Check for updates

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 Guiglielmoni[1], Alexander Brandt[2], Yichen Zheng[1,3], Mohammed Errbii[1], Viktoria Bednarski[1], Christian Becker[4], Kerstin Becker[5], Lea Borgschulte[1], Katharina Atsuko Burak[1], Anne-Marie Dion-Côté[6], Vladislav Leonov[7], Linda Opherden[1], Satoshi Shimano[8], Jens Bast[1]

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

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

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

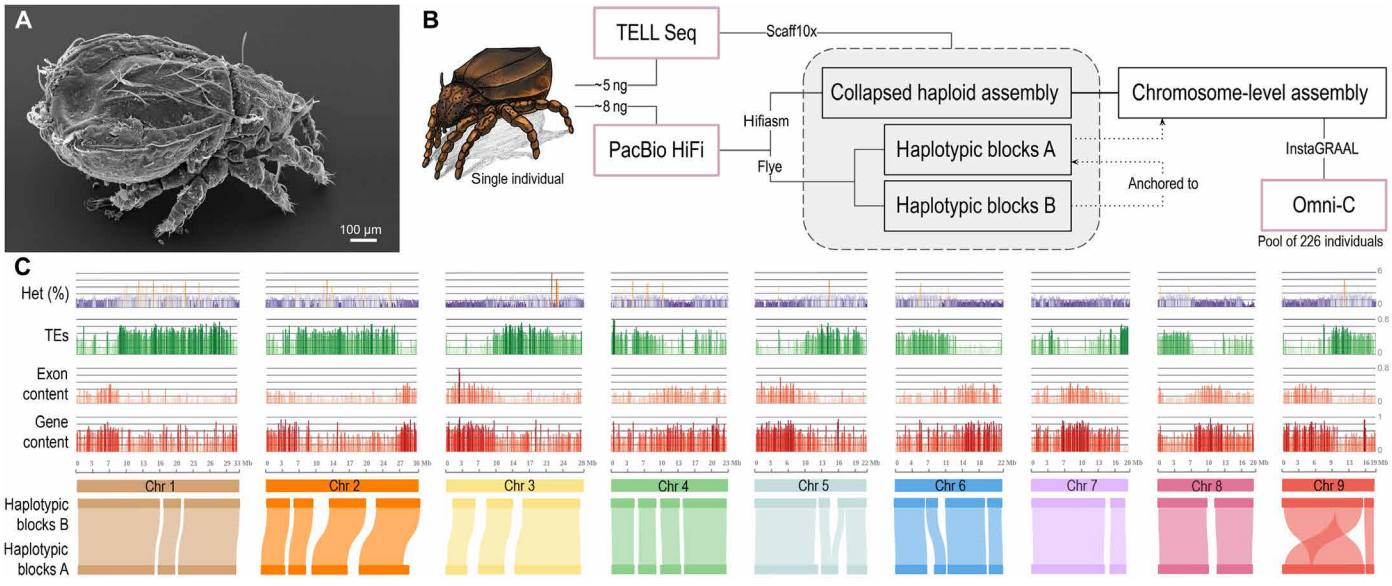


Fig. 1. Assembly strategy and characteristics of the *P. peltifer* genome. (**A**) *P. peltifer* electron microscopy picture. Scale bar, 100 µm. (**B**) Schematic workflow of single-individual sequencing and assembly. (**C**) Proportion of both aligned and collinear regions for haplotypic blocks A (98% 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

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

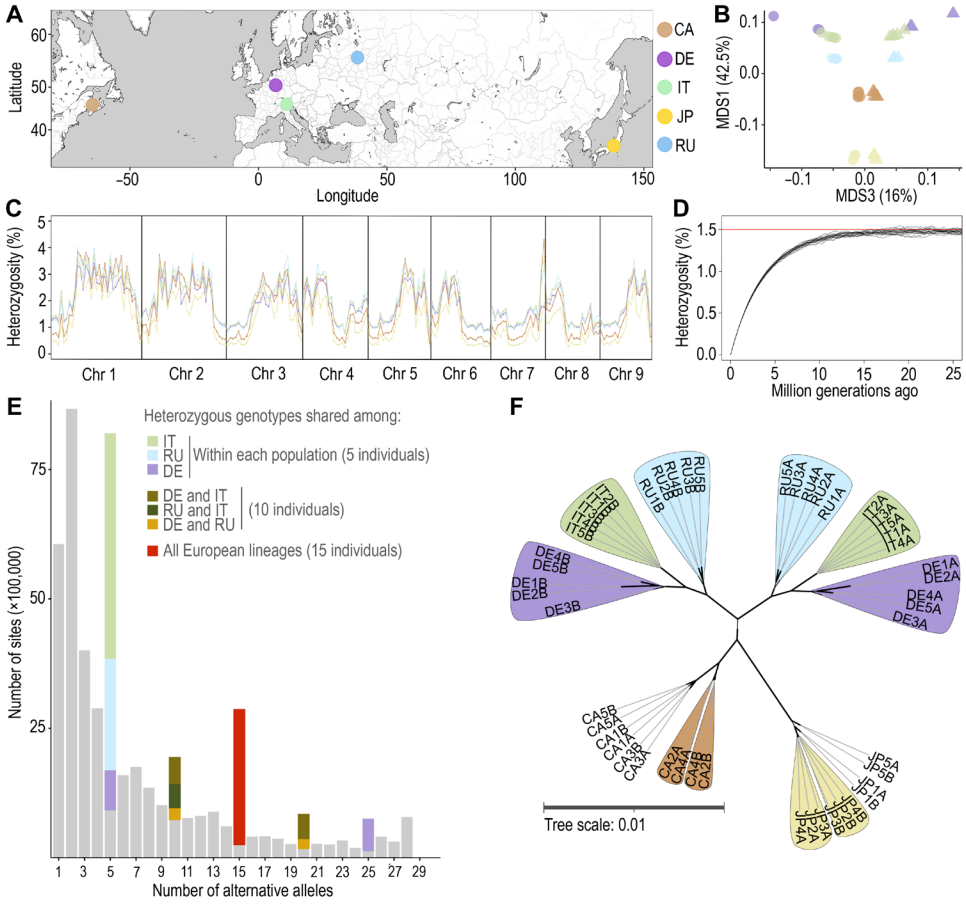


Fig. 2. Haplotype dynamics of worldwide *P. peltifer* populations suggest independent haplotype evolution for at least 20 million years under asexuality of the DE-IT-RU lineages. (**A**) Sampling locations worldwide, with water bodies shaded in gray. Abbreviations: CA, Canada; DE, Germany; IT, Italy; 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

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

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.

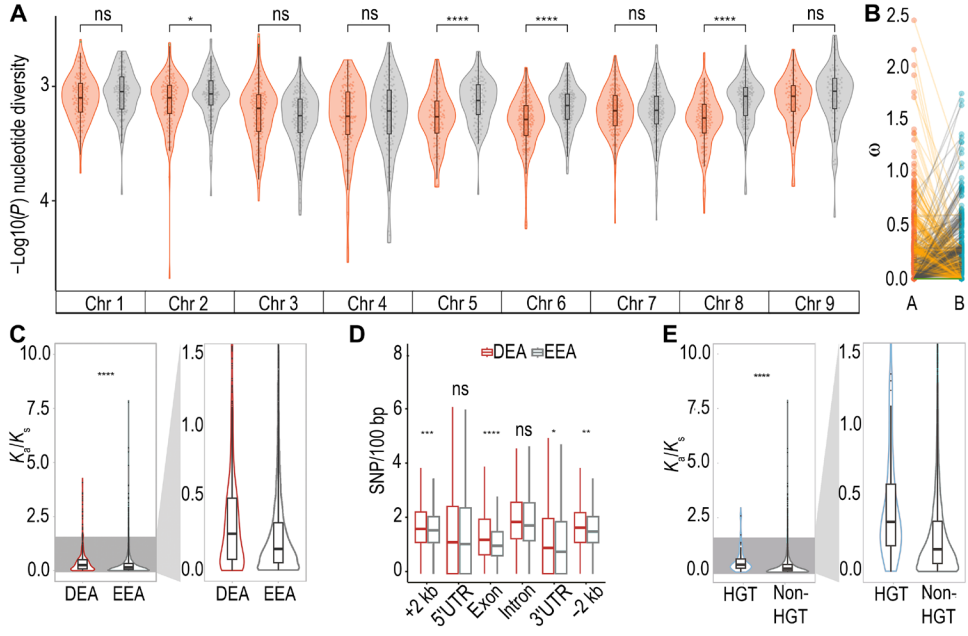


Fig. 3. Haplotype-specific genome evolution. (**A**) Nucleotide diversity (π) for the largest haplotypic blocks in German *P. peltifer*. All estimates were calculated in 25,000 site windows sliding by 10,000 sites. (**B**) Allele-specific evolutionary rates (ω; see table S4 for count data). (**C**) Violin 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; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$).

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

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 quasihomeologs. 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-cm^2 square. Italian samples were collected at Montan Southern Tyrol (46.33514°N, 11.29728°E) in a 30-cm^2 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

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

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 H_2O 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 BigDye 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 N_2, 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 × 10^6 pairs of 150-bp reads.

Linked-read sequencing for populations

In addition to the reference individual, four individuals of the same population were sequenced using Illumina TELL-seq linked paired-end reads only. For details, see 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 × 10^6 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

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

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. BRAKER2 (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, MCScanX (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/merrbii/Pcal-SocialPolymorphism/blob/main/PcalGenomics.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 November 18, 2026

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 TELLread (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).

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

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 ~>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 (~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 November 18, 2026

To estimate omega (ω) or d_N/d_S 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 "log10((best-hit *E*-value for INGROUP) + 1×10^{-200}) − log10((best-hit *E*-value for OUTGROUP) + 1×10^{-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 ≥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 orthogroup 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 ≥ |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 log2 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 NGenomeSyn (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×" 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 November 18, 2026

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 µg/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. Heimburger, K. S. Jaron, M. Labédan, M. Maraun, D. J. Parker, M. Robinson-Rechavi, I. Schaefer, P. Simion, S. Scheu, T. Schwander, J. Bast, Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101485118 (2021).
2. M. Heethoff, K. Domes, M. Laumann, M. Maraun, R. A. Norton, S. Scheu, High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *J. Evol. Biol.* **20**, 392–402 (2007).
3. A. Brandt, I. Schaefer, J. Glanz, T. Schwander, M. Maraun, S. Scheu, J. Bast, Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* **8**, 873 (2017).
4. J. Bast, I. Schaefer, T. Schwander, M. Maraun, S. Scheu, K. Kraaijeveld, No accumulation of transposable elements in asexual arthropods. *Mol. Biol. Evol.* **33**, 697–706 (2016).
5. I. Schaefer, R. A. Norton, S. Scheu, M. Maraun, Arthropod colonization of land – Linking molecules and fossils in oribatid mites (Acari, Oribatida). *Mol. Phylogenet. Evol.* **57**, 113–121 (2010).
6. M. Heethoff, R. A. Norton, S. Scheu, M. Maraun, "Parthenogenesis in oribatid mites (Acari, Oribatida): Evolution without sex" in *Lost Sex: The Evolutionary Biology of Parthenogenesis*, I. Schoen, K. Martens, P. van Dijk, Eds. (Springer Press, 2009), pp. 241–257.
7. R. A. Norton, S. C. Palmer, "The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites" in *The Acari: Reproduction, Development and Life-History Strategies*, R. Schuster, P. W. Murphy, Eds. (Springer Netherlands, 1991), pp. 107–136.
8. M. Maraun, M. Heethoff, K. Schneider, S. Scheu, G. Weigmann, J. M. Cianciolo, R. H. Thomas, R. A. Norton, Molecular phylogeny of oribatid mites (Oribatida, Acari): Evidence for multiple radiations of parthenogenetic lineages. *Exp. Appl. Acarol.* **33**, 183–201 (2004).
9. R. Butlin, The costs and benefits of sex: New insights from old asexual lineages. *Nat. Rev. Genet.* **3**, 311–317 (2002).
10. M. Terwagne, E. Nicolas, B. Hespeels, L. Herter, J. Virgo, C. Demazy, A.-C. Heuskin, B. Hallet, K. Van Doninck, DNA repair during nonreductional meiosis in the asexual rotifer *Adineta vaga*. *Sci. Adv.* **8**, eadc8829 (2022).
11. G. Taberly, Recherches sur la parthénogenèse thélytoque de deux espèces d'acariens oribatides: *Trhypochthonius tectorum* (Berlese) et *Platynothrus peltifer* (Koch). IV: Observations sur les males ataviques. *Acarologia* **29**, 95–107 (1988).
12. S. C. Palmer, R. A. Norton, Genetic diversity in thelytokous oribatid mites (Acari; Acariformes: Desmonomata). *Biochem. Syst. Ecol.* **20**, 219–231 (1992).
13. S. 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 automictic populations. *Genetics* **206**, 993–1009 (2017).
17. B. Dunn, T. Paulish, A. Stanbery, J. Piotrowski, G. Koniges, E. Kroll, E. J. Louis, G. Liti, G. Sherlock, F. Rosenzweig, Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. *PLOS Genet.* **9**, e1003366 (2013).
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. Lenormand, *Cis*-regulator runaway and divergence in asexuals. *Evolution* **72**, 426–439 (2018).
20. L. 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. Schuseil, D. Roze, T. Schwander, Asexual reproduction reduces transposable element load in experimental yeast populations. *eLife* **8**, e48548 (2019).
23. E. B. Chuong, 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. L. Payne, A. Wagner, The causes of evolvability and their evolution. *Nat. Rev. Genet.* **20**, 24–38 (2019).
26. D. Kempson, M. 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.yxmvm3yybl3p/v1.
30. O. Folmer, M. Black, W. Hoeh, R. Lutz, 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. Guiglielmoni, 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. Guiglielmoni, 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. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
36. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
37. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
38. L. Baudry, N. Guiglielmoni, H. Marie-Nelly, A. Cormier, M. Marbouty, K. Avia, Y. L. Mie, O. Godfroy, L. Sterck, J. M. Cock, C. Zimmer, S. M. Coelho, R. Koszul, instaGRAAL:

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

Chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. *Genome Biol.* **21**, 148 (2020).
39. M. Xu, L. Guo, S. Gu, O. Wang, R. Zhang, B. A. Peters, G. Fan, X. Liu, X. Xu, L. Deng, Y. Zhang, TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Gigascience* **9**, giaa094 (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, L. Lebeigle, M. Kirsche, K. Jenike, S. Ou, S. Aganezov, X. Wang, Z. B. Lippman, M. C. Schatz, S. Soyk, Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258 (2022).
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. Li, 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. Liao, K. Chougule, J. R. A. Agda, A. J. Hellinga, C. S. B. Lugo, T. A. Elliott, D. Ware, T. Peterson, N. Jiang, C. N. Hirsch, M. B. Hufford, Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 (2019).
48. E. A. Bell, C. L. Butler, C. Oliveira, S. Marburger, L. Yant, M. I. Taylor, Transposable element annotation in non-model species: The benefits of species-specific repeat libraries using semi-automated EDTA and DeepTE *de novo* pipelines. *Mol. Ecol. Resour.* **22**, 823–833 (2022).
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. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, A. H. Schulman, A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973–982 (2007).
51. S. J. Teresi, M. B. Teresi, P. P. Edger, TE Density: A tool to investigate the biology of transposable elements. *Mob. DNA* **13**, 11 (2022).
52. V. Bandi, C. Gutwin, J. N. Siri, E. Neufeld, A. Sharpe, I. Parkin, Visualization tools for genomic conservation. *Methods Mol. Biol.* **2443**, 285–308 (2022).
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. Neph, M. S. Kuehn, A. P. Reynolds, E. Haugen, R. E. Thurman, A. K. Johnson, E. Rynes, 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. Maraun, S. Scheu, S. L. Cameron, The complete mitochondrial genome of the sexual oribatid mite *Steganacarus magnus*: Genome rearrangements and loss of tRNAs. *BMC Genomics* **9**, 532 (2008).
56. R. Allio, A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, F. Delsuc, MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Mol. Ecol. Resour.* **20**, 892–905 (2020).
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 Acid Res.* **40**, 2833–2845 (2011).
58. M. Bernt, A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsch, J. Pütz, M. Middendorf, P. F. Stadler, MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **69**, 313–319 (2013).
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. Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, I. L. 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. L. A. Bergeron, S. Besenbacher, T. Turner, C. J. Versoza, R. J. Wang, A. L. Price, E. Armstrong, M. Riera, J. Carlson, H.-Y. Chen, M. W. Hahn, K. Harris, A. S. Kleppe, E. H. López-Nandam, P. Moorjani, S. P. Pfeifer, G. P. Tiley, A. D. Yoder, G. Zhang, M. H. Schierup, The Mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *eLife* **11**, 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. Lunter, 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, L. 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, L. L. Wang, M. Coon, T. Nguyen, L. 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. L. 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. Guiglielmoni, D. L. 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. Le, Y. Li, 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. Li, 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, PAL2NAL: 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. L. Kosakovsky Pond, A. F. Y. Poon, R. Velazquez, S. Weaver, N. L. 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, L. 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, L. Xu, K. Yu, X. Fang, NGenomeSyn: An easy-to-use and flexible tool for publication-ready visualization of syntenic relationships across multiple genomes. *Bioinformatics* **39**, 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. Guiglielmoni, A. Brandt, Y. Zheng, M. Errbii, V. Bednarski, C. Becker, K. Becker, L. Borgschulte, K. A. Burak, A.-M. Dion-Côté, V. Leonov, L. Opherden, S. Shimano,

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

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.L., 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., L.O., C.B., K.B., A.B., and V.L. Visualization: H.Ö., S.G., N.G., Y.Z., V.B., L.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., L.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 has 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 Universitaets und Stadtbibliothek Koeln on November 18, 2026

ScienceAdvances

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

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

Sci. Adv. **11** (4), eadn0817. DOI: 10.1126/sciadv.adn0817

View the article online
https://www.science.org/doi/10.1126/sciadv.adn0817
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org at Universitaets und Stadtbibliothek Koeln on November 18, 2026

Use of this article is subject to the Terms of service

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

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).