seq) library was constructed by the CCG using TruSeq Stranded Total RNA with Ribo-Zero Globin, and 43 × 10⁶ pairs of 100-bp reads were sequenced.

k-Nucleotide oligomer analyses

27-Nucleotide oligomers in the HiFi reads were analyzed using KAT version 2.4.2 (31) with the modules kat hist and kat gcp (default parameters). Ploidy was further investigated using kmc version 3.2.1 with parameters -k 27 -ci 1 -cs 10000 and Smudgeplot (version 0.2.5) (32) with default parameters.

Chromosome-level collapsed and phased assemblies

Genome assembly followed best practices using state-of-the-art approaches for chromosome-level assembly and haplotype resolution (33, 34).

De novo genome assembly

HiFi reads were assembled using Hifiasm (version 0.16.1-r375) (35) with default parameters into initial collapsed haploid contigs and using Flye (version 2.9) (36) 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 (version 1.0.2). The barcodes were formatted to 10x Genomics standard by using ust10x (version 1.0.2) from TELL-seq conversion_tool. The collapsed haploid contigs (Hap0: 211 Mb, N = 33) were scaffolded using Scaff10X (version 4.2) (<https://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 (version 0.6.5) (<https://github.com/FelixKrueger/TrimGalore>) with parameters “-j 30 -q 30 --fastqc --paired” and mapped to the draft haploid assembly using hicstuff (version 3.1.1) and bwa (version 0.7.15) (37) with parameters --enzyme 100 --iterative --aligner bwa. InstaGRAAL (version 0.1.6 no-opengl branch) (38) was run with parameters --level 5 --cycles 100. The scaffolds were curated with instaGRAAL-polish to reduce misassemblies and add 10 Ns in gaps. A total of 99.75% of the assembly was 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 (version 1.11) with default parameters. Gaps in the collapsed scaffolds Hap0 were filled using TGS-GapCloser (version 1.1.1) (39) with the parameters “-tgstype pb --minmap_arg -x asm20 --ne,” and the final haploid (Hap0) assembly was polished using HyPo (version 1.0.3) (40) with the PacBio HiFi reads.

Haplotype scaffolding

Primary and alternative haplotypes were separated using minimap2 and purge_dups (version 1.2.5) (41) and are subsequently designated as haplotypic blocks A and B. Haplotypes were reciprocally scaffolded using RagTag (version 2.1.0) (42) with the other haplotype as the reference.

Anchoring of haplotypes

Haplotypic blocks A were mapped against the collapsed assembly (Hap0) using minimap2 (version 2.24-r1122) with parameter -x

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

Science Advances | Research Article

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 version 5.0.0) (43) against the Arthropoda odb10 lineage (1066 orthologs) and the Arachnida odb10 lineage (2934 orthologs). *k*-Nucleotide oligomer completeness was evaluated using KAT (version 2.4.2) 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` (version 2.24-r1122) with parameters `-ax map-hifi`, and the mapped reads were sorted with `SAMtools` (version 1.11) (44, 45). Hap0 was aligned against the nucleotide database using the Basic Local Alignment Search Tool (BLAST version 2.6.0) (46) 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` (44).

Genome annotation

Repeat and TE annotation and masking

A repeat library including TEs was built using EDTA (version 1.9.6) (47) with parameters `--sensitive 1 --anno 1`. TEs were annotated and classified to the superfamily level using the `FaSTe` pipeline (48). The hardmasked assemblies were converted to softmasked assemblies using `BEDTools` (version 2.26.0) with `mask mode`.

RNA-seq mapping

Adapter sequences were removed from RNA-seq raw reads using `TrimGalore` (version 0.6.5) with parameters `--j 30 -q 30 --fastqc --paired`. Trimmed reads were mapped to the softmasked reference haploid assembly Hap0 and the two haplotypic block assemblies haplotypic blocks A and haplotypic blocks B using `STAR` (version 2.5.1a) (49) with parameters `--readFilesCommand zcat --outSAMtype BAM SortedByCoordinate --outSAMstrandField intronMotif --outFilterIntronMotifs RemoveNoncanonical`.

Gene structure prediction

`Trinity` (version 2.1.1) was used to assemble the RNA-Seq reads de novo, and the resulting 61,067 transcripts were aligned to the assemblies using `PASA` (version 2.5.2). Mapped reads were provided to `StringTie` (version 2.2.0) to predict transcripts. Then, `TransDecoder` (<https://github.com/TransDecoder/TransDecoder/wiki>; last accessed 14 November 2018) was used to find the open reading frames for each gene. `BRACKER2` (version 2.1.6) was used to predict gene structures with RNA-seq evidence. These results were provided to `EVidenceModeler` (version 1.1.1) with weight parameters `"ABINITIO_PREDICTION AUGUSTUS 4; TRANSCRIPT_assembler-database.sqlite 7; OTHER_PREDICTION_transdecoder 8"` to select genes. Last, `PASA` was used to refine UTR regions. The final annotation was converted into protein sequences using `gffread` (version 0.12.1).

Functional annotation

We used `EggNog-mapper` (version 2.1.6) with protein sequences and generated KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway annotations. Furthermore, we used `InterProScan` (version 5.61-93.0) with protein sequences to generate reactome pathway annotations. The Gene Ontology annotation was created by combining the results of `EggNog-mapper` (version 2.1.6) 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 (50) 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 (51) 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` version 2.6.0 with parameters `-evalue 1e-10 -outfmt 6`. To find blocks of gene synteny, `MCSanX` (commit 97e74f4) was run between Hap0, haplotypic blocks A, and haplotypic blocks B with default parameters. Collinearity was visualized with `SynVisio` (52) with a minimum match score ≥ 3950.895 . Heterozygosity between the haplotypes was estimated using `Mash` (version 2.3) (53).

Genome architecture

To calculate TE, gene, and exon content along the genome, `BEDTools` and `BEDOPS` (version 2.4.41) (54) were used in conjunction with the TE and gene annotations in 100-kb nonoverlapping windows, following the pipeline described in <https://github.com/merrbi/Peal-SocialPolymorphism/blob/main/PealGenomics.md#te-and-gene-content>.

Mitochondrial genome assembly and annotation

TELL-seq reads were trimmed using `TrimGalore` and assembled using `MitoFinder` version 1.4.1 with the mitochondrial genome of *Steganacarus magnus* as seed (55, 56). The assembly was annotated using `MitFi` version 0.1 (57) and then using `MITOS` (58), `MITOS2` (59) (both available at <https://gitlab.com/Bernt/MITOS/>), `ARWEN` version 1.2.3 (60), and `tRNAscan-SE` version 2.0 (61) 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 on the basis of their predicted secondary structure and their minimum free energy, computed using the `RNAfold` web server (<http://rna.tbi.univie.ac.at/>) (62).

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 (63, 64)]. *P. peltifer* individuals were sampled from the same population as the reference genome individual (Dahlem, Germany) in early June 2021. For three individual

Downloaded from <https://www.science.org at Universitaets und Stadtbibliothek Koeln on February 23, 2025>

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

mothers (M1, M2, and M3), their “offspring daughters,” i.e. eggs, were extracted after removing the genital plates. Eggs were cleansed in 0.05% NaClO (DonKlorix; CP GABA GmbH, Hamburg, Germany) with a brush to remove tissue from their mothers. One daughter (D1) for M1, one daughter (D2) for M2, and three daughters (D3, D4, and D5) for M3 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, US) in combination with UMI Adaptor “NEBNext Multiplex Oligos for Illumina (Unique Dual Index UMI Adaptors DNA Set 1)” (New England Biolabs, Ipswich, US) 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 (minimum, 100.3×10^6 ; maximum, 103.0×10^6) and a mean of 100.9×10^6 (minimum, 81.2×10^6 ; maximum, 111.0×10^6) for the eggs.

Identifying de novo mutations

Reads were demultiplexed using UMI-tools extract (version 1.1.2). Adapters were trimmed using TrimGalore (version 0.6.5) 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* (version 0.7.17-r1198-dirty) and filtered with the Genome Analysis ToolKit (GATK version 4.1.9). Duplications were removed with UMI-tools (version 1.1.2). Potential variant sites were called and filtered using GATK (version 4.1.9) following an in-house pipeline (all scripts and parameters are outlined in detail at GitHub). In short, variants were filtered by read depth (DP), phred-scaled likelihood of the genotype (PL), allele depth, and homozygous to heterozygous mutations. In addition, following common practice (64, 65), 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 version 2.8.13) 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 table 3 in data S1).

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 three mothers, similarly filtered as the de novo mutations. 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 DP, genotype, PL, and allele depth: DP was filtered like above; sites with DP between 28 and 124 were accepted; for genotype, all homozygous sites from the mother were selected; PL for sites must have fulfilled $PL_{\text{second most likely}} - PL_{\text{first most likely}} < 120$; and only sites where allele depth 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 (version 1.0.3) using *bwa mem* (version 0.7.15) with default parameters to the collapsed softmasked haploid reference genome (Hap0). The resulting alignment was sorted using SAMtools (version 1.11), and duplications were removed using Picard MarkDuplicates (version 2.26.2 Broad Institute). Coverage was calculated with SAMtools flagstat. Following, variants were called using the Genome Analysis ToolKit (GATK version 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.

Site frequency spectrum analysis

To visualize the genotype profiles of *P. peltifer* populations on a site frequency spectrum, we filtered the population-specific variants to only include biallelic SNPs with a mean coverage between 15.7 and 124 and at least five reads per genotype. In addition, we retained SNPs with a minor allele count of at least 1 (i.e., polymorphic sites) with no missing data, resulting in 7,710,168 SNPs. The filtering thresholds were selected on the basis of the distribution of these metrics in the original call set. An in-house script was then used to count the number of genotypes across all individuals at each SNP position, and R (version 4.3.2; R Core Team, 2023) was used to produce a bar plot depicting the number of sites with a given number of nonreference variants.

Population data coverage filter

To compare population statistic metrics among the different populations (Germany, Russia, Italy, Japan, and Canada), 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) is used for the population statistics. The high variation of coverage distribution among individuals (median ranging from 28 in IT3 to 124 in IT1, although most are between 70 and 100) called for individual-based coverage filtering.

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 10 possible pairs are used and the mean value is used. This ratio can show whether the heterozygosity and site frequency distribution is more similar to that of 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).

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

Science Advances | Research Article

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 1000 bp 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 version 38.63; (66)]. Reads with \sim 5 noncontiguous substitutions were discarded (minratio = 0.9). Reads that mapped to both haplotypic block assemblies were kept (ambiguous2 = split) and merged with each set of split reads mapping best to haplotypic blocks A and with the set of split reads mapping best to haplotypic blocks B per individual. This was done to avoid biasing the analysis toward 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 version 0.7.17 and sorted using SAMtools version 1.15.1 (as described above). Optical and sequencing duplicates were removed using picard MarkDuplicates version 2.26.2 (Broad Institute). Variants were called for each split set of resequencing reads, i.e., each haplotype, separately using gatk version 4.2.6.1 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 because of 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 because of 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 in phase, i.e., no phase switching possible within a region, scaffolds of haplotypic block assembly haplotypic blocks A were aligned to the collapsed assembly using D-Genies and standard parameters (see table S3) (67). Per chromosome, the longest alignment block was extracted from the haplome alignment using Geneious (68) and subdivided into 1 Mio Mb long bins using msa_split (69). This resulted in a region of 47,131,932 bp (\sim 21% of the collapsed assembly). Best-fitting ML trees were reconstructed for each bin using IQ-tree (version 1.6.1) (70) 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 Fig. 2 and fig. S9).

Multidimensional scaling

To conduct a multidimensional scaling analysis, the haplotype-specific SNPs were filtered using VCFtools [version 0.1.16; (71)] to

include only biallelic SNPs with less than 20% missing data, resulting in 4,463,885 SNPs. Next, SNPs were pruned based on linkage disequilibrium with a threshold of $r^2 < 0.2$ ($--indep-pairwise$ 1 kb 1 0.2), plink [version 1.90b7.2; (72)]. The resulting 416,588 linkage disequilibrium-pruned SNPs were then used to perform multidimensional scaling using the “--mds-plot 5 eigvals” command in plink.

Nucleotide diversity calculation

Overall nucleotide diversity as well as diversity at synonymous and nonsynonymous sites in the five populations of *P. peltifer* was calculated considering each haplotype separately. For this, haplotype-specific genotypic calls, excluding those with more than 20% missing data, were used, resulting in 101,833,929 genotyped sites (including both variant and invariant positions). Invariant positions were extracted separately, and the effects of all variant sites on protein-coding genes were annotated using SnpEff [version 5.2c; (73)]. From these annotated variants, retrieved SNPs were annotated as “synonymous_variant” or “missense_variant” separately with SnpSift [version 5.2c; (73)]. Each of these types of variants was combined with the set of invariant positions, resulting in two sets of calls: one containing the synonymous variants and the other containing the nonsynonymous variants.

The two VCF files were converted into genotype files. Nucleotide diversity at synonymous and nonsynonymous sites in 25,000 site windows sliding by 10,000 sites was calculated using scripts available at https://github.com/simonhmartin/genomics_general. Unlike VCFtools, this approach does not assume missing sites to be invariants but requires their inclusion with the variant sites to compute more reliable estimates of nucleotide diversity (74). Overall levels as well as nucleotide diversity at synonymous and nonsynonymous sites were compared between haplotypic blocks A and B for the two longest haplotypic blocks on each chromosome. This analysis was restricted to populations from Germany and Russia (showing the Meselson effect) and Canada and Japan (no Meselson effect). Individuals from Italy were excluded as they showed very little genetic variation.

Allele-specific evolutionary rates

Omega (ω or d_N/d_S ratio) was calculated for single-copy orthologous genes between *P. peltifer* using haplotype-specific polymorphisms in two populations: one exhibiting haplotype-specific divergence (Germany) and one without haplotype-specific divergence (Japan) using *Hermannia gibba* as an outgroup species (75). From the call set generated for the “Haplotype-specific analyses” section, 92,105,324 positions (including invariant calls) were successfully genotyped in at least 90% of the samples from both German and Japanese populations.

Next, to produce consensus FASTA files with variant and invariant sites specific to each haplotype in the two populations (four consensus FASTA files in total), the most represented allele (allele frequency > 0.6) was applied in each haplotype and population separately to the reference genome using BCFtools consensus [version 1.19; (45)]. To extract the longest protein isoform and corresponding CDS for each gene, SeqKit [version 2.8.0; (76)] was used. To retrieve all single-copy orthologs, OrthoFinder [version 2.5.5; (77)] was used. Amino acid alignments for each of the 8853 single-copy ortholog groups were generated using ClustalO [version 1.2.4; (78)] and converted into nucleic acid alignments using Pal2Nal [version 14; (79)]. After excluding one gene that failed the previous step, phylogenetic trees were generated for the remaining 8852 genes using RAXML [version 8.2.12; (80)].

Downloaded from <https://www.science.org at Universitaets und Stadtbibliothek Koeln on February 23, 2025>

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

Science Advances | Research Article

To estimate omega (ω) or d_S/d_N ratio, the FitMG94 model was applied, by implementing the HyPhy package [version 2.5.60; (81)], to fit the Muse-Gaut model (82) of DNA sequence evolution, which estimates the number of synonymous and nonsynonymous substitutions per nucleotide site for each branch on a given tree. Gene alignments containing premature stop codons or not being a multiple of three were excluded, resulting in a total of 8600 gene alignments to which the FitMG94 model was applied. Further genes not overlapping any of the applied SNP genotypes were excluded, resulting in 7404 genes, with ω rates estimated for the branch leading to each allele in the two populations.

Horizontal gene transfer

To detect horizontally transferred genes from different organisms, a previously developed pipeline (<https://github.com/reubwn/hgt>) (83) was used, which applies several lines of evidence for detecting HGT candidates (HGTc). In short, DIAMOND (version 0.9.21.122) BLASTP was used in “sensitive” mode to align the protein sequences from the reference and two haplotypic block assemblies to the UniRef90 database. Then, taxid information from the National Center for Biotechnology Information (NCBI) Taxonomy database and three measures were used to identify putative candidate HGTs.

The three measures including the (i) HGT index using the best-hit bitscores from the BLASTP results and defined “(best-hit bitscore for OUTGROUP) – (best-hit bitscore for INGROUP);” (ii) alien index based on *E*-values also from the BLASTP results to evaluate HGT defined by “ $\log_{10}((\text{best-hit } E\text{-value for INGROUP}) + 1 \times 10^{-200}) - \log_{10}((\text{best-hit } E\text{-value for OUTGROUP}) + 1 \times 10^{-200})$ ” (21), and then (iii) consensus hit support using the sum of bitscores across all hits, and the HGTc must be designated as the outgroup and the designation of rest hits $\geq 90\%$. 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. 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. The mapped read count was divided by gene length for normalization. In addition, the orphan HGTc were checked for syntenic neighboring gene content.

Variation calling and annotation

PacBio HiFi reads from the reference individual were mapped to the collapsed assembly (Hap0) using minimap2 (version 2.24-r1122). Subsequently, the resulting BAM file was sorted using SAMtools (version 1.11). The above-mentioned pipeline was applied, and the SNP was annotated by ANNOVAR (version 2020-06-07 23:56:37-0400).

Orthology and K_a/K_s

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 (version 2.5.4) (77) to detect single-copy ortho-group 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 (version 2.0) (84)–integrated KaKs_Calculator (version 1.2) was used to calculate the K_a/K_s value for the haplotypic orthologous single-copy gene pairs.

Differential expression of alleles

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 (version 2.5.1a) with parameters “--readFilesCommand zcat --outSAMtype BAM SortedByCoordinate --outSAMstrandField intronMotif --outFilterIntronMotifs RemoveNoncanonical --outFilterMismatchNmax3 --outFilterMismatchNoverLmax 0.1 --outFilterMismatchNoverReadLmax 0.5.”

Differential expression of alleles between the two haplotypic blocks haplotypic blocks A and haplotypic blocks B was performed with GFOLD version 1.1.4 (85). The GFOLD value $\geq |0.2|$ and the RPKM (Reads Per Kilo bases per Million reads) > 1 of all samples were set as cutoff values (the GFOLD value could be considered as a reliable \log_2 fold change). All identified DEAs were further submitted for functional enrichment with R package clusterProfiler (version 4.6.2) using all RPKM > 1 single-copy genes as the background with the “enricher” function.

Wolbachia

Previously published *Wolbachia* genomes on NCBI used ASM1658442v1 as a query. To identify candidate genes integrated into the mite genome, we used the DIAMOND (version 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 NGenome-Syn (version 1.41) (86).

Simulations

Single lineage–based long-term simulation

Considering a lack of exchange of genetic material between individuals in a cloning species, long-term 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 the value of 0) or heterozygosity (represented by the value of 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 “ $2 \times$ ” 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 $\mu = 2.05 \times 10^{-9}$ (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 haplotypes of the populations). These parameters are set as constant so that the equilibrium state is constant across simulations. We set the mean gene conversion track length, $E(L_{GC})$, as a variable with values of 50, 100, 200, 500, 1000, 2000, and 5000; and the L_{GC} for each gene conversion event is either constant or geometrically distributed. This is to observe the effect of gene conversion track length on how heterozygotic sites are distributed on the chromosome. 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

Downloaded from <https://www.science.org at Universitaets und Stadtbibliothek Koeln> on February 23, 2025

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

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 2×10^7 to 5×10^7 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*, flow cytometry was used following the method in (87). *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 of buffer was 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 10 μl of Galbraith buffer, after which 440 μl of budder was added. Next, 50 μl of the *D. melanogaster* mix was added to the mite cells and homogenized. The mix containing the sample and reference was filtered through a 45- μm nylon mesh, and 20 μl of propidium iodide (50 $\mu\text{g}/\text{ml}$) was added and homogenized. The sample was then incubated at 4°C for 2 hours before 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. 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 were deemed good quality if the peaks were separated and 4C had the double mean FL3-A value as the 2C peak. Several replicates were run, and the best-quality peak was chosen as a result.

Supplementary Materials

The PDF file includes:

Supplementary text

Figs. S1 to S26

tables S1 to S6

Legend for data S1

Other Supplementary Material for this manuscript includes the following:

data S1

REFERENCES AND NOTES

1. A. Brandt, P. tran van, c. Bluhm, Y. Anselmetti, Z. dumas, e. Figuet, c. M. François, n. Galtier, B. heimbürger, K. S. Jaron, M. Iabédan, 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**, e2101485118 (2021).
2. M. heethoff, K. domes, M. Iauemann, 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, 1991), pp. 107–136.
8. M. Maraun, M. heethoff, K. Schneider, S. Scheu, G. Weigmann, J. M. cianciollo, 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, I. 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 parthénogenèse thélytoque de deux espèces d’acariens oribatides: *Triphochthonius 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. c. Palmer, R. A. norton, Further experimental proof of thelytokous parthenogenesis in oribatid mites (Acari: Oribatida: desmonomata). *Exp. Appl. Acarol.* **8**, 149–159 (1990).
 14. t. Pfingstl, h. Schatz, A survey of lifespans in Oribatida excluding Astigmata (Acari). *Zoosymposia* **20**, 7–27 (2021).
 15. c. W. Birky Jr., heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144**, 427–437 (1996).
 16. J. engelstädter, Asexual but not clonal: evolutionary processes in automic populations. *Genetics* **206**, 993–1009 (2017).
 17. B. dunn, t. Paulish, A. Stanbery, J. Piotrowski, G. Koniges, e. Kroll, e. J. Iouis, G. Iiti, 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).
 18. 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).
 19. F. Fyon, t. Ienormand, Cis-regulator runaway and divergence in asexuals. *Evolution* **72**, 426–439 (2018).
 20. I. Boto, horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc. Biol. Sci.* **281**, 20132450 (2014).
 21. e. A. Gladyshev, M. Meselson, i. R. Arkhipova, Massive horizontal gene transfer in bdelloid rotifers. *Science* **320**, 1210–1213 (2008).
 22. J. Bast, K. S. Jaron, d. Schusel, d. Roze, t. Schwander, Asexual reproduction reduces transposable element load in experimental yeast populations. *eLife* **8**, e48548 (2019).
 23. e. B. chung, n. c. elde, c. Feschotte, Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017).
 24. A. Wagner, “Gene duplications and innovation” in *The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems* (Oxford Univ. Press, 2011).
 25. J. I. Payne, A. Wagner, the causes of evolvability and their evolution. *Nat. Rev. Genet.* **20**, 24–38 (2019).
 26. d. Kempson, M. I. Lloyd, R. Ghelardi, A new extractor for woodland litter. *Pedobiologia* **3**, 1–21 (1963).
 27. G. Weigmann, *Hornmilben (Oribatida) (Die Tierwelt Deutschlands)* (Goecke & evers, 2006).
 28. S. A. Miller, d. d. dykes, h. F. Polesky, A simple salting out procedure for extracting dNA from human nucleated cells. *Nucleic Acids Res.* **16**, 1215 (1988).
 29. h. Öztoprak, J. Bast, Modified salting out method for high molecular weight gDNA extraction (oribatid mites) (2023). <https://doi.org/10.17504/protocols.io.xmvm3yybl3p/v1>.
 30. O. Folmer, M. Black, W. hoeh, R. Iutz, R. vrijenhoek, dNA primers for amplification of mitochondrial cytochrome c oxidase subunit i from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299 (1994).
 31. 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).
 32. 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).
 33. n. Guiguelmoni, A. houtain, A. derzelle, K. van doninck, J.-F. Flot, Overcoming uncollapsed haplotypes in long-read assemblies of non-model organisms. *BMC Bioinformatics* **22**, 303 (2021).
 34. n. Guiguelmoni, R. Rivera-vicéns, R. Koszul, J.-F. Flot, A deep dive into genome assemblies of non-vertebrate animals. *Peer Community J.* **2**, 10.24072/pcjournal.128 (2022).
 35. h. cheng, G. t. conception, X. Feng, h. Zhang, h. Ii, haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
 36. M. Kolmogorov, J. Yuan, Y. Ilin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
 37. h. Ii, R. durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
 38. I. Baudry, n. Guiguelmoni, h. Marie-nelly, A. cormier, M. Marbouty, K. Avia, Y. I. Mle, O. Godfroy, I. Sterck, J. M. cock, c. Zimmer, S. M. coelho, R. Koszul, instaGRAAL:

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

- chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffold. *Genome Biol.* **21**, 148 (2020).
39. M. Xu, I. Guo, S. Gu, O. Wang, R. Zhang, B. A. Peters, G. Fan, X. Liu, X. Xu, I. Deng, Y. Zhang, tGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Gigascience* **9**, gaa094 (2020).
40. R. Kundu, J. Casey, W.-K. Sung, hyPo: Super fast & accurate polisher for long read genome assemblies. *bioRxiv* 882506 [Preprint], 20 December 2019. <https://doi.org/10.1101/2019.12.19.882506>.
41. 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).
42. M. Alonge, I. Ieagle, M. Kirsche, K. Jenike, S. Ou, S. Aganezov, X. Wang, Z. B. Ippman, 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).
43. 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).
44. R. Challis, e. Richards, J. Rajan, G. Cochrane, M. Blaxter, BlobToolKit – interactive quality assessment of genome assemblies. *G3 Genes Genomes Genet.* **10**, 1361–1374 (2020).
45. 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. Ii, twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
46. 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).
47. S. Ou, W. Su, Y. Iiao, K. Chougule, J. R. A. Agda, A. J. Helling, c. S. B. Iugo, 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).
48. e. A. Bell, c. I. Butler, c. Oliveira, S. Marburger, I. 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).
49. 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).
50. t. Wicker, F. Sabot, A. Hua-van, J. I. 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).
51. S. J. Teresi, M. B. Teresi, P. R. Edger, te density: A tool to investigate the biology of transposable elements. *Mob. DNA* **13**, 11 (2022).
52. v. Bandi, c. Gutwin, J. n. Siri, e. Neufeld, A. Sharpe, i. Parkin, visualization tools for genome conservation. *Methods Mol. Biol.* **2443**, 285–308 (2022).
53. 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).
54. S. Nepf, M. S. Kuehn, A. P. Reynolds, e. Haugen, R. e. Thurman, A. K. Johnson, e. Ryne, M. t. Maurano, J. vierstra, S. Thomas, R. Sandstrom, R. Humbert, J. A. Stamatoyannopoulos, BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**, 1919–1920 (2012).
55. K. Domes, M. Marau, S. Scheu, S. I. Cameron, the complete mitochondrial genome of the sexual oribatid mite *Steganacarus magnus*: Genome rearrangements and loss of tRNAs. *BMC Genomics* **9**, 532 (2008).
56. R. Allio, A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, F. Delsuc, MitoFinder: efficient automated large-scale extraction of mitochondrial data in target enrichment phylogenomics. *Mol. Ecol. Resour.* **20**, 892–905 (2020).
57. J. Frank, P. Joern, M. Bernt, A. Donath, M. Middendorf, c. Florentz, P. F. Stadler, improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Nucleic Acids Res.* **40**, 2833–2845 (2011).
58. M. Bernt, A. Donath, F. Jühling, F. Externbrink, c. Florentz, G. Fritsch, J. Pütz, M. Middendorf, P. F. Stadler, MitOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **69**, 313–319 (2013).
59. 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).
60. d. Laslett, B. Canbäck, ARWen: A program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**, 172–175 (2008).
61. P. P. Chan, t. M. Lowe, tRNAscan-Se: Searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* **1962**, 1–14 (2019).
62. R. Lorenz, S. h. Bernhart, c. h. Z. Siederdisen, h. tafer, c. Flamm, P. F. Stadler, i. I. Hofacker, ViennaRNA package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).
63. 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).
64. I. A. Bergeron, S. Besenbacher, t. Turner, c. J. Versoza, R. J. Wang, A. I. 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**, e73577 (2022).
65. 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).
66. B. Bushnell, BBMap (2015). <https://sourceforge.net/projects/bbmap/>.
67. F. Cabanettes, c. Klopp, d-GenieS: dot plot large genomes in an interactive, efficient and simple way. *PeerJ* **6**, e4958 (2018).
68. 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).
69. M. J. Hubisz, K. S. Pollard, A. Siepel, PhAST and RPhAST: Phylogenetic analysis with space/time models. *Brief. Bioinform.* **12**, 41–51 (2011).
70. 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).
71. P. Danecek, A. Auton, G. Abecasis, c. A. Albers, e. Banks, M. A. dePristo, R. e. Handsaker, G. Hunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, 1000 Genomes Project Analysis Group, the variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
72. c. c. Chang, c. c. Chow, I. c. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLink: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
73. P. Cingolani, A. Platts, I. I. Wang, M. Coon, t. Nguyen, I. Wang, S. J. Land, X. Lu, d. M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnPeff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
74. K. I. Korunes, K. Samuk, pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol. Ecol. Resour.* **21**, 1359–1368 (2021).
75. S. Wulsch, h. Öztoprak, n. Guiguelmoni, d. I. Jeffries, J. Bast, A female heterogametic ZW sex-determination system in Acariformes. *bioRxiv* 563255 [Preprint], 29 October 2023. <https://doi.org/10.1101/2023.10.24.563255>.
76. W. Shen, S. Ie, Y. Ii, F. Hu, SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS ONE* **11**, e0163962 (2016).
77. d. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
78. F. Sievers, A. Wilm, d. Dineen, t. J. Gibson, K. Karplus, W. Ii, R. Lopez, h. McWilliam, M. Remmert, J. Söding, J. d. Thompson, d. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
79. M. Suyama, d. Torrents, P. Bork, PA12nAl: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609–W612 (2006).
80. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
81. S. I. Kosakovsky Pond, A. F. Y. Poon, R. Velazquez, S. Weaver, n. I. Hepler, B. Murrell, S. d. Shank, B. R. Magalis, d. Bouvier, A. Nekrutenko, S. Wisotsky, S. J. Spielman, S. d. W. Frost, S. v. Muse, hyPhy 2.5-A customizable platform for evolutionary hypothesis testing using phylogenies. *Mol. Biol. Evol.* **37**, 295–299 (2020).
82. S. v. Muse, B. S. Gaut, A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Mol. Biol. Evol.* **11**, 715–724 (1994).
83. 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).
84. Z. Zhang, J. Xiao, J. Wu, h. Zhang, G. Liu, X. Wang, I. Dai, ParaAT: A parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* **419**, 779–781 (2012).
85. 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).
86. W. He, J. Yang, Y. Jing, I. 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**, btad121 (2023).
87. 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, 2012), vol. 772 of *Methods in Molecular Biology*, pp. 3–12.
88. h. Öztoprak, "evolutionary persistence and speciation under ancient asexuality" dissertation, Universität zu Köln (2023). <https://kups.ub.uni-koeln.de/72279>.
89. h. Öztoprak, S. Gao, n. Guiguelmoni, A. Brandt, Y. Zheng, M. Erbil, v. Bednarski, c. Becker, K. Becker, I. Borgschulte, K. A. Burak, A.-M. Dion-Côté, v. Leonov, I. Opherden, S. Shimano,

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

Science AdvAnces | ReSeARch Article

J. Bast, data and code for chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*. <https://doi.org/10.5281/zenodo.12180072>.

Acknowledgments: We thank M. Maraun and K. Wehner for the electron microscopy picture of *P. peltifer*, M. Solbach for the *P. peltifer* drawing, M. Jean for collection of mites in Canada, and N. Kaneko (Fukushima University) for providing information about the previous study site as a habitat for *P. peltifer*. This work was supported by the dFG Research Infrastructure West German Genome Center (407493903) and nGS_cc (project 479) as part of the next Generation Sequencing competence network (project 423957469). We acknowledge computational support of the Zentrum für Informations und Medientechnologie, especially the hPc team (high Performance computing) at the Heinrich-Heine University. We sincerely thank the reviewers for insightful feedback and constructive suggestions, which have improved the quality of the manuscript. **Funding:** This work received support from German Research Foundation Emmy Noether grant BA 5800/3-1 (to J.B.) and German Research Foundation grant BA 5800/4-1 as part of the dFG Sequencing call (to J.B.). This work was supported by the dFG Research Infrastructure West German Genome Center (project 407493903) as part of the next Generation Sequencing competence network (project 423957469). nGS analyses were carried out at the West German Genome Center and production sites in Cologne and Düsseldorf. The work was also supported by nSeRc discovery grant RGPIn-2019-05744 (to A.-M.d.-c.). **Author contributions:** conceptualization: J.B. Resources: h.Ö., v.I., A.-M.d.-c., S.S., c.B., and K.B. Methodology: J.B., h.Ö., S.G., n.G., Y.Z., K.A.B., v.B., M.e., and A.B. Investigation: h.Ö., I.O., c.B.,

K.B., A.B., and v.I. Visualization: h.Ö., S.G., n.G., Y.Z., v.B., I.B., and M.e. Funding acquisition: J.B. Project administration: h.Ö. and J.B. Supervision: J.B., h.Ö., n.G., and S.G. Data curation: n.G. and S.G. Formal analysis: h.Ö., S.G., A.B., n.G., Y.Z., v.B., I.B., K.A.B., and M.e. Writing—original draft: J.B. and h.Ö., with input from all authors. Writing—review and editing: J.B. and h.Ö. Both h.Ö. and S.G. contributed substantially to this work. h.Ö. led the overall project (i.e., coordination and writing), organismal research, lab work, and analytical work; S.G. made substantial contributions, specifically through the analysis of the German reference genome, focusing on haplotypic differences. Both authors were additionally contributing as stated in the above-listed roles. For specific details on the overall project, please contact J.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. They are part of the dissertation of h.Ö., which provides additional details to the methodology (88). All codes for the analyses can be found at https://github.com/theBastLab/Ppr_evolution and have been deposited to Zenodo. The genome and the reads are available at NCBI and at the SRA under Bioproject PRJNA1003031. All data and codes to reproduce the analyses are additionally available at Zenodo under <https://doi.org/10.5281/zenodo.12180072> (88, 89).

Submitted 21 November 2023

Accepted 24 December 2024

Published 24 January 2025

10.1126/sciadv.adn0817

Downloaded from <https://www.science.org at Universitäts- und Stadtbibliothek Köln on February 23, 2025>

3. Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

5. General discussion

In this dissertation, I used two non-model species that reproduce primarily asexually to detect the biological implications of sexual reproduction. Including a facultatively sexual reproducing protist, *Fisculla terrestris* and an obligately asexual reproducing oribatid mite, *Platynothrus peltifer*. Investigating putative sexual reproduction in protists and examining the genomic features of long-term asexual mite lineages can provide new insights into how organisms persist, evolve and adapt with sex.

Sex is a conserved, ancient and widespread attribute of eukaryotic life

Sexual reproduction is thought to be a critical feature of LECA and is putatively present in most eukaryotes (Speijer 2016; Fu et al. 2019). Protists represent basal taxa of eukaryotes, however, it remains unclear whether sexual reproduction or a sexual cycle exists prevalently (da Silva and Machado 2022). A key contribution of my thesis is the detection of meiosis-associated genes and their expression change between the “lack of food” and “supplied food” treatments. Further, the potential occurrence of a cryptic sexual cycle in *F. terrestris* was presumed when cell fusion was induced by starvation. These findings implied there might exist more facultatively sexual organisms and sexual reproduction could be more prevalent than previously assumed. Under the microscope, when supplied with food, *F. terrestris* fed, digested, and underwent typical mitotic fission. In contrast, when food was depleted, most individuals became dormant. Notably, approximately 5% of the population fused into large aggregates with an extensive, highly active pseudopodial mass (**Figure. 1**, Chapter 1). I detected 11 meiosis-specific and 40 meiosis-related genes and found the meiosis-associated genes belong to the recombination-related pathway upregulated in the lack of food treatment, which suggests the starvation might have triggered the recombination process (**Figure. 3**, Chapter 1). Despite lack of direct evidence, it seems very likely that *F. terrestris* engaged in some form of cryptic sex. One evidence is starvation-induced cell fusion, a mechanism that might be widely conserved as a trigger for sexual reproduction across eukaryotes. Starvation can trigger the meiosis process in yeast (Mitchell and Herskowitz 1986). Similarly, it was also reported that mouse spermatogonia can enter meiosis under

conditions of nutrient restriction and in the presence of retinoic acid in vitro. This mechanism appears to be applicable to all mammals, as retinoic acid (RA) is a chordate morphogen, and its signaling has been the primary emphasis in mammals, whereas neither RA nor nutrient restriction alone can initiate meiosis (Zhang et al. 2021). Taken together, the evidence of starvation-induced cell fusion, along with the upregulation of recombination-related genes in the starvation treatment, strongly suggests the occurrence of cryptic sex and the potential existence of sexual reproduction in *F. terrestris*.

Sexual reproduction is the predominant mode of reproduction in most metazoans (Bell 1982). However, some oribatid mites are a notable exception, with approximately 10% of species being obligate asexuals (Norton et al. 1993). Asexuality is generally expected to lead to the extinction of lineages due to the accumulation of deleterious mutations and a reduced ability to adapt to environmental changes (Muller 1964; Felsenstein 1974). While many examples of asexuality exist, the asexual lineages capable of enduring and evolving over the long term are uncommon (Judson and Normark 1996). A key finding of this thesis is the detection of the Meselson effect in *P. peltifer* across different populations as a long-term asexual characteristic, which confirmed *P. peltifer* has been asexual for a long time, approximately 20 million years. In other words, *P. peltifer* is a rare example of obligate ancient asexuals and this rarity may indicate the prevalent use of sexual reproduction (Chapter 2).

New insight into sexual reproduction from *F. terrestris* and *P. peltifer*

While numerous theories have been proposed regarding how sex facilitates adaptation and evolution, “Why sex” is still an enigma in evolutionary biology (Otto 2009). By studying the role of sexual reproduction in basal, putatively facultatively sexual organisms and the genomic features of obligate ancient asexual organisms, I aim to provide a new insight on the adaptation and abundance of sexual reproduction.

Our investigation into the cryptic sex of *F. terrestris* revealed that starvation serves as a key trigger (abiotic factor) for cell fusion, potentially facilitating the initiation of a sexual reproduction cycle. Under microscopic observation, *F. terrestris* exhibited cell fusion following prolonged starvation, forming large aggregates characterized by highly active pseudopodial masses. Once food was reintroduced, the aggregates dispersed, and individual cells resumed their normal activity. This aggregation behavior may represent

an adaptive response to food scarcity, potentially facilitating cell fusion necessary for the exchange of genetic material. A previously similar hypothesis also said that the emergence of sexual reproduction resembled a cannibalistic event among unicellular organisms on primordial earth (Sagan and Margulis 1987).

Our results showed that recombination-related genes are upregulated in conditions of food deprivation and cell fusion compared to food-sufficient conditions. Recombination among fused cells could theoretically generate beneficial genetic combinations, enhancing survival during periods of environmental stress. Alternatively, it may serve to increase genetic variation before the cells separate, providing a mechanism to maintain genetic diversity despite predominantly asexual reproduction. Engaging in sexual reproduction under nutrient-deprived conditions may offer evolutionary advantages. It mitigates the potential drawbacks of continuous asexual reproduction under favorable conditions, such as the accumulation of deleterious mutations and reduced adaptability (Muller 1964; Felsenstein 1974). At the same time, it does not interfere with the rapid population expansion enabled by asexual reproduction when resources are abundant (Chapter 1). In this instance, cell fusion induced by starvation may serve as a defensive mechanism for *F. terrestris* against the stress of starvation.

However, it is rare for metazoans to reproduce asexually or facultatively sexual, and it has become widely accepted that sexual reproduction is essential for long-term evolutionary survival (Speijer et al., 2015). In this study, we confirmed an ancient asexual example, challenging the assumption that asexuality is an evolutionary dead end. One key finding of this study is the observation of conservation of one haplotype while relaxation of the other in *P. peltifer*. This pattern may indicate an adaptive mechanism in oribatid mites to mask the accumulation of harmful mutations, and relaxed selection one haplotype could provide a source of genetic novelty (Chapter 2). However, the pattern is likely influenced by multiple factors, including but not limited to environmental stress, body size, population size, etc. In addition, ancient asexuality might likely require independent haplotypes, but it might be restrained by the specific transition to asexuality and cytology that is needed to maintain haplotypic independence. If this hypothesis is correct, then the evolutionary strategies that allow obligate asexual mites to persist may not be easily expanded to other organisms compared to sex. After all, the special strategy of evolution and persistence is only ever found in parthenogenesis mites. Furthermore, it could provide proof that the evolutionary process does not readily replace sexual reproduction.

Sequence technology help studies on tiny non-model organisms

Both *F. terrestris* and *P. peltifer* are tiny organisms. *F. terrestris* is a unicellular protist that can be observable only under a microscope. While oribatid mites can be roughly visible to the naked eye, they are still very small, ranging from approximately 150 to 1400 μm in size. Neither organism (*F. terrestris* and *P. peltifer*) can be fully cultured in the laboratory to observe their whole life cycle, and due to their minute size, tracing their entire life span in natural environments is nearly impossible.

While *F. terrestris* can be fed and maintained under laboratory conditions for a period of time (wet lab work of Solbach and Dumack), the dynamics of its gene expression during starvation-induced cell fusion remain unknown. A key evidence in the *F. terrestris* study is the detection of the recombination-related gene expression changes. Without sequencing data, we would be unable to confirm the presence of meiosis-associated genes and track gene expression dynamics during cell fusion. This would significantly hinder our exploration of cryptic sex in protists (Chapter 1).

Prior to the release of our phased chromosome-level genome of *P. peltifer*, research on the Meselson effect in asexual mites faced several limitations. Schaefer et al. (2006) attempted to detect the Meselson effect using only two genes but were unsuccessful due to the low number of genes (Schaefer et al. 2006). The first successful detection of the Meselson effect in oribatid mites came from the genomic study of *Oppiella nova*, although the genome assembly remained fragmented (Brandt et al., 2021). We utilized high-quality, chromosome-level genome and haplotype-resolved assemblies to identify the clear genetic signatures of independent haplotype evolution in asexual mites (known as the Meselson effect) and their genomic feature, i.e., one haplotype is conserved and the other is relaxed. Importantly, uncovering the evolutionary implications of this phenomenon requires phased genomic data from single individuals sampled from natural populations. Such analysis has only become possible recently due to advances in sequencing technologies used in this study (Chapter 2).

Summary

In summary, this study utilized transcriptomic and genomic data from the facultatively sexual protist *F. terrestris* and the obligately asexual mite *P. peltifer* to explore the evolution of reproductive strategies and the genomic features of asexuality. Evidence of

starvation-induced cell fusion and the upregulation of meiosis-related genes in *F. terrestris* suggests a likely cryptic sexual process, challenging the long-standing assumption that most protists reproduce exclusively asexually. In contrast, earlier research suggested that protists reproduce asexually, likely due to limited observational data and a lack of detailed life cycle studies. In *P. peltifer*, the detection of the Meselson effect across populations confirms its long-term asexuality and reveals a key genomic characteristic: one haplotype copy is conserved and one under relaxed selection. These findings support the hypothesis that sex is a conserved and widespread feature of eukaryotic life. Advances in sequencing technologies have helped overcome these challenges, enabling the study of tiny, non-model organisms at the genomic level. Together, these two case studies offer a new perspective on the conservation and prevalence of sexual reproduction by illustrating a putative sexual basal protist and an exception of ancient asexual mites.

Outlook

In the future, the mites may serve as model organisms for studying long-term asexual evolution and offer new perspectives on the necessity of sexual reproduction. *F. terrestris* can be a representative of Rhizaria for exploring sexual reproduction in old taxa. Next, studies to detect sexual reproduction in protists would aim to detect genetic recombination of *F. terrestris* and explore if they exchanged genetic materials when the cell fused. To explore that, we need to sequence three different treatments: the mitosis stages, prolonged aggregation and re-supplied food, through comparing the exchanges of genotypes. Furthermore, we need to conduct additional cytological observations and experiments with population genetics to further detect sexual reproduction, like allele sharing and LD decay. Moreover, *F. terrestris* is a good model for studying meiotic sex of protists (Rhizaria). There are still many questions that need to be answered, like if *F. terrestris* produces gametes after prolonged aggregation, if they use conventional meiotic sex, or what the detail of potential cryptic sex is (if it exists). These questions are worth exploring in the future and provide new insights into the evolution of meiosis. The findings (Chapter 2) render oribatid mites as the best examples for ancient asexuality and one of the best model systems to explore the singularities that allow for successful asexuality over time. As one example showing the Meselson effect, they are confirmed to be obligately long-term asexuals. Future research may compare asexual lineages to

closely related sexual mites to elucidate the mechanisms and timeline of their transition to asexuality, as well as investigate how the evolution of novel traits may facilitate their adaptation.

Conclusion

In conclusion, this thesis represents a significant step towards a better understanding of the implication of sexual reproduction in eukaryotes via studying facultative and obligate asexual organisms. Additionally, the analyses pipelines and methods have far-reaching implications for the study of tiny and basal eukaryotes.

References of General Discussion

- Bell G. 1982. The masterpiece of nature: the evolution and genetics of sexuality. Routledge
- Brandt A, Van PT, 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.*
- Da Silva VS, Machado CR. 2022. Sex in protists: A new perspective on the reproduction mechanisms of trypanosomatids. *Genet. Mol. Biol.* 45:e20220065.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737–756.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737–756.
- Flot J-F, Hespeels B, Li X, Noel B, Arkhipova I, Danchin EGJ, Hejnal A, Henrissat B, Koszul R, Aury J-M, et al. 2013. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* 500:453–457.
- Fu C, Coelho MA, David-Palma M, Priest SJ, Heitman J. 2019. Genetic and genomic evolution of sexual reproduction: echoes from LECA to the fungal kingdom. *Curr. Opin. Genet. Dev.* 58–59:70–75.
- Judson OP, Normark BB. 1996. Ancient asexual scandals. *Trends Ecol. Evol.* 11:41–46.
- Otto SP. 2009. The evolutionary enigma of sex. *Am. Nat.* 174 Suppl 1:S1–S14.
- Sagan D, Margulis L. 1987. Cannibal's relief: the origins of sex. *New Sci.* 1971 115:36–40.
- Schaefer I, Domes K, Heethoff M, Schneider K, Schön I, Norton RA, Scheu S, Maraun M. 2006. No evidence for the “Meselson effect” in parthenogenetic oribatid mites (Oribatida, Acari). *J. Evol. Biol.* 19:184–193.

5. General discussion

- Smith RJ, Kamiya T, Horne DJ. 2006. Living males of the 'ancient asexual' Darwinulidae (Ostracoda: Crustacea). *Proc. R. Soc. B Biol. Sci.* 273:1569–1578.
- Speijer D, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci.* 112:8827–8834.
- Speijer D. 2016. What can we infer about the origin of sex in early eukaryotes? *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150530.
- Mitchell AP, Herskowitz I. 1986. Activation of meiosis and sporulation by repression of the RME1 product in yeast. *Nature* 319:738–742.
- Norton RA, Kethley JB, Johnston DE, O'Connor BM, Wensch DL, Ebbert MA. 1993. Phylogenetic perspectives on genetic systems and reproductive modes of mites.
- Muller HJ. 1964. The relation of recombination to mutational advance. *Mutat. Res. Mol. Mech. Mutagen.* 1:2–9.
- White MJD. 1954. Animal cytology & evolution. *Cambridge university press*
- Zhang X, Gunewardena S, Wang N. 2021. Nutrient restriction synergizes with retinoic acid to induce mammalian meiotic initiation in vitro. *Nat. Commun.* 12:1758.

6. Appendix

Chapter 1 Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*

6. Appendix

Table. 1 DESeq2 expression level of the 41 DEGs showed the raw expression.

name	131270_S158	131272_S159	131274_S160	131276_S161	131278_S162	131280_S163	131282_S164	131284_S165	131286_S166	131288_S167
HOP1	2645.88	3472.62	3622.21	3562.62	3130.79	1411.29	1295.89	1242.48	1305.37	1276.61
EXO1	867.83	884.85	921.46	977.13	830.58	327.90	378.34	369.19	409.19	368.41
SLX1	3350.15	3512.88	3488.83	3589.23	2861.94	1446.24	1500.08	1515.04	1554.63	1474.67
SMC6	747.49	867.47	931.25	975.29	859.79	427.60	368.61	407.48	387.58	377.65
SAD1	2193.32	2392.86	2526.98	2244.19	2331.13	1059.75	1165.06	1128.54	1133.91	1205.80
RAD17	600.87	651.52	684.06	778.95	709.47	423.49	375.68	404.74	439.44	447.43
RAD24	49552.15	56165.83	60039.37	59319.56	49720.67	34990.36	35087.05	35490.50	35084.97	33332.47
SPO11	318.66	314.78	336.52	280.75	304.92	192.22	216.57	195.99	200.27	192.93
LIG4	643.25	581.06	628.99	695.46	645.05	377.24	405.74	428.44	413.51	489.50
RAD23	6556.22	9212.73	8599.08	8358.36	7453.76	4862.94	6053.36	5673.67	5055.77	5149.54
MRE11	202.55	190.33	216.60	201.85	188.10	137.74	142.32	125.80	129.67	136.49
RAD51	357.64	323.01	301.04	314.70	402.84	244.64	271.38	245.21	226.21	224.74
REC8	408.49	440.14	516.41	451.41	397.68	319.67	334.14	343.67	334.27	324.28
KU	974.62	1030.35	1036.49	1024.84	999.79	873.71	865.40	912.49	876.01	831.23
SAE2	5249.38	5681.55	5816.34	5977.46	5400.07	5039.74	5191.50	5190.53	5045.69	4712.38
MER3	13815.00	12712.80	10439.55	11948.50	13717.05	14074.88	14229.12	14240.68	14742.29	14726.18
RAD50	13692.97	15642.80	15996.47	16202.00	14885.19	18986.13	18783.28	18507.78	19751.95	19132.74
MMS4	7177.43	8369.06	9270.90	8748.29	7476.95	10373.45	10180.58	10110.32	10429.97	10145.15
SLX4	744.10	611.25	685.28	607.38	718.06	851.09	799.98	895.17	851.51	859.97
SGS1	28986.85	35433.52	36627.16	35306.95	30229.91	41503.06	42653.76	42273.55	43013.70	41476.50
DMC1	1053.44	1162.12	1152.74	1272.56	1002.37	1296.17	1501.85	1477.67	1426.39	1559.85
MND1	199.16	187.59	186.01	148.63	176.08	217.91	289.06	231.54	219.00	217.56
MLH1	144.92	129.02	81.99	112.85	149.45	173.71	157.35	153.15	193.07	181.64
MLH3	119.50	97.00	112.58	105.51	97.06	128.49	156.46	152.23	158.49	137.51
MSH6	44455.31	48292.73	48369.98	49763.88	48773.27	87855.70	85609.83	84090.41	88273.62	84439.99
FEN1	339.00	348.63	456.45	455.08	348.72	360.79	321.76	336.37	334.27	350.97
MUS81	44.92	35.69	55.07	46.79	37.79	37.00	38.89	40.11	43.22	33.87
NBS1	1571.25	1729.45	1798.87	1828.56	1553.80	1541.83	1475.33	1460.35	1749.13	1571.13
MSH2	92.38	58.56	74.65	76.15	77.30	74.01	74.25	72.01	85.01	58.49
SMC5	128498.80	129580.40	145381.60	134313.60	107710.20	137969.40	132795.70	126163.40	120711.80	104528.10
MPH1	7241.84	7866.69	8020.26	8251.01	7080.13	7920.91	8038.74	7633.56	8265.88	7759.21
DNA2	2356.03	1730.36	1726.67	1805.63	1807.18	2093.81	2102.06	2066.55	2068.99	1992.91
PMS1	770.37	754.92	948.38	857.86	804.81	923.04	874.24	929.81	913.47	970.80
RAD54	1097.51	866.55	561.69	733.08	1175.87	1028.92	1046.61	1020.97	1061.87	994.40
HOP2	42.37	51.24	55.07	49.54	49.82	61.67	60.99	47.40	76.36	59.52
ATM	101.70	115.30	138.28	132.12	125.40	134.65	131.71	97.54	119.59	121.09
CDC2	738.17	840.02	840.70	911.99	832.30	432.74	418.11	429.35	374.61	446.40
KU80	493.24	488.64	532.32	539.49	480.14	433.77	422.53	461.26	453.85	482.32
MSH3	173.74	126.28	130.94	147.72	136.57	128.49	130.83	154.06	144.08	139.57
MSH5	10.17	6.41	18.36	11.01	7.73	15.42	15.03	16.41	7.20	14.37
RTEL1	358.49	354.12	349.98	375.25	344.43	474.88	436.68	473.11	453.85	461.80

Chapter 2 Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*



Supplementary Materials for

Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*

Hüsna Öztoprak *et al.*

Corresponding author: Hüsna Öztoprak, h.oeztoprak@uni-koeln.de; Shan Gao, sgao3@uni-koeln.de

Sci. Adv. **11**, eadn0817 (2025)
DOI: 10.1126/sciadv.adn0817

The PDF file includes:

Supplementary Text
Figs. S1 to S26
Tables S1 to S6
Legend for data S1

Other Supplementary Material for this manuscript includes the following:

Data S1

6. Appendix

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 a '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.

$$\mu = \frac{2}{(2 \times 488,839,965)} \rightarrow \mu = \frac{2}{(2 \times 488,839,965)} = 2.05 \times 10^{-9}$$

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*.

6. Appendix

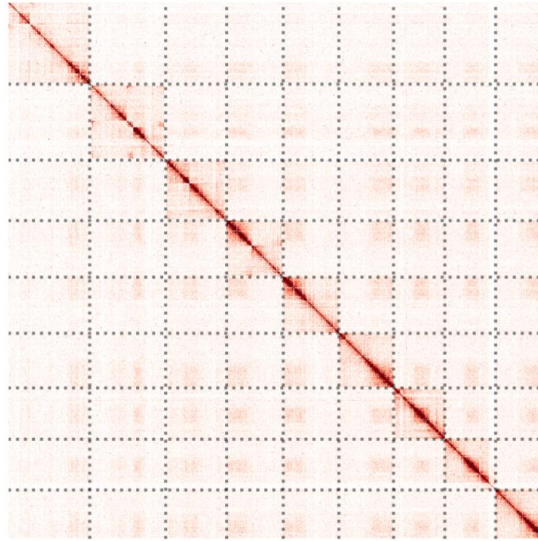


Fig. S1

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.

6. Appendix

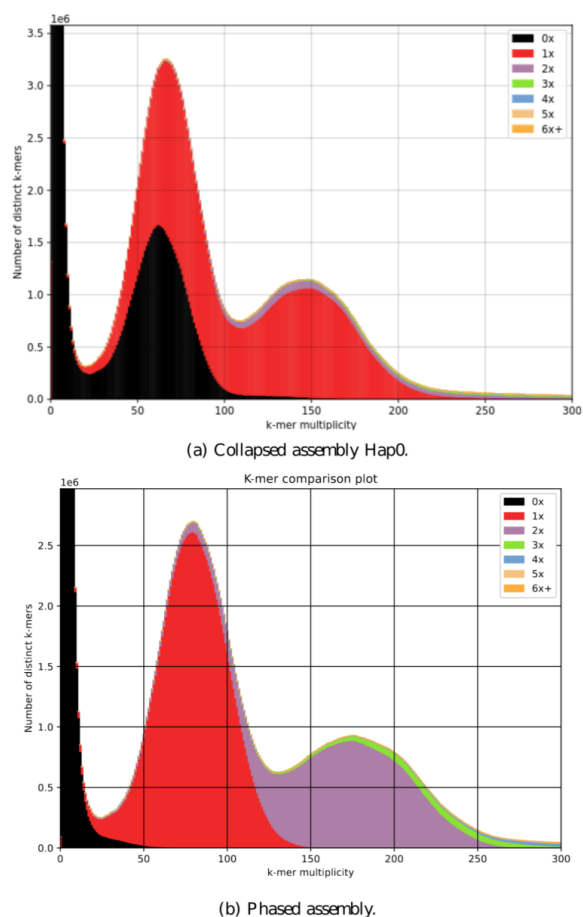


Fig. S2.

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 parts of heterozygous k -mers are represented once (red) and parts 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).

6. Appendix

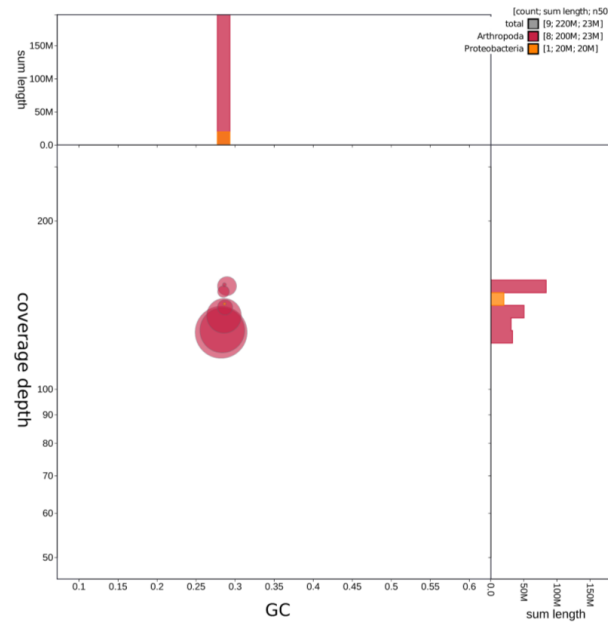


Fig. S3.

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.

6. Appendix

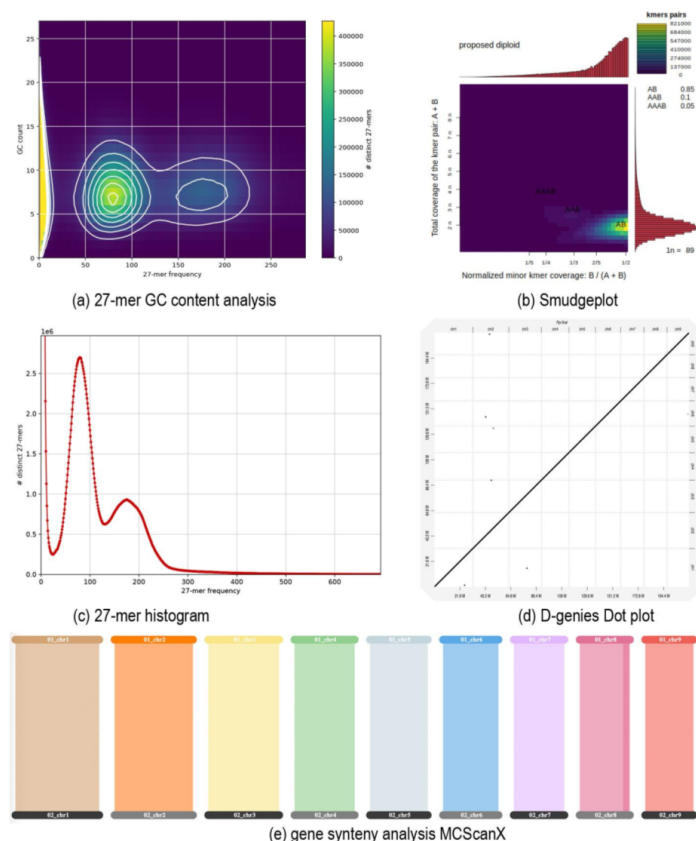


Fig. S4.

Genome property analyses based on k -mers of the PacBio HiFi reads ($k = 27$). a) The two peaks identified in histogram (c) have similar GC content, and there is no extra peak at a higher GC content, which does not suggest bacterial contamination in the reads. b) The Smudgeplot shows a distinct smudge of k -mers with an A/B configuration and the genome is identified as diploid without duplications. c) 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. d) Dot plot generated with synteny analysis with D-GENIES supporting ploidy level. e) Pairwise collinearity between haploid assembly. Synteny blocks within the collapsed assembly were obtained using MCSanX.

6. Appendix

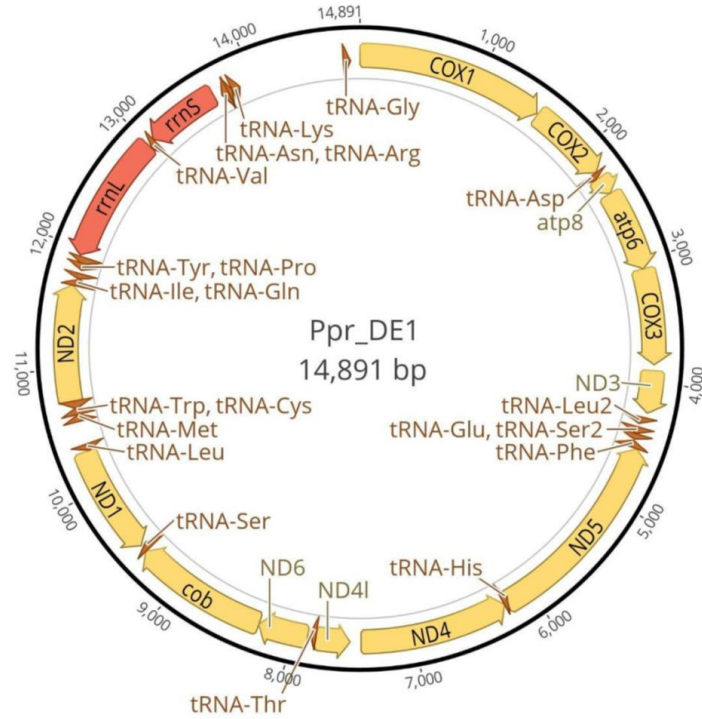


Fig. S5.

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.

6. Appendix

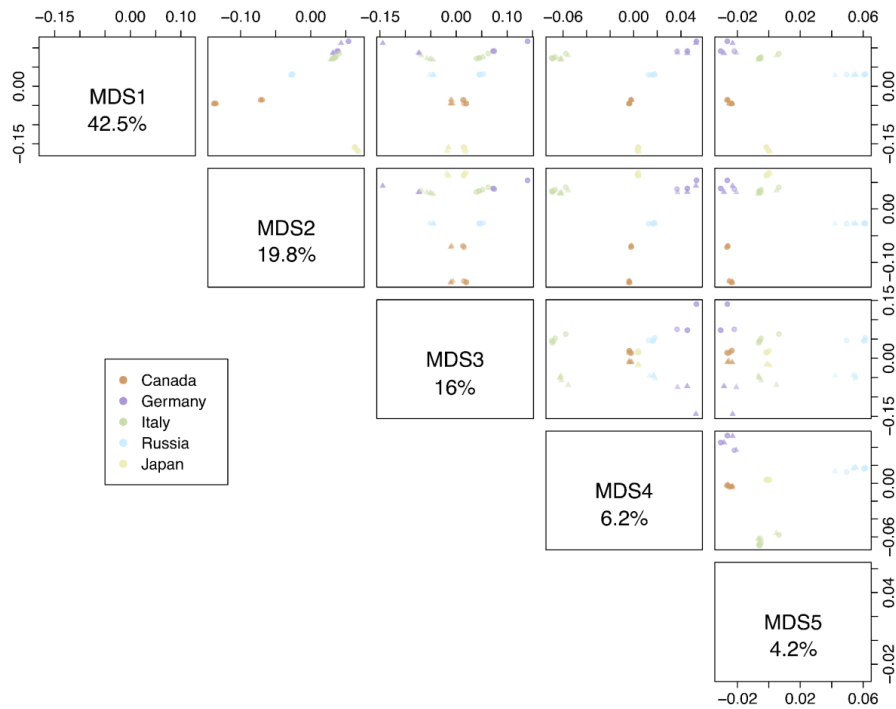


Fig. S6.

Multidimensional scaling (MDS) analysis using 416,588 LD-pruned genome-wide SNPs from five individuals belonging to the five populations of *Platynothrus peltifer*. The scatterplot matrix displays the five eigenvectors and their respective percentages of explained variation. Note: MDS1 primarily separates samples based on their population origin, while MDS3 splits the samples based on the two haplotypes (triangle for haplotype A and point for haplotype B) mirroring each other.

6. Appendix

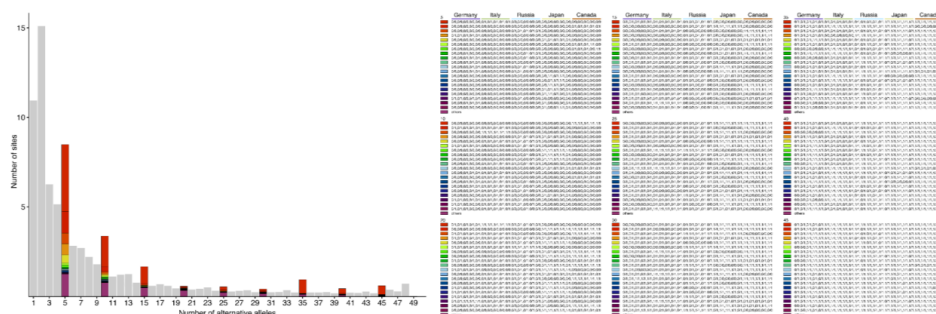


Fig. S7.

Site frequency spectrum (SFS) shows the alternate allele counts on the x-axis in a sample size of $n = 25$ diploid individuals (five from each population), and the number of sites with that count on the y-axis. Colored bars depict the number of sites with a given genotype across all 25 individuals. Horizontal bars at the top of the legend highlight the genotypes of all five individuals in each population. For example, the majority of sites with exactly 5 alternate alleles are those where all Italian individuals are heterozygous, while individuals from other populations are homozygous for the reference allele. In contrast, the majority of sites with an alternate allele count of 15 are those where all 15 individuals from Europe are heterozygous, indicating shared heterozygosity before the split, while individuals from Canada and Japan are homozygous for the reference allele.

6. Appendix

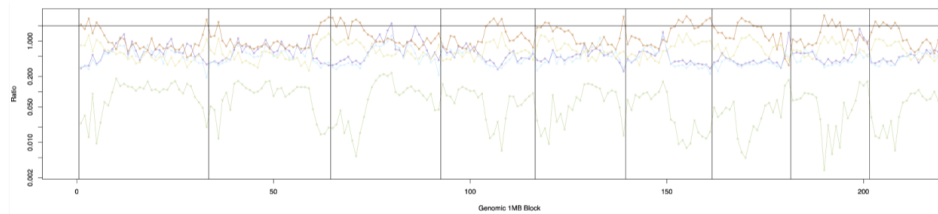


Fig. S8.

The ratio of unshared to shared heterozygous 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 $R_x = 2$. 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 ones will stay in an equilibrium (new mutations vs fixation/loss within haplotype). Given enough time, R_x 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 a few heterozygous sites. 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.

6. Appendix

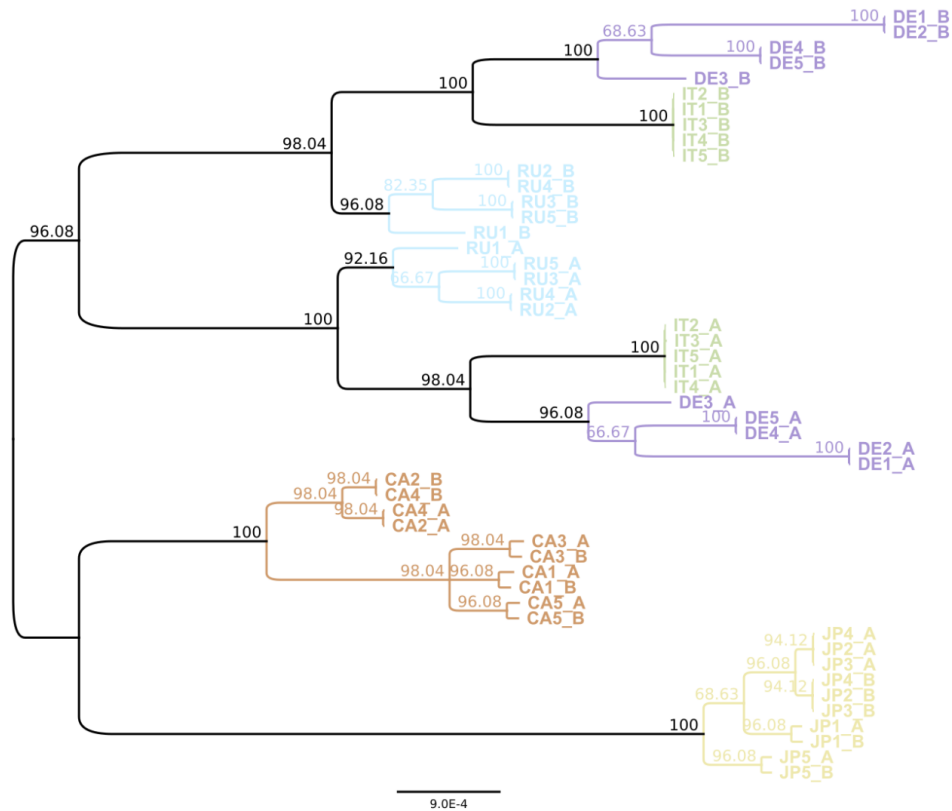


Fig. S9.

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.

6. Appendix

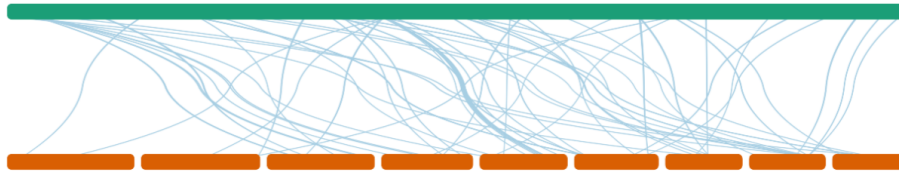


Fig. S10.

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

6. Appendix

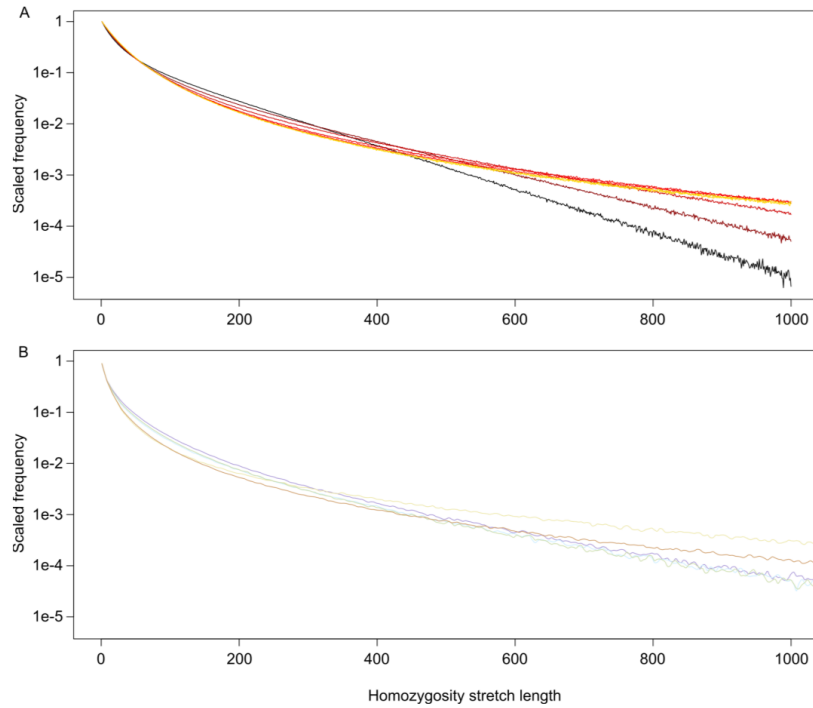


Fig. S11.

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. a 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. 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.

6. Appendix

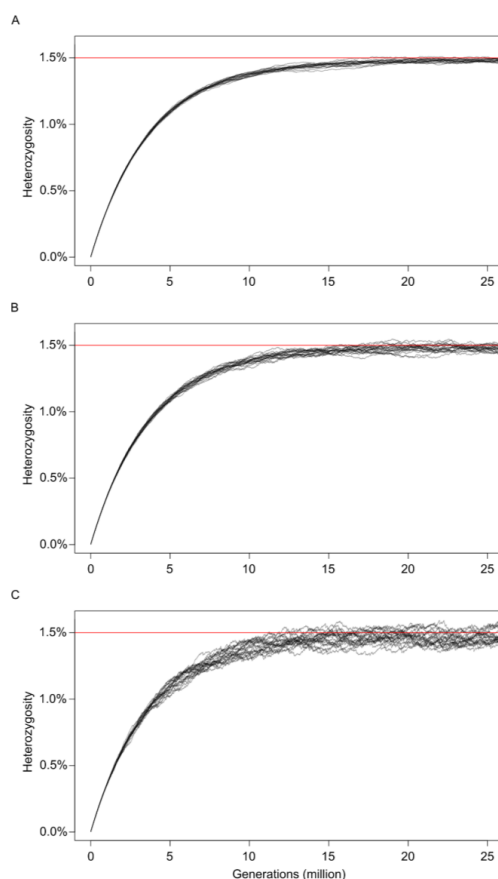


Fig. S12.

Simulations of heterozygosity change (i.e., divergence of haplotypes) over time under different gene conversion track lengths 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 was 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.

6. Appendix

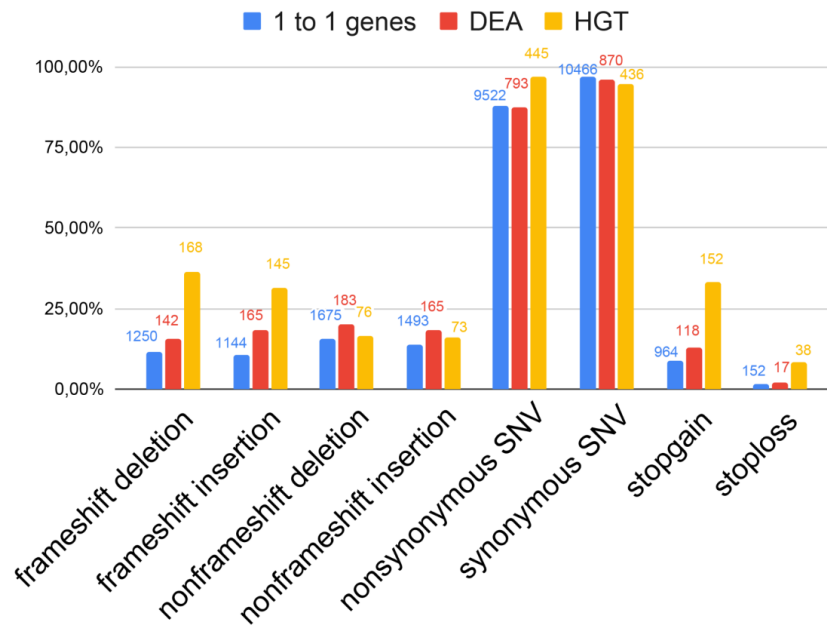


Fig. S13.

Categories of different SNP effects and their percentage in orthologous genes (blue), differentially expressed alleles (DEAs; red), and horizontal gene transfers (HGTs; yellow). The sum of the 1 to 1 copy genes, DEA and HGT genes are 10816, 905 and 460, respectively.

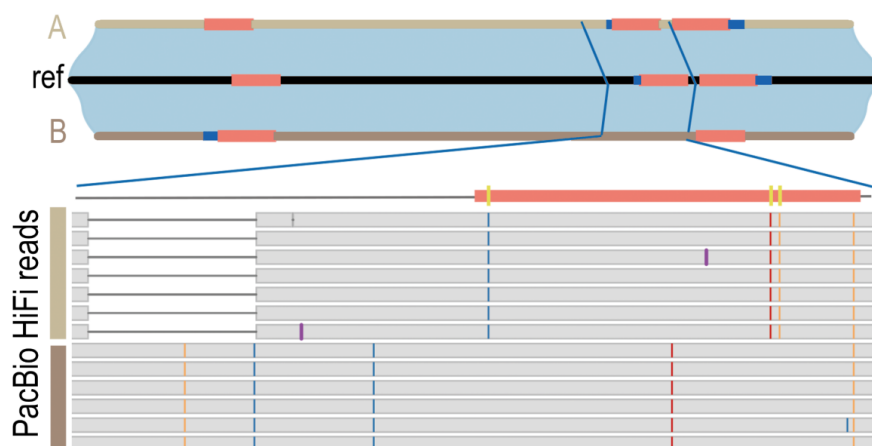


Fig. S14.

Differentially expressed HGT (*Ppr_hap0_g20454:chr8:683873-688817*) 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. This gene was annotated to be a GDSSL-like Lipase/Acylhydrolase-like horizontally transferred gene from the Acidobacteria which are one of the most abundant groups across soil ecosystems. The Blastp results suggest similarity to pectin esterase (identify=96 %, coverage=59.48 %), which was found in all higher plants as well as in some bacteria and fungi. This HGT might have contributed to facilitating the breaking down of cell walls or pectin, which is in line with their function as soil decomposers.

6. Appendix

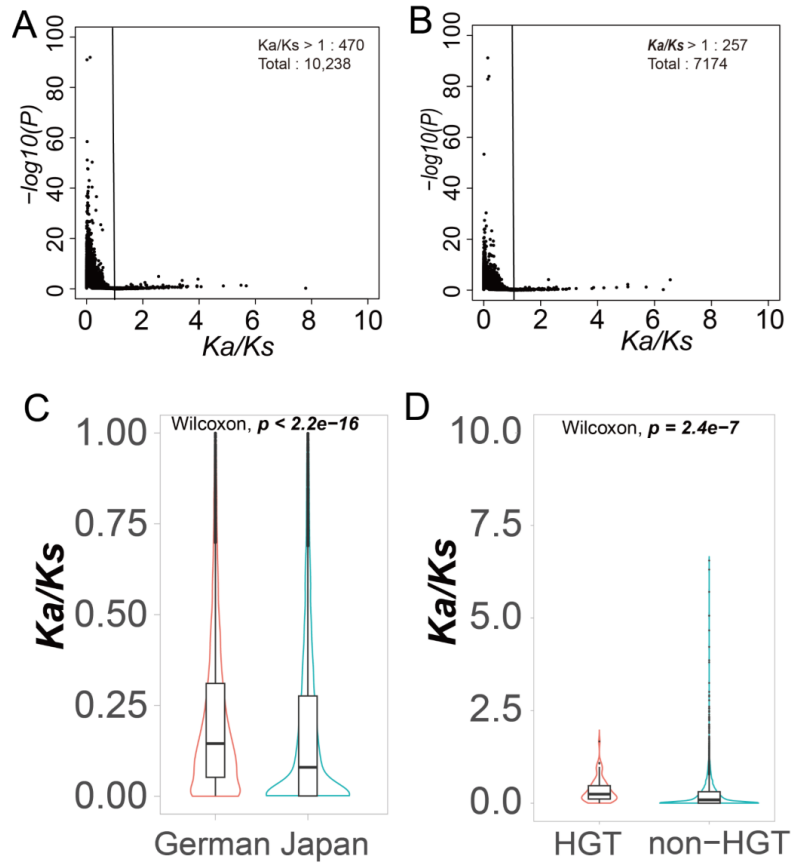


Fig. S15.

(A) Pairwise comparison of the K_a/K_s distribution for German *Platynothrus peltifer*'s allelic genes. Outlier $K_a/K_s=0.225$ p -value= 5.36587×10^{-228} removed from plot. (B) K_a/K_s distribution for allelic genes in the Japanese lineage. This K_a/K_s landscape shows K_a/K_s of 7,174 single copy alleles, among those 257 alleles with $K_a/K_s > 1$, which is less than in the German lineage, consistent with the hypothesis that the German lineage is long-term asexual while the Japanese lineage is not. (C) The K_a/K_s comparison of the single copy alleles between the German and Japanese lineages indicates higher K_a/K_s mean values in German *P. peltifer*. (Effect size=0.181, small). (D) The K_a/K_s comparison between HGT genes and non-HGT genes of the Japanese lineage. Similar to the German lineage, K_a/K_s mean values in Japanese HGTs are higher than in non-HGT regions (Effect size=0.0609, small).

6. Appendix

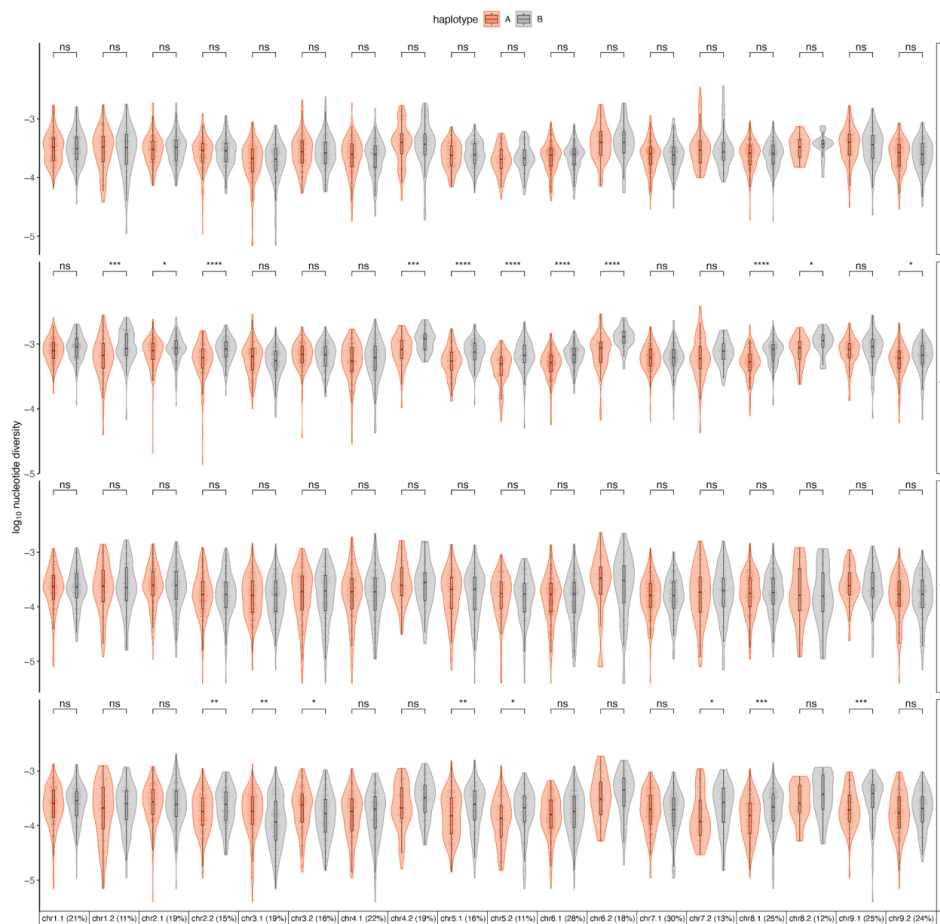


Fig. S16.

Nucleotide diversity (π) for the first and second largest haplotypic blocks in four populations of *Platynothrus peltifer* from Canada, Germany, Japan and Russia. Italy was excluded as individuals from this population showed very little genetic variation. The x-axis shows chromosomes (chr'X') along with the largest (chr'X'.1) and second largest haplotypic blocks (chr'X'.2). The percentage in parentheses indicates the fraction of the chromosome covered by the respective haploblock. All estimates were calculated in 25,000 site windows sliding by 10,000 sites (Wilcoxon rank-sum test; n.s. = non-significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

6. Appendix

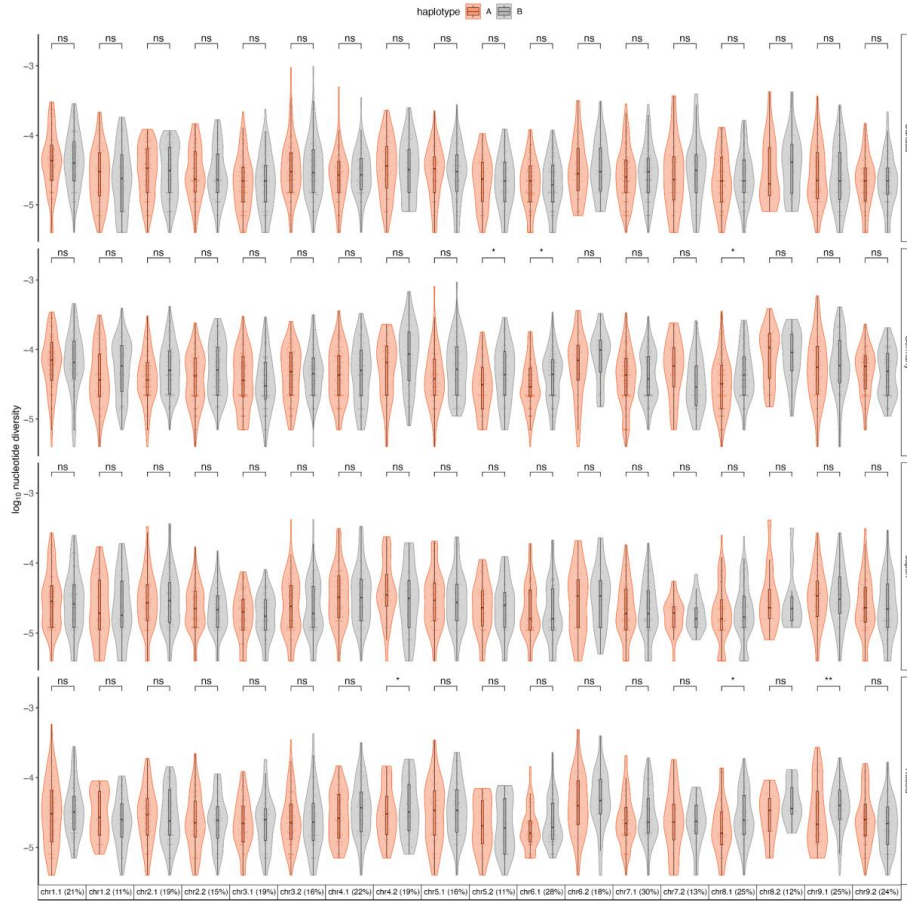


Fig. S17.

Nucleotide diversity (π) at nonsynonymous sites for the first and second largest haplotypic blocks in four populations of *Platynothrus peltifer* from Canada, Germany, Japan and Russia. Italy was excluded as individuals from this population showed very little genetic variation. The x-axis shows chromosomes (chr'X') along with the largest (chr'X'.1) and second largest haplotypic blocks (chr'X'.2). The percentage in parentheses indicates the fraction of the chromosome covered by the respective haploblock. All estimates were calculated in 25,000 site windows sliding by 10,000 sites (Wilcoxon rank-sum test; n.s. = non-significant; * $p < 0.05$; ** $p < 0.01$).

6. Appendix

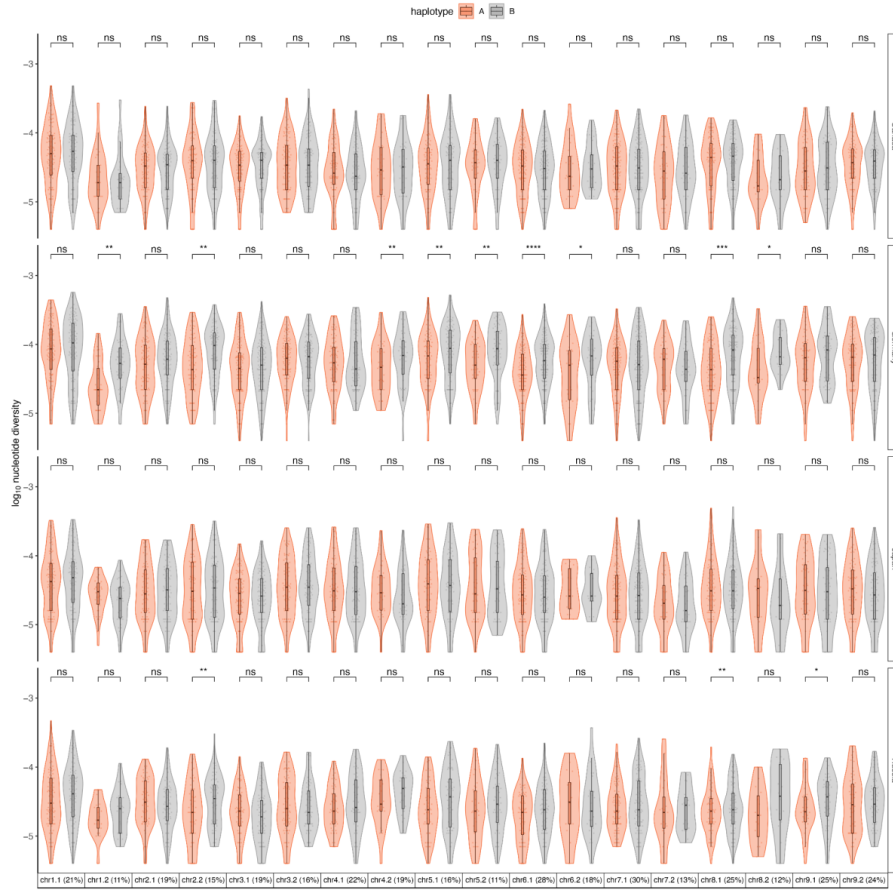


Fig. S18.

Nucleotide diversity (π) at synonymous sites for the first and second largest haplotypic blocks in four populations of *Platynothrus peltifer* from Canada, Germany, Japan and Russia. Italy was excluded as individuals from this population showed very little genetic variation. The x-axis shows chromosomes (chr'X') along with the largest (chr'X'.1) and second largest haplotypic blocks (chr'X'.2). The percentage in parentheses indicates the fraction of the chromosome covered by the respective haploblock. All estimates were calculated in 25,000 site windows sliding by 10,000 sites (Wilcoxon rank-sum test; n.s. = non-significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

6. Appendix

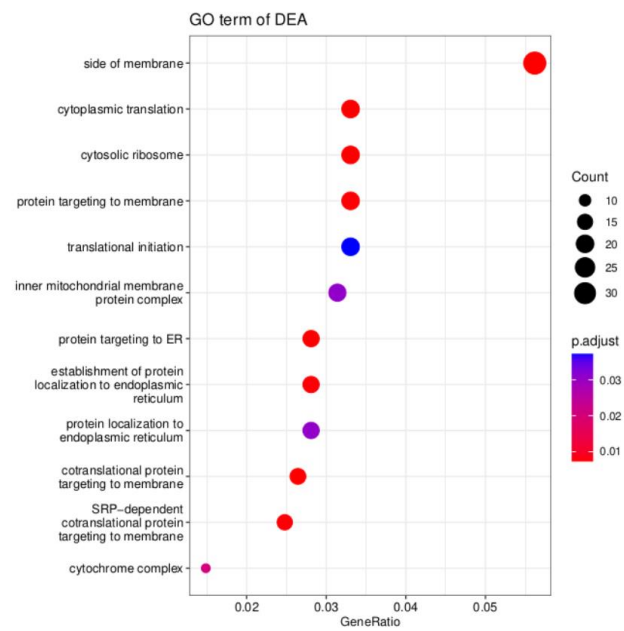


Fig. S19.

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

6. Appendix

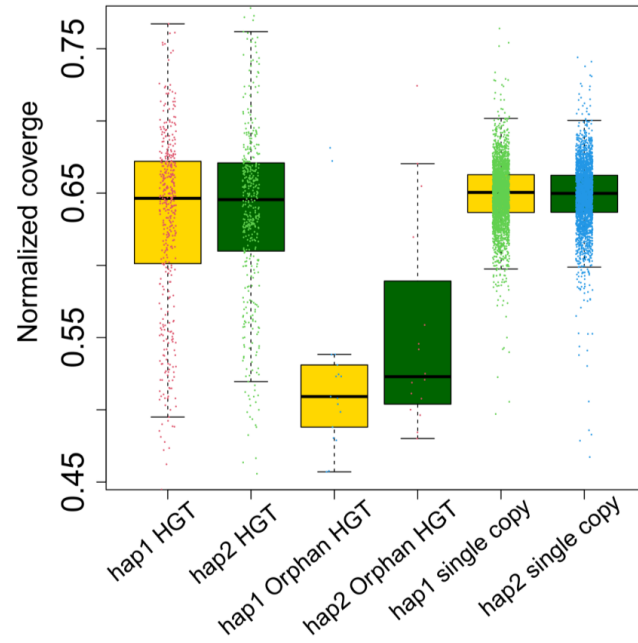


Fig. S20.

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 (hap1 HGT=428, hap2 HGT=415, hap1 orphan HGT=17, hap2 orphan HGT=16, hap1 single copy=2694, hap2 single copy=2677).

6. Appendix

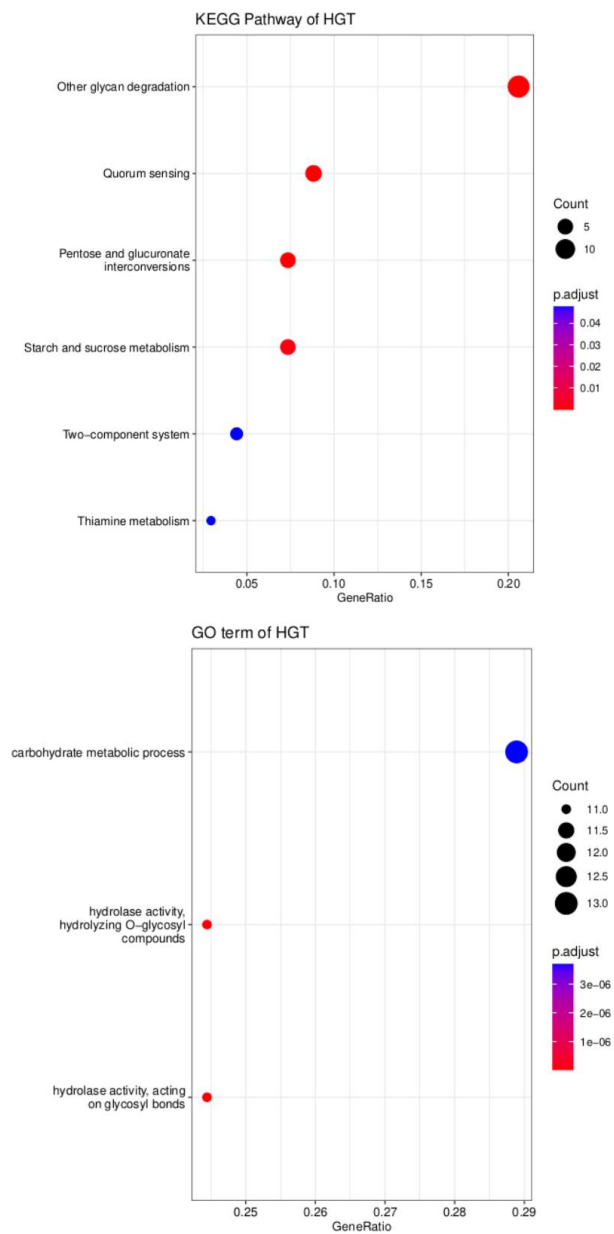


Fig. S21.

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

6. Appendix

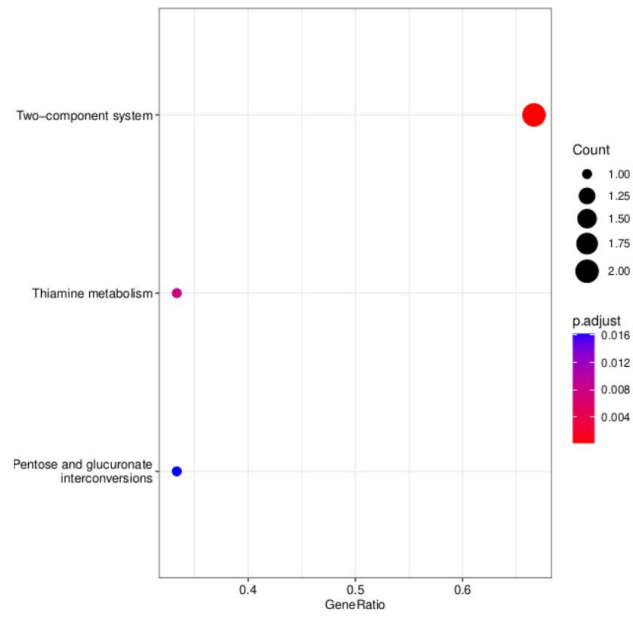


Fig. S22.

KEGG pathways enrichment of HGT genes that exhibit differential expression.

6. Appendix

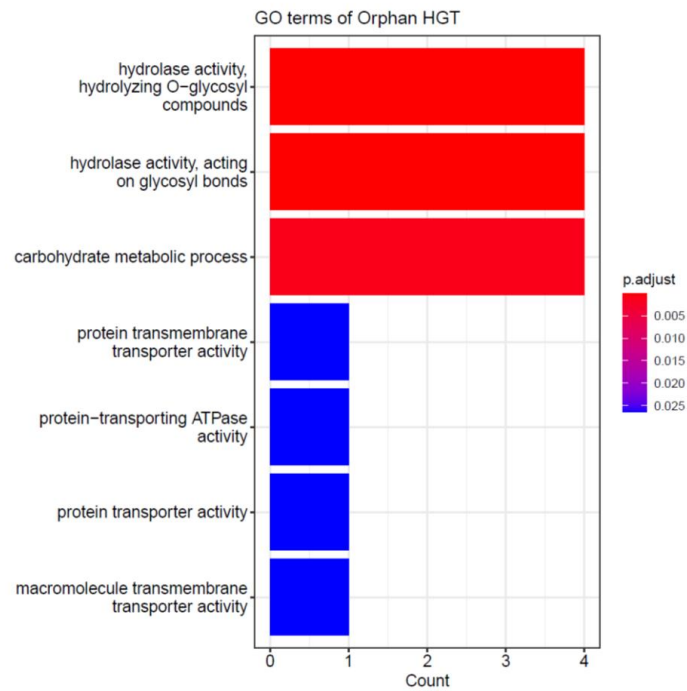


Fig. S23.

GO term enrichment for the orphan HGTs.

6. Appendix

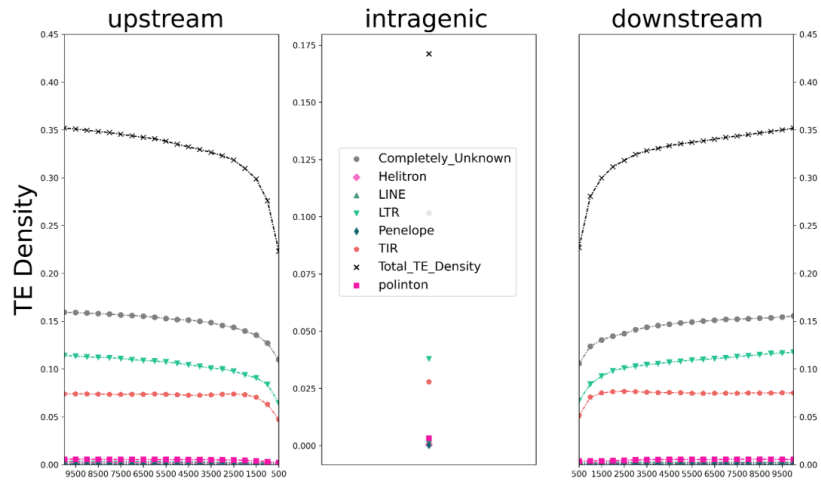


Fig. S24.

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).

6. Appendix

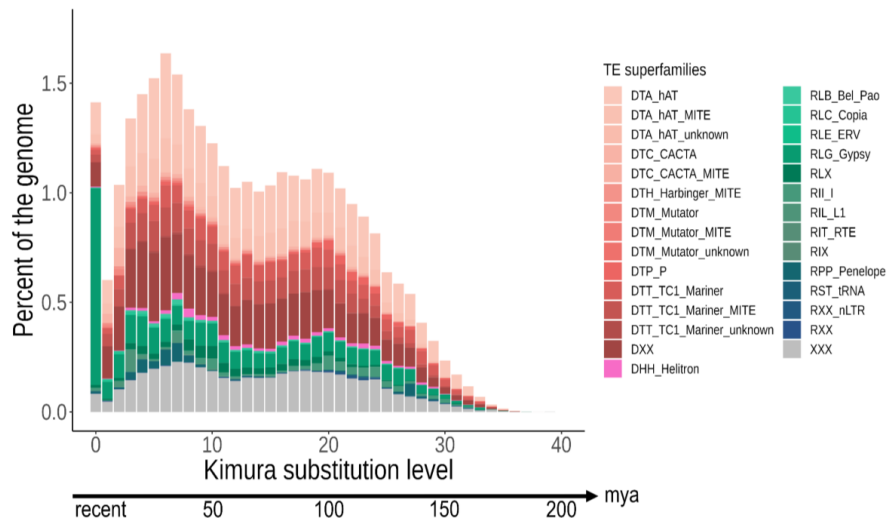


Fig. S25.

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 29-59 my ago (6 %-12 % divergence), coinciding with the increase of global temperature dynamics.

6. Appendix

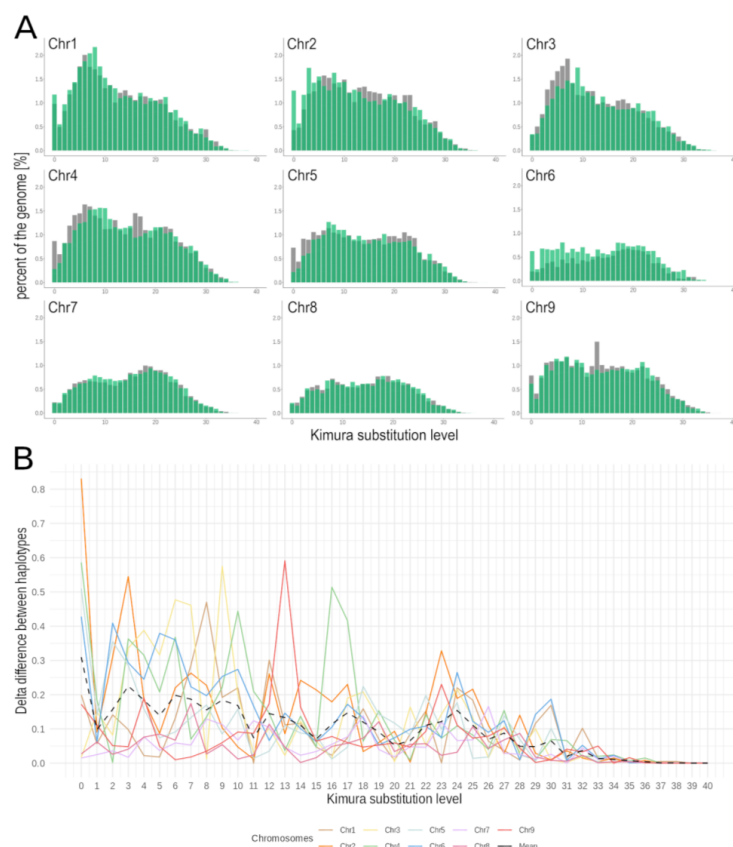


Fig. S26.

(A) Transposable element divergence landscapes of the largest haplotypic blocks per chromosome show slight but noticeable haplotype-specific historical activity. Green bars represent haplotypic block A and gray bars haplotypic block B. While haplotype TE dynamics differ among chromosomes ($p=0.0002$, non-parametric permutation ANOVA), dynamics differ between haplotypes overall in chromosome 6 ($p=0.029$, non-parametric permutation ANOVA). Chromosomes 7 and 8 show no difference between haplotypes in TE dynamics. Chromosomes 1,2,3,5,6,9 feature different episodes of differential TE activity between haplotypes. (B) Delta-differences between the haplotypes of each chromosome regarding the TEs as percent of the genome for each position of Kimura substitution level including the mean delta-difference across all chromosomes.

6. Appendix

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

6. Appendix

Table S2.

Fragment sizes [bp] of haplotypic blocks A and B (in bold the longest scaffold in each chromosome used for 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

6. Appendix

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

6. Appendix

Table S3.

Summary statistics for the longest in-phase region 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 (collapsed assembly region) length
alt1_chr 1_1	14957471	3454174	9999998	+	chr1	32608290	3587612	10326678	5228739	6864089 (6739067)
alt1_chr 2_5	10752214	28598	5717847	+	chr2	30492103	89265	5854202	4686177	5862286 (5764938)
alt1_chr 3_3	5119837	32321	5119734	+	chr3	27587832	45418	5170894	4366896	5196015 (5125477)
alt1_chr 4_1	4929078	15814	4928941	+	chr4	23397169	14530458	19553083	4185504	5063857 (5022626)
alt1_chr 5_1	13159702	2582264	6054325	+	chr5	22442908	4686733	8214567	3064880	3552017 (3527835)
alt1_chr 6_5	7722251	1802320	7722233	+	chr6	21557542	13485156	19401970	5149241	6050008 (5916815)
alt1_chr 7_1	15067032	7073018	12787819	+	chr7	19673716	7426596	13193208	4927722	5883881 (5766613)
alt1_chr 8_1	10309134	3280164	7993029	+	chr8	19511964	8588947	13277971	4100966	4785435 (4689025)
alt1_chr 9_1	17136559	7136559	11645221	+	chr9	18765566	7084104	11663639	3571460	4657571 (4579536)

6. Appendix

Table S4.

Number of genes exhibiting differential omega (d_N/d_S) values between the two alleles (haplotypes) in the German (showing haplotype independence) and Japanese (without haplotype independence) populations. Out of 7,404 tested genes, over 320 show differential omega values between alleles (omega values close to 0 for one allele and close to 1 for the other). In Japan, only one gene shows such a pattern, while the remaining 7,403 genes show no difference in omega values.

Germany vs Japan		
Population	Allele subject to purifying selection	Count
Germany	A	321
Germany	B	329
Germany	none	6754
Japan	B	1
Japan	none	7403

6. Appendix

Table S5.

The effect size of the SNP density on five kinds of gene regions compared between differentially expressed alleles (DEA) and equivalently expressed alleles (EEA).

comparison	group1	group2	effect size	n1	n2	magnitude
exon_SNP	DEA	EEA	0.0669	936	11174	small
intron_SNP	DEA	EEA	0.0192	822	9139	small
UTR3_2kb_SNP	DEA	EEA	0.0311	936	11174	small
UTR3_SNP	DEA	EEA	0.0377	467	5816	small
UTR5_2kb_SNP	DEA	EEA	0.0283	936	11174	small
UTR5_SNP	DEA	EEA	0.0127	466	5567	small

6. Appendix

Table S6.

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
Total TEs in genome	27.389

6. Appendix

Data S1. (separate file)

Datasheets include detailed information about assembly and annotation metrics of *Platynothrus peltifer*, including fragment size of haplotypic blocks and longest alignment blocks; details on population data, mutation rate and telomeric regions, *Wolbachia* homology to *P. peltifer*, horizontal gene transfers (HGT) and orphan HGT candidates and categories of different SNP effects.

6. Appendix

7. Declaration

„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:

Gao S, Solbach MD, Bast J, Dumack K. 2025. Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*. *BMC Biol.* 23:140.

Öztoprak H, **Gao S**, et al. 2025. Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer* *Sci. Adv.* 11:eadn0817.

Datum, Name und Unterschrift
2025.02.23 Shan Gao



8. Acknowledgements

First and foremost, I would like to express my sincere gratitude to my dissertation chair and committee members for their time and efforts on my paper. My appreciation goes to Prof. Dr. Gunther Doehlemann for serving as my committee chair and host the defense. I am also extremely grateful to my committee members, Prof. Dr. Bart Thomma and Prof. Dr. Thomas Wiehe, for their time and comments (will happen soon) enhanced the quality of this work.

I would like to extend my heartfelt thanks to my mentor, Dr. Jens Bast, whose unwavering support, patience, and wisdom have been instrumental in my academic growth. Your mentorship has shaped not only my study but also my development as a scientist.

To my colleagues in the AG Bast and AG Bonkowski, thank you for creating such a collaborative and supportive environment. Special thanks to Dr. Kenneth Dumack for the stimulating discussions, technical assistance, and collaboration that made this journey both educational and enjoyable. I am also grateful to Dr. Isabell Witt from the GSfBS for the organization of the workshops and help. I also want to thank the colleagues and friends from AG Zhang in the University of Shanxi, Prof. Dr. TingTing Zhang, Prof. Dr. Siegfried Roth.

I am profoundly grateful to my family for their unconditional love and support throughout my academic pursuits. To my parents, Changye Gao (高昌业) and Xianzhen Zhang (张先珍), your encouragement and faith in me have been a constant source of strength. To my friends, thank you for providing the much-needed balance in my life and for your understanding during the demanding times of this research.

9. Curriculum vitae

Shan Gao



Personal Contact

Address Institute for Zoology, University of Cologne, Zùlpicher Str 47b, 50674 Köln, Germany

Contact +49 1634003337 (mob.)

E-mail gaoshan12138@gmail.com, sgao4@smail.uni-koeln.de

Born May 11, 1994 in XianYang, China

Nationality China

Academic interests

Genome assembly, Comparative genomics, Differently expressed Alleles, Evolution of Asexuality, Reproduction, Meiosis and Mitosis, Population genomics

Education and Work Experiences

2020.11-Now	Ph.D in Biology , Institute for Zoology, University of Cologne, Germany
2016.7-2019.12	M.Sc. in Animal Breeding , College of Animal Science and Technology, Northwest A&F University, China
2012.9-2016.7	B.Sc. in Biotechnology , College of Innovation and Experiment, Northwest A&F University, China
2006.9-2012.7	Rainbow High School , Xianyang, China
2000.9-2006.7	Rainbow Primary School , Xianyang, China
1997.9-2000.7	Rainbow Kindergarten , Xianyang, China

Academic Performance and Publication

Doctoral work:

- ◆ Öztoprak H*, **Gao S***, Guiglielmoni N, Brandt A, Zheng Y, Errbii M, Bednarski V, Becker C, Becker K, Borgschulte L, et al. 2025. Chromosome-scale genome dynamics reveal signatures of independent haplotype evolution in the ancient asexual mite *Platynothrus peltifer*. *Science Advances* 11:eadn0817.
- ◆ **Gao S**, Solbach MD, Bast J, Dumack K. 2025. Meiosis-associated expression patterns during starvation-induced cell fusion in the protist *Fisculla terrestris*. *BMC Biol.* 23:140.

Master work:

- ◆ Li Z, Wang X, Zhang Y, Yu Z, Zhang T, Dai X, Pan X, Jing R, Yan Y, Liu Y, **Gao S**, et al. 2022. Genomic insights into the phylogeny and biomass-degrading enzymes of rumen ciliates. *ISME J* 16:2775–2787.
- ◆ **Gao S**, Nanaei HA, Wei B, Wang Y, Wang X, Li Z, Dai X, Wang Z, Jiang Y, Shao J. 2020. Comparative Transcriptome Profiling Analysis Uncovers Novel Heterosis-Related Candidate Genes Associated with Muscular Endurance in Mules. *Animals* 10:980.
- ◆ Shao J, He S, Pan X, Yang Z, Nanaei HA, Chen L, Li R, Wang Y, **Gao S**, Xu H. 2020. Allele-specific expression reveals the phenotypic differences between thin-and fat-tailed sheep.
- ◆ Cao Y, Xu H, Li R, **Gao S**, Chen N, Luo J, Jiang Y. 2019. Genetic Basis of Phenotypic Differences Between Chinese Yunling Black Goats and Nubian Goats Revealed by Allele-Specific Expression in Their F1 Hybrids. *Frontiers in Genetics*
- ◆ Li R, Fu W, Su R, Tian X, Du D, Zhao Y, Zheng Z, Chen Q, **Gao S**, Cai Y, et al. 2019. Towards the Complete Goat Pan-Genome by Recovering Missing Genomic Segments From the Reference Genome. *Frontiers in Genetics*
- ◆ Wang Y, **Gao S**, Zhao Y, Chen W-H, Shao J-J, Wang N-N, Li M, Zhou G-X, Wang L, Shen W-J, et al. 2019. Allele-specific expression and alternative splicing in horse x donkey and cattle x yak hybrids. *Zoological Research* 40:293–304.

Instrumental Skills

- ◆ Programming with Python, Linux shell and plotting with R and Adobe Illustrator.
- ◆ Bjob system used to submit the jobs on high-performance computing cluster.
- ◆ Genome assembly (HiFiasm, Flye, Canu, Minimap2, Juicer etc.)
- ◆ Population genetics (BWA, GATK, ANGSD, VCFTOOLS, Samtools etc.)
- ◆ RNA-Seq (STAR, StringTie, DESeq2, rMATs etc.)
- ◆ Comparative genomics (Last, LASTZ, Blastn etc.)

Conferences

- ◆ Genomic signatures and consequences of sex and asexuality (ESEB Satellite Symposium, online, 2021)
(Talk) Genome of *P. peltifer* provide a new insight of the ancient asexual genomics landscape
- ◆ Congress of the European Society for Evolutionary Biology (ESEB, Prague, Czech, 2022)
(Poster) Evolutionary innovation in the long-term absence of sex in the oribatid mite *P. peltifer*
- ◆ Reference genomes for biodiversity (DERGA, Cologne, Germany, 2022)
(Poster) Evolutionary innovation in the long-term absence of sex in the oribatid mite *P. peltifer*
- ◆ 13th Acarological Colloquium (Göttingen, Germany, 2022)
(Poster) Evolutionary innovation in the long-term absence of sex in the oribatid mite *P. peltifer*
- ◆ Long-Read Sequencing (Uppsala, Sweden, 2022)
(Talk) Genomic signatures of single *P. peltifer* from chromosome-scale genome assembly with resolved haplotypes
- ◆ Evolution of Unisexual Reproduction (Mainz, Germany, 2023)
(Abstract) Pan-genome comparisons to identify the impact of long-term absence of sex on structural variation
- ◆ Society for Molecular Biology and Evolution (Ferrara, Italy, 2023)
(Poster) Horizontal gene transfer and differential allele expression of the oribatid mites *P. peltifer* provide new insight of asexual genome evolution
- ◆ Sex unfolded: sex, asex, sexes (Roscoff, France, 2023)
(Poster) Horizontal gene transfer as substrate for evolutionary novelty under ancient asexuality
- ◆ The 3rd AsiaEvo Conference (Singapore, 2023)
(Poster) Why sex? insights from asexual genomes