Phylloxera and aphids show distinct features of genome evolution despite similar reproductive modes

Zheng Li¹*, Allen Z. Xue¹, Gerald P. Maeda¹, Yiyuan Li¹,², Paul D. Nabity³, Nancy A. Moran¹

¹Department of Integrative Biology, The University of Texas at Austin, Austin, TX, 78712, USA;
²Institute of Plant Virology, Ningbo University, Ningbo, Zhejiang, 315211, China;
³Department of Botany and Plant Sciences, University of California Riverside, Riverside, CA, 92521, USA;

Corresponding author: Zheng Li, liz7@arizona.edu

Abstract

Genomes of aphids (family Aphididae) show several unusual evolutionary patterns. In particular, within the XO sex determination system of aphids, the X chromosome exhibits a lower rate of interchromosomal rearrangements, fewer highly expressed genes, and faster evolution at nonsynonymous sites compared to the autosomes. In contrast, other hemipteran lineages have similar rates of interchromosomal rearrangement for autosomes and X chromosomes. One possible explanation for these differences is the aphid’s life cycle of cyclical parthenogenesis, where multiple asexual generations alternate with one sexual generation. If true, we should see similar features in the genomes of Phylloxeridae, an outgroup of aphids which also undergoes cyclical parthenogenesis. To investigate this, we generated a chromosome-level assembly for the grape phylloxera, an agriculturally important species of Phylloxeridae, and identified its single X chromosome. We then performed synteny analysis using the phylloxerid genome and 30 high-quality genomes of aphids and other hemipteran species. Unexpectedly, we found that the phylloxera does not share aphids’ patterns of chromosome evolution. By estimating interchromosomal rearrangement rates on an absolute time scale, we found that rates are elevated for aphid autosomes compared to their X chromosomes, but this pattern does not extend to the phylloxera branch. Potentially, the conservation of X chromosome gene content is due to selection on XO males that appear in the sexual generation. We also examined gene duplication

© The Author(s) 2023. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
patterns across Hemiptera and uncovered horizontal gene transfer events contributing to phylloxera evolution.

**Keywords:** chromosomal rearrangement, sex chromosome, horizontal gene transfer

**Introduction**

Aphids (Insecta: Hemiptera: Aphididae) are a monophyletic group of about 5,000 species that feed on plant sap and include some globally distributed agricultural pests. They show remarkable life cycles incorporating cyclical parthenogenesis, in which several asexual, all-female generations are interspersed with a single sexual generation. Cyclical parthenogenesis depends on modifications of meiosis, including non-meiotic reproduction during all-female generations, elimination of an X chromosome to produce XO sons, and production of only X sperm by males to yield only XX daughters (Moran 1992; Davis 2012, Blackman 1980).

Aphids also exhibit distinctive patterns of genome evolution. Recently, chromosome-level assemblies have revealed that aphid X chromosomes display long-term conservation of gene content and arrangement, contrasting with relatively frequent autosomal rearrangements including autosomal translocations (Li et al. 2019; Mathers et al. 2021). Despite this conservation of gene content, aphid X-linked genes show low expression and elevated rates of nonsynonymous substitution, consistent with weak purifying selection, and a trend towards male-biased expression (Jaquiéry et al. 2018; Li et al. 2019; Jaquiéry et al. 2013; Mathers et al. 2021). Genes that are highly expressed, critical to all life stages (for example, those encoding ribosomal proteins), and/or single-copy are heavily concentrated on autosomes and largely lacking in the X chromosome. These observations suggest that selection favors placing functionally critical genes on autosomes, and that genes remaining on the X chromosome undergo less stringent purifying selection.

The cyclical parthenogenesis of aphids is hypothesized to underlie this distinctive pattern of genome evolution (Jaquiéry et al. 2018; Li et al. 2019). If true, these genomic features are predicted to extend to the phylloxera (family Phylloxeridae), a related lineage in the same infraorder as aphids, Aphidomorpha. Aphids and phylloxera diverged over 160 MYA (Ren et al. 2013). These two lineages share features reflecting their shared ancestry, most notably cyclical parthenogenesis and XO sex determination (Morgan 1909; Forneck and Huber 2009). There are also several key differences. One is that, during asexual reproduction, aphids are viviparous,
while phylloxera produce eggs. A second difference is that aphids contain intracellular bacterial endosymbionts that provide amino acids (Shigenobu et al. 2000; Chong et al. 2019) and support these symbionts with genes acquired by horizontal gene transfer (HGT) from bacteria (Nakabachi et al. 2014; Smith et al. 2022), while phylloxera lack endosymbionts and the corresponding HGT genes (Rispe et al. 2020).

The grape phylloxera, Daktulosphaira vitifoliae, has historically had a global impact on the grape and wine industries. A recent study characterized its genome and elucidated its population and global invasion history (Rispe et al. 2020). However, this assembly is highly fragmented, preventing chromosome-level comparisons between aphids and phylloxera.

In this study, we used Dovetail Omni-C technologies to assemble a chromosome-level genome for the grape phylloxera. Using analyses of genomes of 30 other hemipteran species with standard sexual life cycles and XO/XY sex determination, we tested whether the distinctive features of aphid genome evolution originated with the origin of cyclical parthenogenesis, and thus extend to phylloxera but not to other hemipterans. In addition, we used the new genome assembly to examine the extent and pattern of gene duplication in phylloxera as compared to aphids and to identify HGT events that have contributed to phylloxera evolution.

Results:

Genome assembly and annotation of grape phylloxera

The assembled genome was produced using data from a proximity ligation protocol (Dovetail Omni-C) incorporating 11.6 Gb of new Illumina reads and the published draft genome assembly (Rispe et al. 2020). Our genome has a total length of 282.86 Mb and an N50 of 45.89 Mb (supplementary table S1). Five scaffolds are larger than 30 Mb, confirming a haploid chromosome count of n = 5 (supplementary fig. S1), as previously documented (Forneck et al. 1999). The total length of the chromosome-level scaffolds is 259.64 Mb, which is 91.8% of the total length of the assembly.

We used BUSCO (Benchmarking Universal Single-Copy Orthologs) to evaluate the completeness of our genome assembly. Querying the single copy orthologs of Hemiptera resulted in a BUSCO score for the genome assembly of 98.3% complete (96.7% single, 1.6% duplicated, 0.4% fragmented, and 1.3% missing). The BUSCO score for the five chromosome-
level scaffolds alone was 98.1% complete (97.1% single, 1.0% duplicated, 0.4% fragmented, and 1.5% missing).

The NCBI RefSeq annotation pipeline was used to annotate the genome (O’Leary et al. 2016). We used WindowMasker (Morgulis et al. 2006) to mask repetitive elements, which made up 53.26% of the genome. We then aligned 18 phylloxera transcriptomes containing 1,698,647,388 reads onto the repeat-masked genome. We predicted a total number of 17,104 annotated genes and pseudogenes, with 14,650 protein-coding genes. Overall, 15,582 predicted genes were annotated on the five chromosome-level scaffolds, with 2,670 on the X chromosome and 12,912 genes on the four autosomes (Table 1).

**Assignment of the X chromosome**

To identify the X chromosome in the phylloxera genome, we first looked at genome synteny between the grape phylloxera and the pea aphid (*Acyrthosiphon pisum*) (supplementary fig. S2). The second-largest chromosome (47.28 Mb) showed extensive gene synteny with the pea aphid X chromosome. To confirm the identity of the X chromosome in phylloxera, we obtained Illumina reads from sexual males and from females and mapped them to our five chromosome-level scaffolds. The second-largest scaffold had about half of the normalized sequencing read depth ratio for males when compared to other chromosomes (supplementary fig. S3), confirming that it is the X chromosome. We named the four autosomes as chromosomes 1-4, ordered from longest to shortest (Table 1).

**Synteny evolution of grape phylloxera and other hemipteran insects**

Synteny between species was generated by MCScanX (Wang et al. 2012) and was visualized with SynVisio (https://github.com/kirandini/synvisio). Based on comparisons of assemblies for different species, we observed numerous rearrangements and shuffling of syntenic regions among aphid autosomes. In contrast, gene content and synteny of the X chromosome were highly conserved between phylloxera and aphids (Fig. 1). Most Aphidomorpha, including phylloxera, possess a single conserved X chromosome. In the *Hormaphis* (Hormaphidinae) and *Eriosoma* (Eriosomatinae) (Fig. 1), the X is split into two chromosomes but these retain conserved syntenic regions corresponding to the X.

To study whether elevated autosomal rearrangements are shared by Aphidomorpha and are associated with the origin of cyclical parthenogenesis, we tested if this unusual pattern of chromosome evolution extended to phylloxera. We found a high level of autosomal
rearrangements when comparing aphids to phylloxera. However, this elevated rate of autosomal rearrangements could be unique to the aphid branch. To determine whether the difference between autosomes and X chromosome extends to phylloxera, we separately compared aphids and phylloxera with *Adelges cooleyi*, a member of the family Adelgidae, a third related lineage undergoing cyclical parthenogenesis. The *A. cooleyi* genome was recently sequenced, along with phylogenetic analyses strongly supporting adelgids and phylloxera as sister groups within Aphidomorpha (Dial et al. 2023). Thus, if parthenogenesis is responsible for autosomal rearrangements, the elevated autosomal rearrangements should be present in both comparisons. However, the elevation was only found when comparing aphids to the other two lineages. In contrast, the comparison of phylloxera and adelgids showed strong synteny conservation for autosomes, suggesting that elevated autosomal rearrangements are restricted to Aphididae (Fig. 1, supplementary fig. S2).

To further study whether Aphididae uniquely display conservation of gene content of the X chromosome contrasting with a higher rate of interautosomal translocations, we extended our analyses to the other two major hemipteran clades, Heteroptera and Auchenorrhyncha, for which sufficient chromosome-level assemblies were available (supplementary table S2). We found no evident differences in the frequencies of interchromosomal rearrangements between autosomes and X chromosomes in these two clades (Fig. 1).

To quantify the difference in chromosomal evolution between Aphididae and other hemipteran lineages, we selected seven independent pairwise comparisons along different branches of the hemipteran phylogeny (supplementary fig. S4). For each pairwise comparison, we estimated the number of interchromosomal rearrangement events per chromosome using pairwise genome synteny analysis. We then calculated approximate rates of interchromosomal rearrangement per chromosome by calibrating with estimated divergence dates of the included species (supplementary table S3). These events include translocations and fusions/fissions; however, most of them are translocations. Because the signature of small chromosomal rearrangements can be ambiguous and difficult to infer, we focused on major rearrangements between species pairs with a minimum 50 MY of divergence (supplementary table S3, S4). We classified chromosomes into four categories: Aphididae X chromosomes, Aphididae autosomes, other hemipteran X chromosomes, and other hemipteran autosomes. Aphididae autosomes have a mean of 0.027 major translocation events per MY. Averages for the other three categories range...
from 0.015 to 0.025 events per MY. Overall, Aphididae autosomes have a significantly higher rate of major rearrangements compared to Aphididae X chromosomes (Mann–Whitney $U = 100$, $p < 0.0006$), other hemipteran autosomes (Mann–Whitney $U = 1603$, $p < 10^{-15}$), and X chromosomes (Mann–Whitney $U = 182$, $p < 10^{-4}$). However, rates of rearrangement of Aphididae X are not significantly different from the other hemipteran autosomes (Mann–Whitney $U = 340$, $p = 0.22$) and X chromosomes (Mann–Whitney $U = 36$, $p = 0.74$) (Fig. 2, supplementary table S5).

**Sequence evolution of aphids, phylloxera, and adelgid**

Compared to genes on autosomes, X-linked genes in aphids show a faster rate of gene sequence evolution, particularly at nonsynonymous sites (Jaquiéry et al. 2012, 2013; Purandare et al. 2014; Jaquiéry et al. 2018; Li et al. 2020). To test if this faster-X pattern is shared by phylloxera, we estimated sequence divergence between phylloxera and pea aphid and between phylloxera and adelgid for gene pairs or orthologs within syntenic blocks (supplementary table S6). Between phylloxera and pea aphid, we found that the mean divergence at nonsynonymous sites is higher for X-linked genes (mean $d_{NX} = 0.30$) than for genes on autosomes (mean $d_{NA} = 0.23$, Mann–Whitney $U = 296745.5$, $p < 10^{-5}$). The divergence at synonymous sites is similar for genes on the X chromosome and on autosomes (mean $d_{SX} = 1.20$, mean $d_{SA} = 1.25$, Mann–Whitney $U = 393876.5$, $p < 0.04$). The $d_N/d_S$ ratio is higher for genes on the X compared to genes on autosomes (mean $d_{NX}/d_{SX} = 0.30$, mean $d_{NA}/d_{SA} = 0.21$, Mann–Whitney $U = 297275.5$, $p < 10^{-5}$). Similarly, between phylloxera and adelgid, the mean divergence at nonsynonymous sites and $d_N/d_S$ ratio are higher for X-linked genes (mean $d_{NX} = 0.20$, mean $d_{NX}/d_{SX} = 0.17$) than for genes on autosomes (mean $d_{NA} = 0.15$, Mann–Whitney $U = 657048$, $p < 10^{-10}$, mean $d_{NA}/d_{SA} = 0.14$, Mann–Whitney $U = 685454$, $p < 10^{-7}$) (supplementary table S6).

**Gene duplication in aphids and phylloxera**

Based on recent genomic studies, aphids genomes show high but variable rates of gene duplication (The International Aphid Genomics Consortium 2010; Julca et al. 2020; Mathers et al. 2017; Li et al. 2019; Fernández et al. 2020) and ancient large-scale gene duplications (Julca et al. 2020). This variation has resulted in wide differences in the total number of protein-coding genes, ranging from 14,089 to 31,435 genes (supplementary table S7). In the grape phylloxera genome, we identified 9,137 genes with duplicates, around 62% of total genes (supplementary...
fig. S5, supplementary table S7), a value near the lower end of the range for Aphidomorpha overall (54-81% of total genes). Relatively low numbers of gene duplications are also observed in Hormaphidinae and Schlechtendalia compared to many Aphidinae. By using synteny and gene locations on chromosome-level assemblies, we further classified gene duplications into four modes: tandem (immediately adjacent), proximal (within 20 flanking genes on a chromosome), dispersed (not within 20 flanking genes), and segmental duplications (anchored within colinear gene blocks) (supplementary fig. S5). We found little evidence of segmental or other large-scale gene duplications. The majority of gene duplications in aphids and phylloxera are dispersed. However, this may reflect the long evolutionary time periods involved, during which larger duplication events are broken up by chromosome rearrangements in autosomes.

Discussion: Life cycles that incorporate unusual reproductive modes present an opportunity to understand the forces acting on genome evolution, including those acting on sex chromosomes. The recent availability of many more chromosome-level genome assemblies can illuminate how these forces may vary due to life cycle variation. In this study, we used a new chromosome-level assembly for the grape phylloxera to determine whether the unusual life cycle shared by aphids and phylloxera has affected their genome evolution similarly. Our new assembly is significantly improved compared to the previous phylloxera genome assembly (Rispe et al. 2020). Our haploid chromosome count agrees with previous karyotyping studies (Maillet 1957; Forneck et al. 1999). Overall, the agreement between our assembly and flow cytometry estimates of genome size and karyotyping supports the high quality and accuracy of this genome assembly.

Elevated rates of interautosomal translocations in Aphididae Our findings are consistent with previous results in aphids, showing long-term conservation of gene content despite elevated gene sequence evolution for the X chromosome, contrasting with frequent interautosomal translocations (Li et al. 2019; Mathers et al. 2021). One hypothesis for the high rate of interautosomal translocations is their accumulation during aphids’ asexual generations. Other hemipterans have fewer annual generations, and all generations are sexual, likely constraining opportunities for translocations. Substantial evidence supports high rates of translocations during asexual aphid generations. Karyotypes in aphids that are exclusively
asexual evolve rapidly (Blackman 1980), and translocations and fusions occur during the short-
term evolution of asexual pest aphid species (Monti et al. 2012b) and in laboratory cultures
(Spence and Blackman 2000). In predominantly asexual aphids, exemplified by species of the
tribe Tramini, karyotypes are highly variable (Blackman et al. 2000; Normark 1999). However,
by comparing synteny between aphids, phylloxera, and adelgids, we found these features of
chromosome evolution have uniquely characterized Aphididae and not shared with Phylloxeridae
and Adelgidae.
Our findings reject the hypothesis that cyclical parthenogenesis itself can explain the pattern of
chromosome evolution in aphids. Cyclical parthenogenesis originated in the common ancestor of
Aphidomorpha, as supported by numerous observations, including cytological observations of
similar chromosome behavior during the derived meiosis (Davis 2012). Yet, our results show
that the elevated rates of autosomal rearrangements in Aphididae most likely evolved after
Aphididae diverged from phylloxera and adelgids. A potential explanation for this elevation is
modifications in the number of asexual generations interspersed between sexual generations.
This varies across aphid species, but the numbers for phylloxera and adelgids fall within the
range observed in aphids (Davis 2012). Some aphid lineages lose the sexual phase entirely, but
such losses are absent or very recent in the lineages we examined, and thus would not impact
large-scale patterns of chromosome evolution (Blackman et al. 2000; Normark 1999). Thus the
elevated rate of autosomal rearrangements in aphids is not readily explained by numbers of
asexual generations.
An alternative explanation for the high rate of autosomal rearrangements is the presence of
holocentric chromosomes in aphids (Schubert and Lysak 2011; Mandrioli and Manicardi 2020).
In plants, holocentric chromosomes can facilitate rearrangements, yielding variable karyotypes
(Hofstatter et al. 2022), and elevated chromosomal evolution has been observed in some other
groups with holocentric chromosomes (Höök et al. 2023; Ruckman et al. 2020; Fradin et al.
2017). However, in aphids, a high rate of rearrangements is likely not driven by holocentricity.
All hemipteran insects have holocentric chromosomes (Tree of Sex Consortium 2014), yet they
have much lower rates of translocations than aphids. Indeed, most hemipterans show lower rates
of chromosome evolution than do most other insect orders (Ruckman et al. 2020; Kuznetsova et
al. 2021).
Thus the explanation for the elevated rate in aphids remains unclear. Potentially it reflects some altered molecular mechanisms involved in chromosome replication during the asexual and/or sexual generations of aphids, but not of adelgids and phylloxera.

**Conservation of X chromosome gene content in Aphidomorpha**

Most likely, aphids undergo relatively frequent rearrangements in all chromosomes, but the sexual generation imposes a stronger selective sieve on the X than on autosomes. For example, in asexual populations of *Myzus persicae*, chromosomal translocations involve both autosomes and the X chromosome (Monti et al. 2012a). Furthermore, the X chromosome does undergo intrachromosomal rearrangements (Fig. 1), indicating that it is not immune to recombination events that affect large-scale gene synteny. The unusually conserved feature of the X chromosome is its gene content, not its gene order.

Translocations involving autosomes or the X chromosome will often have little fitness consequence during asexual female generations, which are freed from constraints of homolog pairing during meiosis. However, translocations affecting gene content of the X are expected to be highly negatively selected during the sexual generation, as XO males would lack essential genes or suffer deleterious changes in gene dosage, as proposed previously (Li et al. 2019; Mathers et al. 2021). An illustration of this kind of selection can be seen for the ribosomal RNA (rRNA) genes. In sharp contrast to most essential genes, the rRNA genes are located on the X chromosome in aphids (Manicardi et al. 2015). The rRNA locus is sometimes eliminated from one homolog through mitotic recombination during asexual generations in laboratory culture (Blackman and Spence 1996). Clearly, this deletion would be lethal in males. Thus, interchromosomal rearrangements between X chromosomes and autosomes likely have large fitness consequences in males. Such changes are likely eliminated by purifying selection during the sexual generation, thus explaining conservation of gene content of the aphid X.

One distinction between the X chromosome of aphids versus that of adelgids and phylloxerids is that, the rRNA genes appear not to be confined to the X chromosome in phylloxerids or adelgids, as they are in aphids. Of the 35 rRNA operons identified in the grape phylloxera, only two were on the X chromosome, 31 were on autosomes, and 2 were on short contigs not placed on chromosomes.

A recent study proposed that gene dosage for X-autosome fusions under somatic X chromosome loss might explain the conservation of aphid X chromosomes (Roy 2021). However, this
hypothesis cannot explain chromosome evolution in aphids generally, as it assumes intensive inbreeding among aphids. Most aphids outbreed, and outbreeding is enforced by life cycles in many aphid groups in which males and females fly separately to alternative host plants (Moran 1988; Hardy et al. 2015).

**Sequence evolution in aphids, phylloxera, and adelgid**

In contrast to the conservation of gene content of the X chromosome, gene sequence evolution is faster for X-linked genes, as found in our study and in previous investigations (Jaquiéry et al. 2012, 2013; Purandare et al. 2014; Jaquiéry et al. 2018; Li et al. 2020). This elevated sequence evolution has been attributed to relaxed purifying selection, consistent with the smaller effective population size of the X, overall lower expression of X-linked genes, and a tendency of these genes to be expressed more in males than in females. Selection on male-biased genes is weaker than that on female-biased genes, as males are infrequent during the life cycle. The aphid X chromosome is highly methylated and transcribed at a lower rate than autosomes (Mathers et al. 2019), a finding that is consistent with relaxed purifying selection on X-linked genes. Given that phylloxera share XO sex determination and similar X chromosome gene content, the faster-X pattern was predicted to extend to phylloxera. Our findings confirmed that X-linked genes in phylloxera have higher rates of amino acid replacement, despite similar rates of silent substitution, as compared to genes on autosomes.

**Conclusion**

An implication from our study is that the accelerated rate of autosomal rearrangements in aphids cannot be explained by their cyclically parthenogenetic life cycle. To explain the X chromosome conservation in aphids and phylloxera, we hypothesize that the absence of this elevated translocation rate on the X chromosome reflects purifying selection on males during the annual sexual generation. One prediction is that the aphid X will exhibit more translocations to and from autosomes in lineages of Aphidomorpha that have eliminated the sexual phase. Our study raises more questions: What mechanism drives elevated rates of interautosomal rearrangements in Aphidae? Do higher rates of interautosomal rearrangements in aphids lead to higher speciation rates compared to phylloxera and adelgids?
Materials and Methods
Sample preparation for genome sequencing
Multiple *D. vitifoliae* Pcf7 clone females were collected from two individual plants (two different rootstocks of grapes) from the Bordeaux collection. For a high-quality genome assembly, a total of 200 mg of fresh material from 1,500 leaf-galling female phylloxera was frozen and shipped to Dovetail Genomics (Santa Cruz, CA, USA). All individuals were used for DNA extraction and Hi-C library preparation. The library was sequenced on an Illumina HiSeq X platform to produce approximately 30x sequence coverage.

Assembly of the grape phylloxera genome
To assemble the *D. vitifoliae* genome, published de novo draft genome assembly (Rispe et al. 2020) and 11.6 Gb Dovetail proximity ligation reads were used as input data for HiRise genome scaffolding. The detailed assembly method for the draft genome can be found in Rispe et al. 2020. HiRise assembler version v2.1.6-072ca03871cc was used for scaffolding with default parameters (Putnam et al. 2016). Proximity ligation library sequences were aligned to the draft input assembly using bwa v0.7.17-r1188 (Li and Durbin 2009) with defaults. The separations of proximity ligation library read pairs mapped within draft genome scaffolds were analyzed by HiRise to produce a likelihood model for the genomic distance between read pairs, and the model was used to identify and break putative misjoins, score prospective joins, and make joins above a threshold (https://omni-c.readthedocs.io/en/latest/). Overall, 1,850 joins were and 30 breaks were made to the input assembly. To evaluate the completeness of our genome assembly, BUSCO v4.1.4 (Seppey et al. 2019) was used on the chromosome-level assembly with the single-copy orthologous gene set for Hemiptera from OrthoDB version 9 (Zdobnov et al. 2017).

Genome annotation
The NCBI Eukaryotic Genome Annotation Pipeline was used for genome annotation (O’Leary et al. 2016). Repeat families found in the genome assemblies of *Daktulosphaira vitifoliae* were identified and masked using WindowMasker (Morgulis et al. 2006). Over 20,000 transcripts of phylloxera and high-quality proteins of phylloxera and other closely related insects were retrieved from Entrez, aligned to the genome by Splign or ProSplign (Kapustin et al. 2008) and Minimap2 v2.17-r941 (Li 2018). Additionally, 1,698,647,388 reads from 18 phylloxera RNA-Seq datasets were also aligned to the repeat-masked genome. Protein, transcript, and RNA-Seq read alignments were passed to Gnomon for gene prediction. The final annotation set was
assigned to models based on known and curated RefSeq and models based on Gnomon predictions. The overall quality of the annotations was assessed using BUSCO v4.1.4 (Seppey et al. 2019). The detailed annotation pipeline can be found at https://www.ncbi.nlm.nih.gov/genome/annotation_euk/process/.

**Assignment of the X chromosome and autosomes**

The X chromosome was assigned following the method previously used in the pea aphid and psyllid genomes (Li et al. 2019, 2020). We mapped whole genome sequencing reads from male and asexual female individuals back to our chromosome-level genome assembly. The male sequencing reads were generated through this study (BioProject: PRJNA929591, Accession: SRR23285932). The asexual female sequencing reads were obtained from the previous phylloxera genome project (Rispe et al. 2020) through GenBank (BioProject: PRJNA588186, Accession: SRR10412121). The sequencing reads were cleaned with Trimmomatic version 0.38 (Bolger et al. 2014). The clean reads were mapped to the chromosome-level assembly using Bowtie2 version 2.3.4.3 (Langmead and Salzberg 2012) with default parameters. The resulting SAM files were converted to BAM files, sorted, and indexed using SAMtools version 1.9 (Danecek et al. 2021). We estimated the sequencing depth based on 10-kb sliding windows with 2-kb steps, and the sequencing depth of each window was estimated using Mosdepth version 0.2.3 (Pedersen and Quinlan 2018). We normalized the overall sequencing depths among male individuals and female individuals based on methods used in Li et al. 2020. The overall sequencing depth distribution was plotted using a violin plot in ggplot2 version 3.2.1 (Wickham 2016). The X chromosome was assigned to the chromosome that had about half the ratio of sequencing depth between males and females compared to the others.

**Synteny analyses of Hemipteran genomes**

We used MCScanX (Wang et al. 2012) to evaluate whole genome synteny between hemipteran species. Eight chromosome-level aphid genomes were downloaded from Aphidinae comparative genomics resource on Zenodo (https://zenodo.org/record/5908005#.Y255M3bM15Y); other chromosome-level hemipteran genome assemblies and annotations were obtained from datasets published before August 2022 (supplementary table S2). All vs. all blastp searches with genome protein sequences were performed with an e-value of 1e-10. MCScanX was used to generate synteny data between species with defaults. SynVisio (https://github.com/kiranbandi/synvisio) was used to display syntenies. As genomes of multiple species are available for some aphid
clades, we selected *Acyrthosiphon pisum* for Macrosiphini, *Rhopalosiphum maidis* for Aphidini, and *Eriosoma lanigerum* for Eriosomatinae for comparisons (Figure 1).

The level of interchromosomal rearrangements per chromosome was quantified with pairwise syntenic analyses of seven independent comparisons along different branches of the hemipteran phylogeny (supplementary fig. S4). A total of 4 aphid and 10 other hemipteran genome assemblies were used. All selected pairs had a minimum of 50 million years of divergence (supplementary table S3), and pairwise syntenic analyses were performed with MCScanX as described above. Pairwise synteny for mirids (*Apolygus lucorum* vs. *Cyrtorhinus lividipennis*) and phylloxera/adelgid (*D. vitifoliae* vs. *Adelges cooleyi*) were generated, but these data were excluded from this analysis because there was no published divergence time for the mirid species pair and because the adelgid genome was not a chromosome-level assembly. All chromosomes were classified into four categories: aphid autosomes (*n*=17), aphid X chromosomes (*n*=6), other hemipteran autosomes (*n*=96), and other hemipteran X chromosomes (*n*=11) (supplementary table S4). Identification of X chromosomes was based on genomic confirmation from previous studies or sequence homology with a confirmed X chromosome.

We further estimated the rate of interchromosomal rearrangement. For each chromosome, we counted the number of chromosomes of the other species which share syntenic blocks (at least five collinear genes) in each pairwise synteny analysis. To eliminate small chromosomal rearrangements, we excluded chromosomes that shared three or fewer syntenic blocks from the count. Since a count of one (a 1:1 syntenic relationship between two chromosomes) indicates no interchromosomal rearrangements, the count was subtracted by one to generate the number of interchromosomal rearrangement events. To estimate the approximate rate of interchromosomal rearrangement events without reconstructing ancestral chromosomes, we then divided the number of interchromosomal rearrangement events by the estimated divergence dates of the species pair. Finally, we tested if the rates of chromosomal rearrangement are statistically different between the four categories using the Mann-Whitney U test with a Bonferroni correction (supplementary table S5).

**Sequence evolution of aphids, phylloxera, and adelgid**

We used syntenic gene pairs between phylloxera and pea aphid and between phylloxera and adelgid from the synteny analysis described in the previous section. The Perl script *add ka and ks to collinearity.pl* from MCScanX (Wang et al. 2012) was used to calculate synonymous (dS) and...
nonsynonymous (d\(N\)) substitution rates for each syntenic gene pair between the two species. Gene pairs with d\(N\) > 1 or d\(S\) > 2 were removed from the analyses to exclude low accuracy estimates of divergence. In phylloxera and pea aphid, we found 226 gene pairs on the X chromosome, and 3,220 gene pairs on autosomes. In phylloxera and adelgid, we found 436 gene pairs on the X chromosome, and 3,758 gene pairs on autosomes (supplementary table S6). We also calculated the d\(N/dS\) ratio for all gene pairs. The mean d\(S\), d\(N\), and d\(N/dS\) ratios were compared between the X chromosome and autosomes with a Mann-Whitney U test.

**Gene duplication in aphids**

Classification of the types of gene duplication was accomplished with the duplicate gene classifier program in MCScanX (Wang et al. 2012). The protein sequences for each species was used as the query and the database in a blastp search with an e-value of 1e-10. For each genome, the blastp output and a protein annotation file were used as input files for the duplicate gene classifier program. All default parameters were used. The duplicate gene classifier identified genes as duplications if they hit any other proteins in the blastp search or singletons if they did not. The duplications were further classified into (1) tandem, if they differed by one gene rank; (2) proximal, if they differed by more than 1 and less than 20 gene ranks; (3) dispersed, if they differed by greater than 20 gene ranks; or (4) segmental duplication, if they were anchored within collinear blocks of genes according to MCScanX (Wang et al. 2012). The percentages for each mode of gene duplication were calculated as the number of duplicate genes in each mode out of the total number of gene duplications in each species (supplementary table S7).

**Data Availability**

*Daktulosphaira vitifoliae* chromosome-level genome assembly generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA870220. The proximity ligation reads data generated in this study have been submitted to the NCBI BioProject database under accession number PRJNA588186. The sexual male genomic DNA sequencing of the Arizona population generated in this study have been submitted to the NCBI BioProject database under accession number PRJNA929591. The RNAseq sequencing data generated in this study have been submitted to the NCBI BioProject database under accession number PRJNA929654.
Acknowledgments

This work was supported by the NIH award R35GM131738 and NSF award DEB1551092 to NAM. This project has also received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 764840. ZL is supported by the National Science Foundation Postdoctoral Research Fellowships in Biology Program (NSF 2109306). We acknowledge Fabrice Legeai and Denis Tagu (INRAE, UMR IGEPP, France), Claude Rispe (INRAE, UMR BIOEPAR, France), and François Delmotte (INRAE, UMR SAVE, France) for discussion and assistance in this study. We acknowledge Dustin Dial, Kathryn Weglarz, Bryan Brunet, Nathan Havill, Carol von Dohlen, Gaelen Burke for sharing the A. cooleyi genome. We thank Eli Powell and Tyler de Jong for assisting with DNA and RNA extractions. We thank Jo-anne Holley and PJ Lariviere for assisting with insect collecting. We also thank Qiuyu Jiang, Michael McKibben, Geoff Finch, and Chenlu Di for collecting the Arizona population of the grape phylloxera.

References

Blackman RL, Spence JM. 1996. Ribosomal DNA is frequently concentrated on only one X chromosome in permanently apomictic aphids, but this does not inhibit male determination. Chromosome Res. 4:314–320.


Mathers TC, Mugford ST, Percival-Alwyn L, Chen Y, Kaithakottil G, Swarbreck D, Hogenhout


Zdobnov EM, Tegenfeldt F, Kuznetsov D, Waterhouse RM, Simão FA, Ioannidis P, Seppey M,

Table 1. Summary of the chromosome-level assembly of the grape phylloxera genome.

<table>
<thead>
<tr>
<th>Category</th>
<th>Length (Mb)</th>
<th>No. of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome 1</td>
<td>89.58</td>
<td>5,226</td>
</tr>
<tr>
<td>Chromosome 2</td>
<td>45.89</td>
<td>2,936</td>
</tr>
<tr>
<td>Chromosome 3</td>
<td>43.29</td>
<td>2,712</td>
</tr>
<tr>
<td>Chromosome 4</td>
<td>33.59</td>
<td>2,038</td>
</tr>
<tr>
<td>Chromosome X</td>
<td>47.28</td>
<td>2,670</td>
</tr>
<tr>
<td>All Chromosomes</td>
<td>259.64</td>
<td>15,582</td>
</tr>
<tr>
<td>Total Assembly</td>
<td>282.86</td>
<td>16,912</td>
</tr>
</tbody>
</table>
Fig. 1. Pairwise synteny (gene order) of Heteroptera (true bugs), and Auchenorrhyncha (planthoppers, leafhoppers, and relatives), and Sternorrhyncha (aphids and relatives). Bars represent chromosomes. The X chromosome(s) for each species is colored gray and located on the far right. Autosomes are in other colors. The length of the bars is proportional to the length of the chromosome-level scaffolds in the assemblies. The cladogram is based on Johnson et al. 2018. The clade undergoing cyclical parthenogenesis is colored blue.
Fig. 2. The rates of interchromosomal rearrangement of Aphididae and other hemipteran chromosomes (AA, Aphididae autosomes; AX, Aphididae X chromosomes; HA, other hemipteran autosomes; HX, other hemipteran X chromosomes). The rate of interchromosomal rearrangement within each category of chromosome is shown in the violin plots. The black dot indicates the mean, the thick black bars represent the standard deviation of the data, and the color shading represents the density of data points. The comparisons of the Aphididae autosomes to other categories by the two-sample Mann-Whitney U test are provided. ‘***’ represents Bonferroni-corrected $p < 0.002$. 

- AA (Aphididae autosomes)
- AX (Aphididae X chromosomes)
- HA (other hemipteran autosomes)
- HX (other hemipteran X chromosomes)