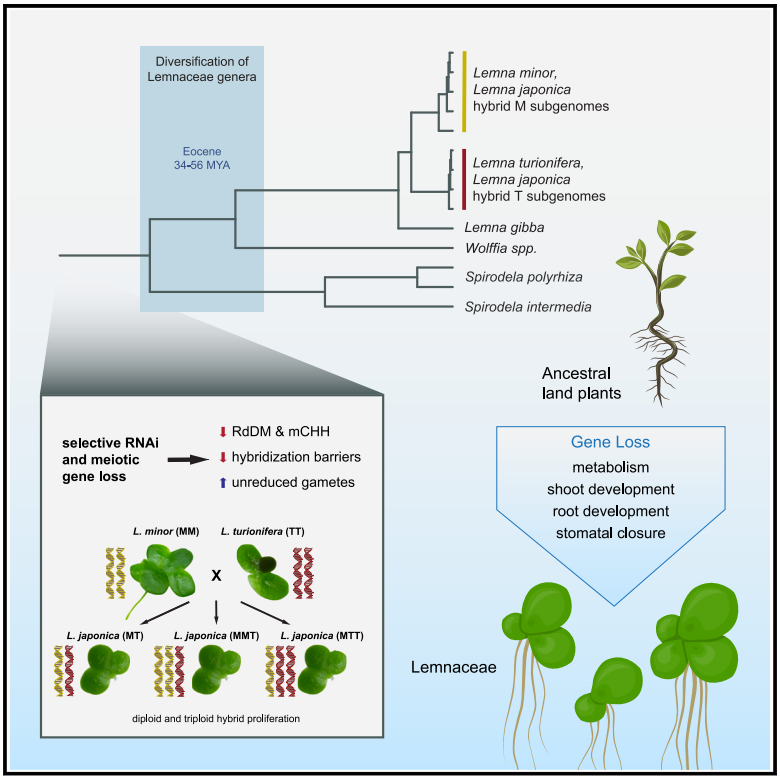
Article

Current Biology

Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction

Graphical abstract



Authors

Evan Ernst, Bradley Abramson, Kenneth Acosta, ..., Eric Lam, Todd P. Michael, Robert A. Martienssen

Correspondence

eric.lam@rutgers.edu (E.L.), tmichael@salk.edu (T.P.M.), martiens@cshl.edu (R.A.M.)

In brief

Duckweeds are the fastest-growing angiosperms. By sequencing and comparing their genomes, Ernst et al. pinpoint genes that could explain adaptations to their ecological niche, providing targets for enhancing their use as energy, protein, and carbon sequestration crops.

Highlights

- Candidate genes found for reduced anatomy, stomata development, and carbon storage
- Loss of RNAi genes involved in reproductive isolation and haploid gamete formation
- Epigenetically defined centromeres retain position and parental sequence in hybrids
- Diversification of duckweed genera coincided with the "Azolla event"

Ernst et al., 2025, Current Biology *35*, 1–20
April 21, 2025
https://doi.org/10.1016/j.cub.2025.03.013

CellPress

Current Biology

CellPress
OPEN ACCESS

Article

Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction

Evan Ernst,[1] Bradley Abramson,[2] Kenneth Acosta,[3] Phuong T.N. Hoang,[4,5] Cristian Mateo-Elizalde,[1] Veit Schubert,[4] Buntora Pasaribu,[3,6] Patrice S. Albert,[7] Nolan Hartwick,[2] Kelly Colt,[2] Anthony Aylward,[2] Umamaheswari Ramu,[1] James A. Birchler,[7] Ingo Schubert,[4] Eric Lam,[3,*] Todd P. Michael,[2,*] and Robert A. Martienssen[1,8,*]

[1]Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, 1 Bungtown Rd, Cold Spring Harbor, NY 11724, USA
[2]Plant Molecular and Cellular Biology Laboratory, the Salk Institute for Biological Studies, 10010 N Torrey Pines Rd, La Jolla, CA 92037, USA
[3]Department of Plant Biology, Rutgers, The State University of New Jersey, 59 Dudley Road, New Brunswick, NJ 08901, USA
[4]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Corrensstraße 3, D-06466 Stadt Seeland, Germany
[5]Biology Faculty, Dalat University, 1 Phu Dong Thien Vuong, Dalat City 670000, Vietnam
[6]Department of Marine Sciences, Faculty of Fisheries and Marine Sciences, Universitas Padjadjaran, Bandung Sumedang Highway KM 21, Jatinangor 40600, Indonesia
[7]Biological Sciences, University of Missouri, Columbia, 105 Tucker Hall, Columbia, MO 65211, USA
[8]Lead contact
*Correspondence: eric.lam@rutgers.edu (E.L.), tmichael@salk.edu (T.P.M.), martiens@cshl.edu (R.A.M.)
https://doi.org/10.1016/j.cub.2025.03.013

SUMMARY

The Lemnaceae (duckweeds) are the world's smallest but fastest-growing flowering plants. Prolific clonal propagation facilitates continuous micro-cropping for plant-based protein and starch production and holds tremendous promise for sequestration of atmospheric CO_2. Here, we present chromosomal assemblies, annotations, and phylogenomic analysis of *Lemna* genomes that uncover candidate genes responsible for the unique metabolic and developmental traits of the family, such as anatomical reduction, adaxial stomata, lack of stomatal closure, and carbon sequestration via crystalline calcium oxalate. Lemnaceae have selectively lost genes required for RNA interference, including Argonaute genes required for reproductive isolation (the triploid block) and haploid gamete formation. Triploid hybrids arise commonly among *Lemna*, and we have found mutations in highly conserved meiotic crossover genes that could support polyploid meiosis. Further, mapping centromeres by chromatin immunoprecipitation suggests their epigenetic origin despite divergence of underlying tandem repeats and centromeric retrotransposons. Syntenic comparisons with *Wolffia* and *Spirodela* reveal that diversification of these genera coincided with the "Azolla event" in the mid-Eocene, during which aquatic macrophytes reduced high atmospheric CO_2 levels to those of the current ice age. Facile regeneration of transgenic fronds from tissue culture, aided by reduced epigenetic silencing, makes *Lemna* a powerful biotechnological platform, as exemplified by recent engineering of high-oil *Lemna* that outperforms oil-seed crops.

INTRODUCTION

The Lemnaceae[1] are a family of freshwater aquatic macrophytes commonly known as duckweeds[2] and are sometimes referred to as water lentils and watermeal. Efficient regeneration of transgenic fronds from tissue culture and reduced epigenetic silencing could make *Lemna* a powerful biotechnological platform, and the Lemnaceae offer a unique opportunity to engineer CO_2 capture and sequestration, as well as biofuel production, in the modern age. A recent study showed *L. japonica* plants simultaneously overexpressing *WRINKLED1*, *DIACYLGLYCEROL ACYLTRANSFERASE*, and *OLEOSIN* can accumulate oil at up to 8.7% of dry mass, illustrating the potential for metabolic pathway manipulation in Lemnaceae.[3]

The Lemnaceae reproduce by reiterative vegetative budding from a "pocket" of meristematic stem cells, doubling once per day under optimal conditions. Free-floating clonal reproduction provides the optimal environment for rapid plant growth, and the Lemnaceae have the shortest biomass doubling time of any flowering plant, making them attractive for micro-farming and for ammonium, nitrate, phosphate, and CO_2 remediation. However, they are true flowering plants, and some species and clones in each of the five genera (*Spirodela*, *Landoltia*, *Lemna*, *Wolffia*, and *Wolffiella*) can produce simple flowers and fruits with 1–5 seeds in response to hormones, nutrients, temperature, and daylength (Figure 1A). The clonal growth habit tolerates a high frequency of polyploidy as well as interspecific hybridization,[4–7] suggesting that reproductive isolation barriers in the seed were lost in the absence of obligate sexual reproduction.[5] In particular, *L. turionifera* (T) and *L. minor* (M) form frequent polyploid hybrids, known as *L. japonica*,[7] which have enhanced vigor compared with diploid relatives

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

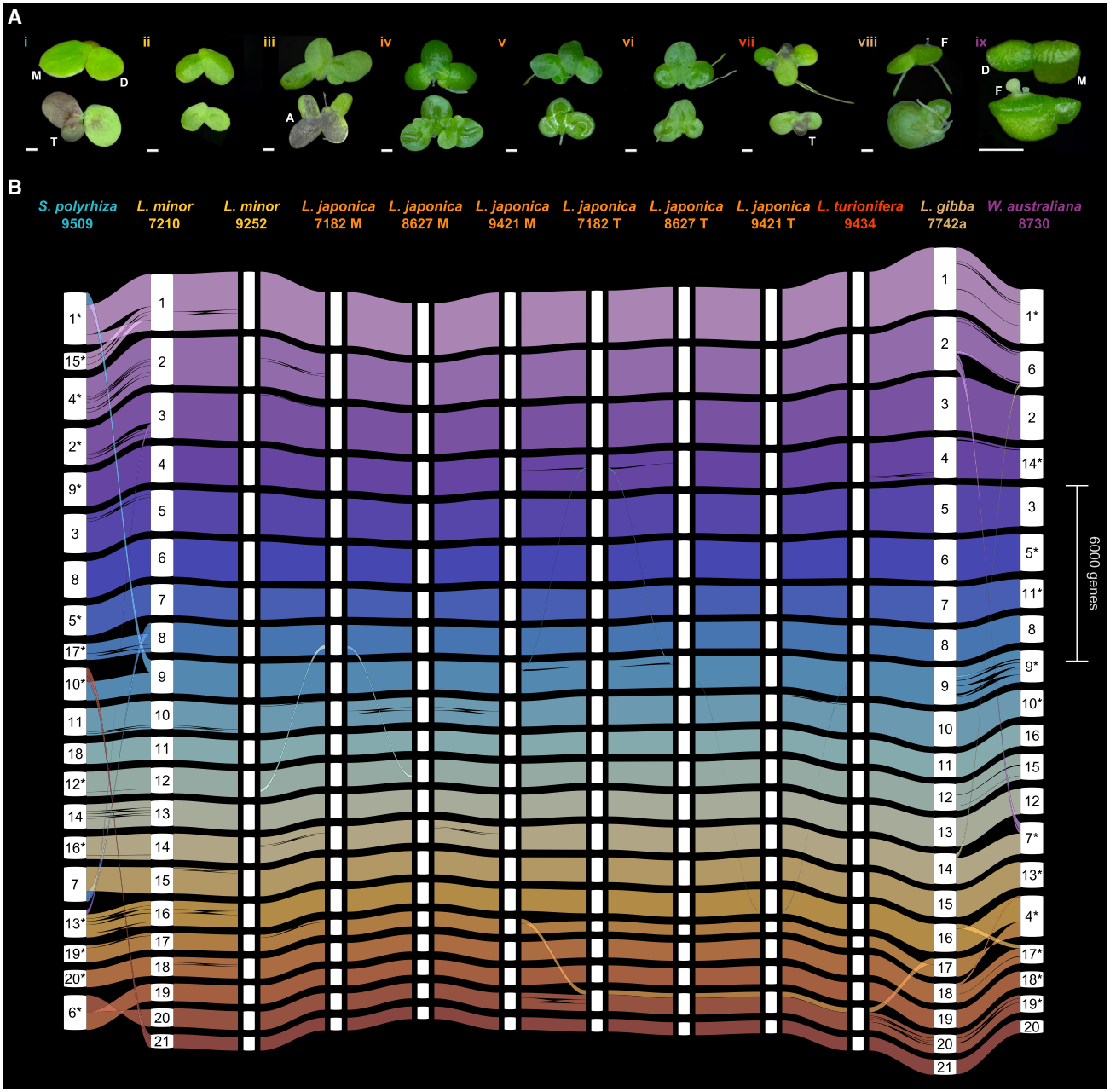


Figure 1. Lemnaceae habit and genomes

(A) Species in this study: (i) *Spirodela polyrhiza* 9509, (ii) *Lemna minor* 7210, (iii) *Lemna minor* 9252, (iv) *Lemna japonica* 7182, (v) *Lemna japonica* 8627, (vi) *Lemna japonica* 9421, (vii) *Lemna turionifera* 9434, (viii) *Lemna gibba* 7742a, and (ix) *Wolffia australiana* 8730. Dark-field microscopy of colonies of mother fronds (M) bearing clonal daughter frond progeny (D) (scale bar, 1 mm). Turions (T) are visible after 40 days of growth in dilute media in *L. turionifera* (vii) colonies but not in those of *L. minor* (ii–iii) or *L. japonica* (iv–vi). *S. polyrhiza* (i) produced turions on dilute media. Starvation elicited a strong anthocyanin (A) response in turion-producing plants and in the fronds of *L. minor* 9252 (iii). Flowers (F) are visible in *L. gibba* (viii) and *W. australiana* (ix) after growth on inductive media.

(B) Gene-level synteny. The genomes and subgenomes of *Lemna* and *Wolffia* species were sequenced using long-read single-molecule sequencing, assembled into 21 (*Lemna*) or 20 (*Wolffia*) pseudomolecules with chromatin conformation capture, and annotated for direct comparison of gene content. *Lemna* chromosomes were numbered by size in *L. minor* 7210 (common duckweed).[4] Ribbons represent blocks of syntenic protein-coding gene loci. * = chromosomes inverted relative to their reference representation to more clearly show syntenic relationships.

See also Figures S1 and S2, Tables S1 and S2, and Data S1.

under certain environmental conditions,[8] perhaps explaining their adventitious selection for biotechnological applications.[9,10] Isolates classified as either *L. minor* or *L. japonica* are difficult to distinguish morphologically or by plastid markers,[11] though genetic barcoding with polymorphic nuclear markers confirms their classification as distinct species.[6]

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

We used single-molecule nanopore sequencing and high-throughput chromatin conformation capture (Hi-C) contact mapping to generate chromosome-resolved genome assemblies of *Lemna* species, including the first assemblies of *L. japonica* interspecific hybrids, revealing that they form with variable parental dosage as diploids and reciprocal triploids. We determined the organization of Lemna centromeres by analyzing retrotransposons and tandem repeats recovered by immunoprecipitation (IP) of the centromeric H3 histone, CENH3. We also determined the patterns of DNA methylation as well as small RNA (sRNA) accumulation, and we suggest that the loss of genes required for RNA-dependent DNA methylation (RdDM) could account for the high frequency of polyploids in Lemnaceae. This is because RdDM is required in flowering plants for the "triploid block": a reproductive barrier in which triploid seeds abort.[12,13]

RESULTS

Chromosome-resolved Lemnaceae genome assemblies

Genome architectures and synteny

We used Oxford Nanopore Technologies (ONT) single-molecule long reads, paired with Hi-C contact mapping or reference-based scaffolding, to generate chromosome-resolved *de novo* genome assemblies for 8 duckweed accessions representing 5 species and protein-coding gene annotations spanning 3 genera (Figures 1 and S1; Table S1). Mean raw ONT read coverage ranged from 38× to 105× (Data S1A), and contig N50s varied from 3.2–13.9 Mbp (Table S1). *Lemna japonica* 8627 (previously classified as *Lemna minor*) was among our initial targets for whole-genome sequencing due to its amenability to genetic transformation and use as a recombinant expression platform.[14,15] Individual ONT reads are long enough to span distant tracts of single-nucleotide and structural variants (SNVs and SVs), enabling the separation of two homologous chromosome sets in draft assemblies of this accession. This prompted us to sequence three additional *L. japonica* accessions and their founder species to better understand hybridization and genome variation within the genus (Figure 2A).

We resolved M (*L. minor*) and T (*L. turionifera*) subgenomes into two haplotype-collapsed sets of 21 chromosomes for Lj7182, Lj8627, and Lj9421, confirming that the *L. japonica* taxon represents distinct interspecific hybrids of *L. minor* and *L. turionifera* (Figures 1A, 1B, and 2B). Genomic read mapping was used to assess the dosage of each parental haplotype between the three accessions, and together with nuclear genome size estimates by flow cytometry, this indicated that hybrids form both as MT diploid (Lj9421) and reciprocal MTT (Lj7182) and MMT (Lj8627) triploids (Figures 2B and S2A). Whole-genome alignment and synteny mapping supported a consistent *L. turionifera* karyotype distinguished from *L. minor* and the more distantly related *L. gibba* by the translocation of 3.5 Mbp of one arm of chromosome 17 (Chr17) to Chr20 (Figure 1B). Further highlighting the difficulty in discriminating *L. minor* from *L. japonica* hybrids, the assembly of accession 9252, originally labeled *L. japonica*, lacked an *L. turionifera* subgenome, consistent with a previous report that it is a heterozygous diploid *L. minor*.[6] By contrast, Lm7210 from South Africa had a rate of heterozygous short variant calls comparable to the single-copy subgenomes of the hybrids, which represents the false discovery rate, indicating that Lm7210 could be a natural doubled haploid (Figure 2C). Although heterozygosity was evident in the other diploid genomes (Figure 2C), it was very low compared with terrestrial plants, as it is in *Spirodela* and other aquatic plants.[16–19] The genome dosage and parental composition of Lj8627 were independently confirmed by genome *in situ* hybridization (GISH) using the *L. minor* and *L. turionifera* genomes as probes (Figure 2D). Structured illumination microscopy (SIM) was used to resolve the chromosome sets from each subgenome, revealing 21 chromosomes from *L. turionifera* and 42 chromosomes from *L. minor*.

rDNA repeats are rearranged in triploid hybrids. In most cases, we were able to determine the chromosomal locations of highly conserved ribosomal DNA (rDNA) repeat arrays, which showed evidence of karyotypic plasticity (Figure S2B). DNA fluorescence *in situ* hybridization (FISH) studies detected just one 45S rDNA locus in the majority of duckweed accessions surveyed, yet some *Wolffia* and *Wolffiella* clones had two loci.[20] In the case of *W. australiana* 8730 (Wa8730), a sole intact locus was assembled on Chr14, which was homologous to *Lemna* Chr4. A remnant locus was also present on Wa8730 Chr4, which was homeologous to Chr16, Chr17, Chr18, and Chr20 in *Lemna spp.*, near the Chr17:Chr20 fusion breakpoint (Figures 1B, S1B, and S2B). In the diploid *L. minor* accessions assembled here, the intact array was located at the end of Chr20, while in the *L. turionifera* subgenomes of the hybrids, the best-conserved array was translocated to Chr17. This was consistent with an rDNA array on the ancestral homeolog of Wa8730 Chr4 migrating to Chr20 in *L. minor* and Chr17 in *L. turionifera* lineages. The translocation of the Chr17 terminus to Chr20 was shared in all *L. turionifera* assemblies, suggesting this may have occurred at the same time (Figures 1B and S1B). However, among the *L. japonica* hybrids, intact and likely active rDNA repeats were assembled at distinct positions. In the MTT hybrid 7182, a conserved array was assembled only at Chr17T, and the remaining loci, including Chr20M, are degraded. In the MMT hybrid 8627, only remnant arrays were present on Chr17T and three other locations, but the highly conserved array sequence was unanchored, as it was in *L. turionifera* Lt9434. Only in the case of the diploid *L. japonica* hybrid 9421, intact rDNA copies were assembled on both subgenomes. In addition, a consistently degraded remnant locus appeared on Chr6 in the diploids Lm9252 and Lt9434 and on both M and T subgenomes of the *L. japonica* hybrids. Thus, active rDNA array degradation occurred in triploid but not diploid hybrids, possibly reflecting dosage of the parental chromosomes on which they reside. Similar rearrangements are frequent in other examples of polyploid hybrids.[21]

Centromere identification and characterization. Centromeres are the sites of kinetochore attachment during cell division, defined epigenetically in plants by incorporation of the histone 3 variant CENH3. Most plants exhibit a regional monocentric organization, with a single attachment site per chromosome, however, metapolycentric organization in legumes and holocentric organization in sedge and rush have also been observed.[22,23] In *Lemna* spp., no centromeric constriction was obvious in metaphase chromosome preparations (Figure 2D; see Hoang et al.[24]), and gene-rich regions were found

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

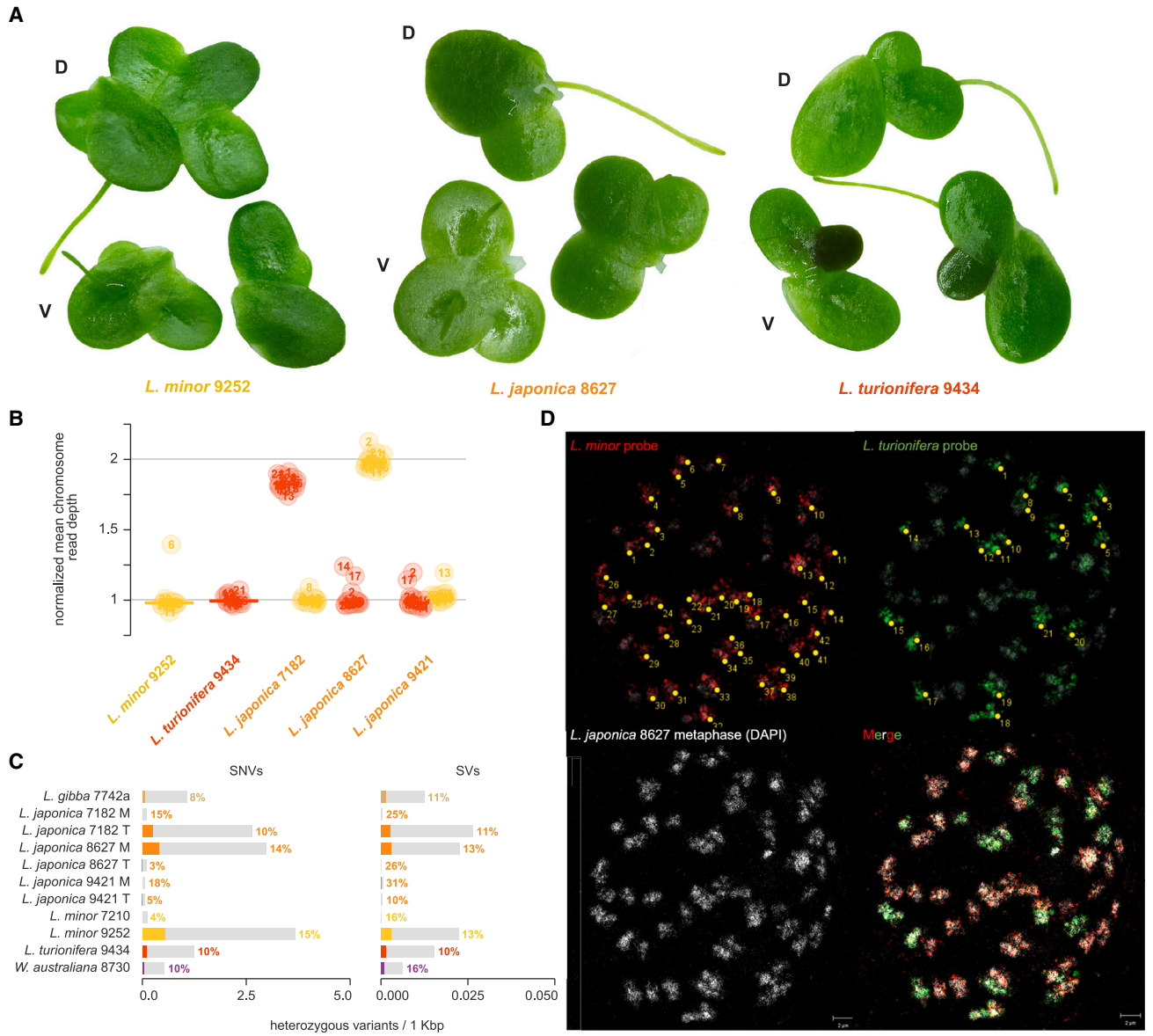


Figure 2. Subgenome dosage and heterozygosity in diploid and triploid *L. japonica*

(A) Dorsal (D) and ventral (V) views of typical *L. minor*, *L. japonica*, and *L. turionifera* frond colonies. Turions are forming in one of the two meristematic recesses of *L. turionifera*.

(B) Read depth per pseudomolecule in *L. japonica* hybrids and representative diploid parent assemblies, normalized to the mode of read depth across each subgenome assembly. Pseudomolecules are numbered as in Figure 1.

(C) Heterozygous variant calls made from long reads mapped back to each haplotype-collapsed assembly. Single nucleotide variants (SNVs) and INDELs <50 bp, and structural variants (SVs) >50 bp counts per kilobase of reference genome sequence are shown. The subgenomes of *L. japonica* hybrids are separately indicated by their *L. minor* (M) or *L. turionifera* (T) karyotypes. Variants overlapping protein-coding gene annotations are highlighted with color, and their proportion is shown to the right of the bars. SNV and SV levels in diploid subgenomes reflect substantial heterozygosity, except for Lm7210, where levels are at the false positive rate observed in monoploid subgenomes 7182 M, 8627 T, 9421 M, and 9421 T. Lm7210 thus appears to be a doubled haploid.

(D) Genome *in situ* hybridization (GISH) of genomic DNA of *L. minor* (in red) and *L. turionifera* (in green) on metaphase chromosomes of *L. japonica* 8627 visualized by 3D-SIM.

See also Figures S2 and S6.

to be dispersed along the length of chromosomes (Figure S3A; see Cao et al.[25]), rather than concentrated toward chromosome arms as in the model dicot *Arabidopsis*.[23] Often, the genomic sequence underlying monocentromeres comprises high copy-number tandem repeat arrays (satellites) spanning hundreds of kilobases to multiple megabases in length. Monomer lengths typically range from 100–250 bp, and higher-order repeats (HOR) consisting of reiterated clusters of monomer variants

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

CellPress
OPEN ACCESS

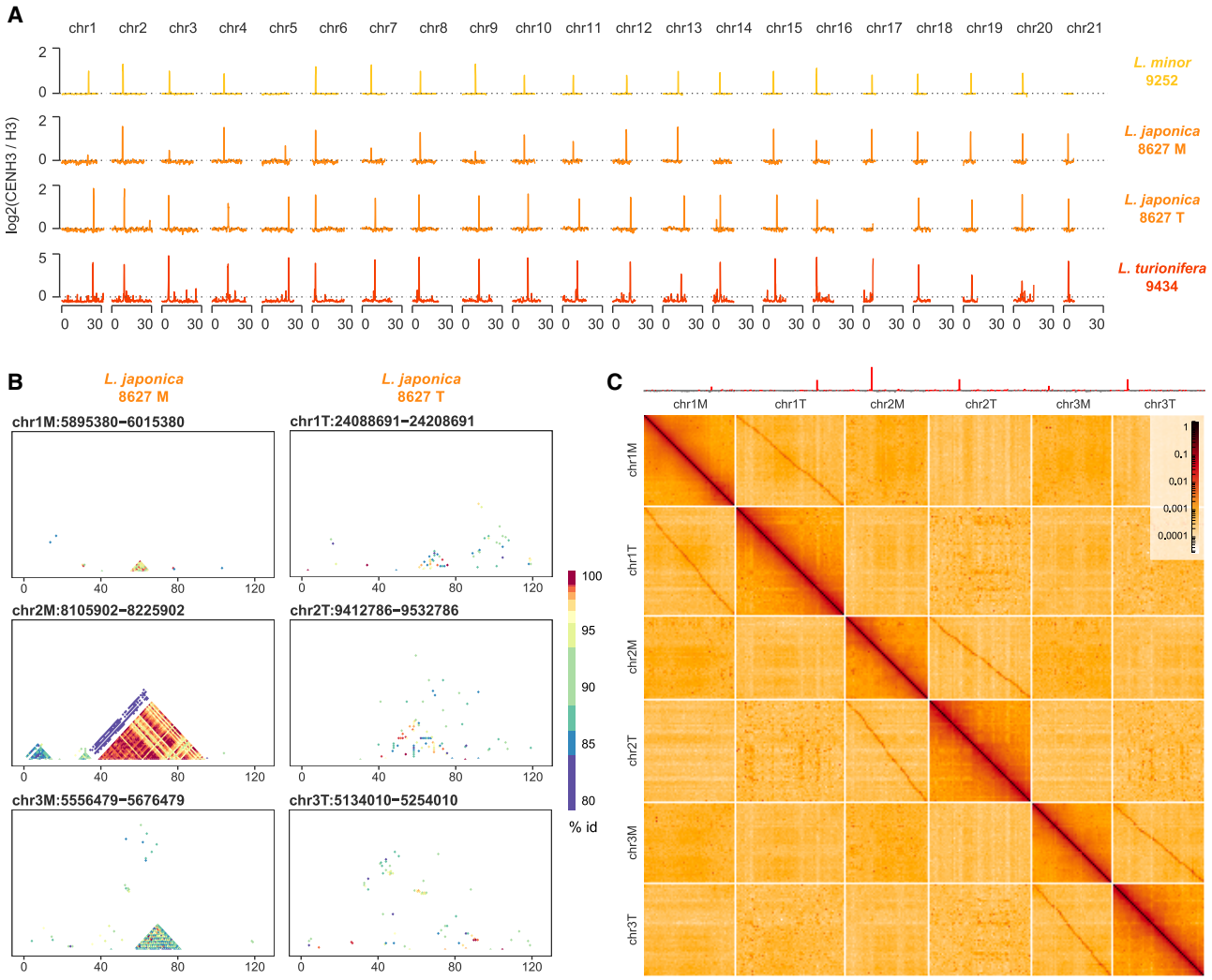


Figure 3. *Lemna* centromeres and their associated repeats

(A) Centromere positions on the chromosomes of *L. minor*, *L. turionifera*, and the homoeologous M and T *L. japonica* subgenomes. CENH3 ChIP-seq enrichment (CPM-normalized, $\log_2$[CENH3 / H3]) was computed over 100 kbp bins and averaged with adjacent bins.

(B) Self-alignment identity of 120 kbp genomic windows centered on predicted monocentromeres in *L. japonica* homeologous chromosomes 1–3.

(C) CENH3 ChIP-seq enrichment (top) and Hi-C chromatin contact maps (bottom) of *L. japonica* homeologous chromosomes 1–3. Contacts are displayed at 500 kbp resolution after balancing with the ICE method. None of the Lemnaceae chromosomes have clear centromeric contact clustering, and inter-homeolog contacts are evident in the *L. japonica* hybrid.

See also Figures S3 and S4.

are common.[26–28] Candidate centromeric satellites were previously identified in *S. polyrhiza* and *W. australiana*; however, these sequences were not highly abundant in the assembled genomes.[19,29,30] We searched the *Lemna* genomes for tandem repeats and found that *L. gibba* lacked a prominent array (Figure S3B). By contrast, *L. minor* (Lm7210 and Lm9252) had prevalent tandem repeats with monomers of 154/174/187 bp, while *L. turionifera* (Lt9434) repeats had distinct 60/105 bp monomers (Figure S3B). In all three of the hybrid *L. japonica* genomes (Lj7182, Lj8627, and Lj9421), both the *L. minor* and *L. turionifera* repeats were found on their respective subgenomes, consistent with these arrays being specific to the parental lines (Figure S3B). *L. minor* tandem repeat density tended to be enriched at a single region on each chromosome, while *L. turionifera* tandem repeats were more dispersed (Figure S3A).

Finding a lack of clear evidence of monocentric organization in at least *L. turionifera*, we raised antibodies against a peptide sequence uniquely common to *L. minor* and *L. turionifera* from the N-terminal tail of CENH3 to localize its deposition using chromatin immunoprecipitation sequencing (ChIP-seq). Among all of the chromosomes of Lm9252, Lt9434, and Lj8627, only three lacked a single predominant peak of CENH3 enrichment, indicating that all of these genomes have monocentric organization (Figures 3A and S3A). Remarkably, the position of these peaks on each chromosome was closely conserved across species, indicating that the 3 missing peaks were likely to be artifacts of centromere underassembly. We estimated the size of each

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

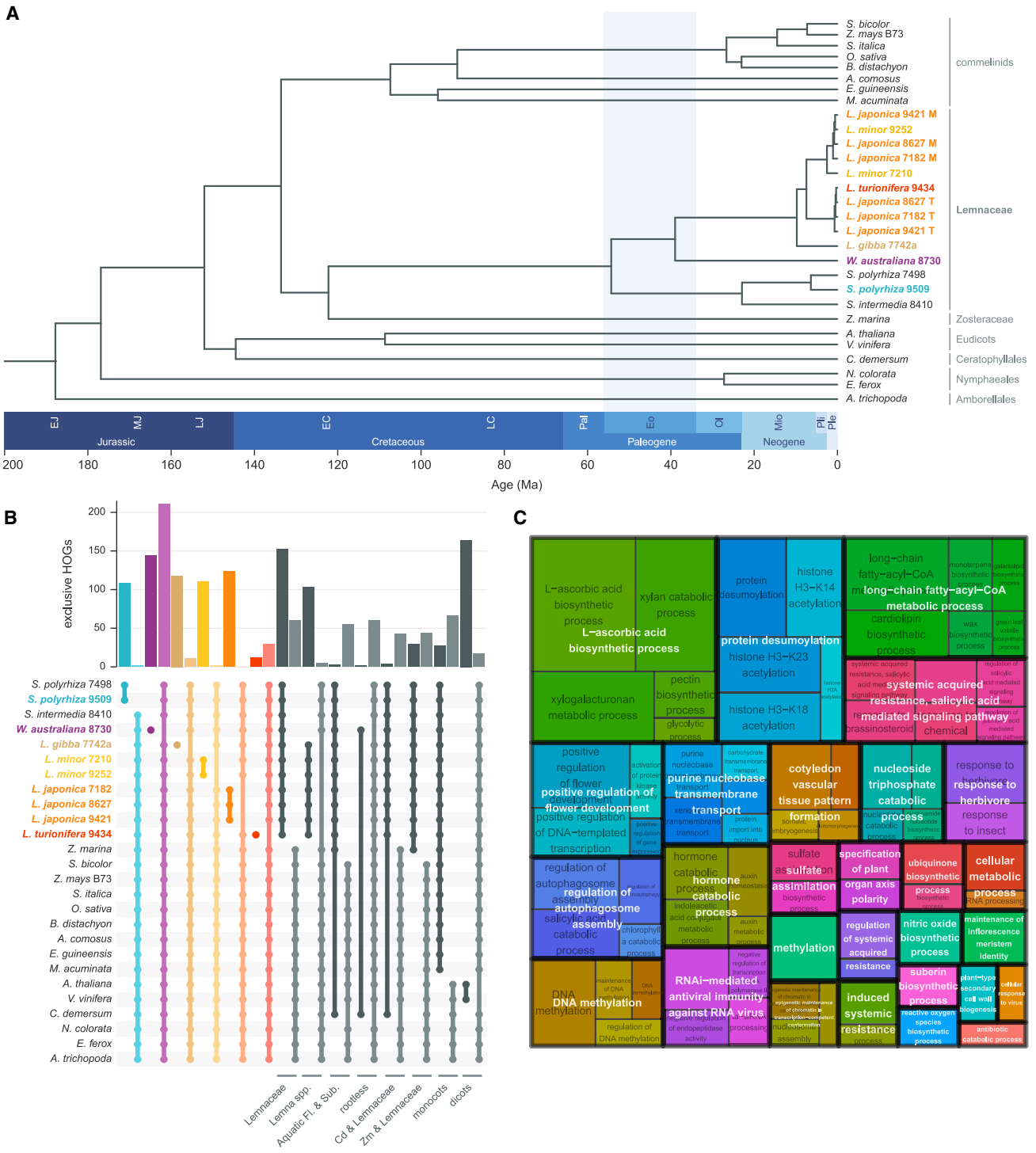


Figure 4. Phylogenomic analysis of the Lemnaceae

(A) Evolutionary relationship of chromosome-resolved Lemnaceae accessions to the angiosperms. The species tree topology was estimated from a concatenated supermatrix alignment (AA) of 854 genes identified as single copy in 87% of the analyzed species, including 11 Lemnaceae clones, 16 other monocots and dicots, and *Gnetum montanum* (not shown) as the gymnosperm outgroup. Blue shading highlights the divergence of the Lemnaceae during the Eocene (Eo). (B) Unique paralogs and missing OGs. Groups of accessions from the full phylogenomic analysis are indicated by connected dots beneath counts of the common hierarchical orthogroups (HOGs) in each group. Groups are arranged in pairs, with the first showing unique paralogs, and the second showing missing HOGs present in all other accessions. Accessions with genomes annotated in this work are highlighted in color. Higher-order phylogenetic, phenotypic, and ecological groupings are represented by gray bars. Fl. & sub., floating and submerged.

(legend continued on next page)

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

centromere region computationally (see STAR Methods) and found mean sizes of 47, 52, 70, and 109 Kbp for Lm9252, Lj8627M, Lj8627T, and Lt9434, respectively. Genomic read coverage was consistent with surrounding regions across the majority of the predicted centromeres, indicating that the small footprints are not due to collapsed repeats in the assemblies in most cases. Tandem repeats account for about 30% of the sequence content of *L. minor* centromeres but less than 2% of *L. turionifera* centromeres (Figure S3C). By contrast, *L. turionifera* centromeres contain twice the proportionate sequence content of structurally intact *Ty3* long terminal repeat (LTR) retrotransposons (LTR RT) as compared with *L. minor*. These two distinct centromere organizations—satellite-rich and satellite-poor—persisted on their respective subgenomes in hybrid *L. japonica*. We annotated protein domains and classified the intact LTR RTs underlying centromeres with DANTE,[31] and found that 43 of the 57 classified elements belong to the plant-specific CRM clade (named after the centromeric retrotransposon of maize elements)[32] of the ''chromovirus'' *Ty3* lineage. CRMs are major constituents of gymnosperm and angiosperm centromeres[33] and were the only class of intact LTR RT found in *L. minor* centromeres. We aligned the reverse transcriptase (RT) sequences of *Lemna* elements with chromovirus annotations, along with those of centromeric transposable elements (TEs) previously cataloged.[33] We found that *Lemna* centromeric chromoviruses are most closely related to ''group B'' CRMs, which lack a known targeting sequence within their integrase domain (Figure S4A), in contrast to group A and group C elements that have a CR motif or type 2 chromodomain, respectively.[33] Next, we aligned full-length nucleotide sequences of all *Lemna* centromeric elements with any chromovirus domain annotation and found that one clade (C2) is unique to *L. turionifera* genomes, and another subclade (C3) encodes only GAG domains (Figure S4B). Intact centrophilic *Ty1 ALE* elements and non-chromovirus *Ty3 ATHILA*s were also found, but exclusively in *L. turionifera* centromeres.

Chromatin contact mapping can reveal distinct interchromosomal associations typical of monocentric, polycentric, and holocentric chromosomes, respectively.[23] None of the species in this study had the stark, transverse centromeric clustering of Hi-C contacts typical of monocentric chromosomes (Figure S1C), which likely reflects the small size of the centromeres. Exceptionally, the hybrid *L. japonica* contact map had a consistent pattern of interhomeolog associations extending along the length of each chromosome (Figures 3C and S1C). Similar contact patterns have been reported between homeologs of allopolyploids as well as phased homologs in holocentric sedges.[34–37]

Gene family gain and loss in the Lemnaceae

The genomes of the freshwater and marine plants *Spirodela polyrhiza*, *Wolffia Australiana*, and *Zostera marina* have a dramatically reduced gene set. For example, many genes for stomatal development are absent from *Zostera*,[38] while genes for root development and disease resistance were lost in *Wolffia*. Consistent with a reduced morphology, single-nucleus RNA sequencing (RNA-seq) of the invasive *Lemna minuta* yielded a reduced molecular cell-type atlas.[19,30,39,40] We undertook the first comprehensive multi-genera phylogenomic analysis of Lemnaceae together with other aquatic plants based on the proteomes of 11 Lemnaceae accessions (9 of which were annotated in this study), 15 additional angiosperms, and one gymnosperm outgroup. We determined that the divergence of the Lemnaceae occurred at the beginning of the Eocene, approximately 56 mya (Figure 4A). We used OrthoFinder2 (see STAR Methods) to infer phylogenetic relationships among the genes of these species (hierarchical orthogroups or ''HOGs'') to discover common gene family losses and phylogenetically distinct paralogs across groupings of accessions (Figure 4B). To specifically examine adaptations to clonal reproduction and aquatic habits, we included *Ceratophyllum demersum*, a submerged, rootless freshwater coontail species considered to be sister to all eudicots, and *Zostera marina*, a monocot seagrass phylogenetically close to the Lemnaceae. Both species, like those in the Lemnaceae, exhibit facultative asexual reproduction.

The high quality of single-molecule genome assemblies and accurate proteome prediction (Figure S5) enabled us to pinpoint gene family losses and differentiations exclusive to species that share adaptations to clonal, aquatic, and reproductive habits. In total, we detected 60 missing HOGs in Lemnaceae, which were conserved in all other angiosperms, while 152 paralogous HOGs were found to be unique to this family (Figure 4B; Data S2A). Gene ontology (GO) term analysis grouped the predominant missing HOGs and included genes required for flower and root development, organ polarity, stomatal closure, and metabolic traits, and the striking loss of genes required for DNA methylation and RNA interference relative to the functionally annotated genomes of rice and *Arabidopsis* (Figure 4C). These candidate genes set the stage for rigorous hypothesis testing using genetic modification to establish their functional relevance.

Reduced morphology and growth habits

Lemnaceae lack root hairs and lateral roots due to the absence of a pericycle,[41] yet we found that Lemnaceae do possess orthologs of key root development genes recently reported to be lost in *S. polyrhiza*.[42] Namely, the Os*ZFP*, Os*NAL2/3* (*WOX3*), Os*ORC*3, Os*SLL1*, and Os*SNDP* families were present in all duckweeds. However, as observed in *S. polyrhiza*,[39] we found that all Lemnaceae have lost the root hair specific expansins At*EXPA7* and *18*, along with At*MYB93*, a very long-chain fatty acid responsive transcriptional regulator of lateral root development genes.[43] *AtCML1*, a Ca^{2+} sensor that regulates auxin response during primary root development,[44] and *XAL2*, a transcription factor required for root stem cell and meristem patterning,[45] were also absent in all duckweeds. *W. australiana* is rootless and lacks *WOX5*,[30,46] which encodes the homeobox transcription factor required for genesis of the meristem initials to start primary root development.[30] This absence was shared exclusively with the other rootless plant in this study,

(C) Overrepresented GO terms in the set of HOGs missing from all Lemnaceae but present in both *A. thaliana* and *O. sativa*. GO terms were grouped by semantic similarity with the relevance method and reduced with a cutoff of 0.7. Size of the rectangles is proportional to $-\log_{10}(p\ \text{value})$ using Fisher's exact test and a cutoff of 0.01.

See also Figures S5 and S7, Table S2, and Data S2.

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

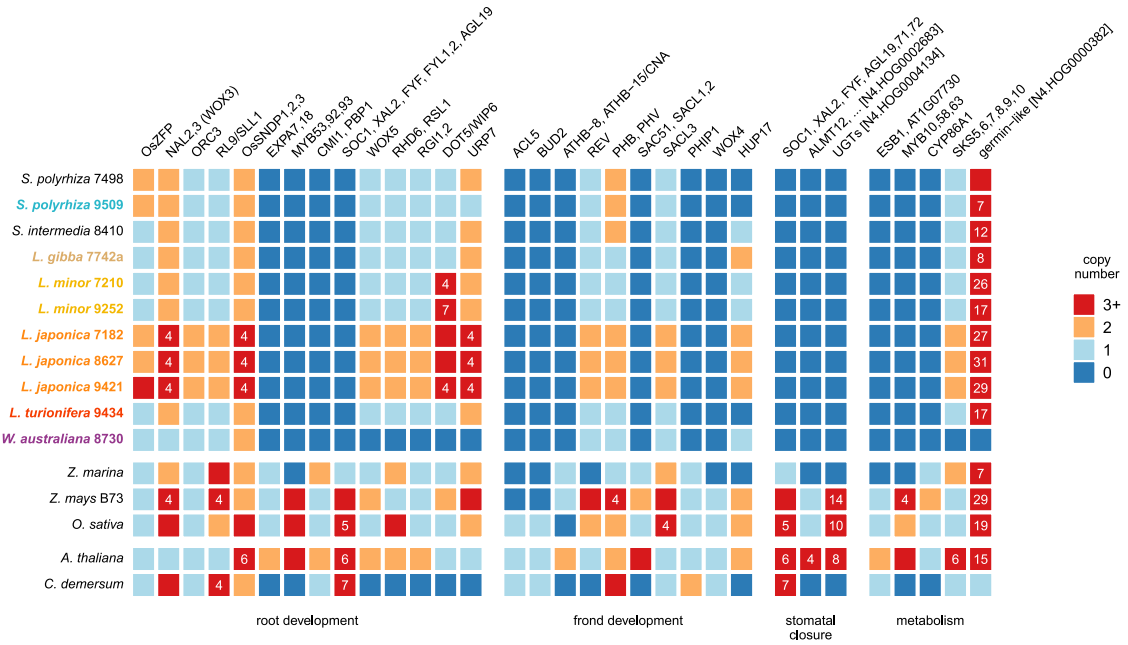


Figure 5. Gene losses shape Lemnaceae development, physiology, and metabolism

Selected gene families (hierarchical orthologous groups) were compared between the Lemnaceae and submerged marine (*Z. marina*) and freshwater (*C. demersum*) aquatic plants and compared with *Arabidopsis*, maize, and rice (Figure 4). Gene copy numbers are shown for selected genes involved in root development, frond development (stature, vascular patterning, polarity, and turion formation), stomatal closure, and metabolism. Copy numbers from 0–3 are color coded as shown, higher copy numbers are indicated.

See also Data S2B.

C. demersum, along with 59 other orthogroups (OGs) (Figure 4B; Data S2A). These include numerous root development genes also missing from rootless carnivorous and parasitic plants[47,48] (*ARF5*, *RHD6*, *RGI1* and *2*, *DOT5*, and *URP7*) (Figure 5).

Turions are the dormant buds induced by cold temperatures and low phosphate and are found in many duckweed species including *S. polyrhiza* and *L. turionifera* (Figures 1A and 2A)[49] and *W. australiana*. By contrast, *L. minor* does not form turions, prompting us to examine whether this trait was retained in hybrids. We assayed turion induction in multiple accessions of *L. minor*, *L. turionifera*, and their hybrid *L. japonica*, including Lm9252, which was phylogenetically closer to the M subgenomes of *L. japonica* hybrids than Lm7210 from South Africa (Figure 4A). Although growth rates were comparable, reaching a maximum of 24 mg per mg starting weight after only 7 days, we found that while *L.turionifera* readily formed turions under inductive conditions, 5 different accessions of *L. minor* and 6 of *L. japonica* did not form turions (Table S2). One interpretation is that *L. minor* has a dominant inhibitor of turion formation missing from *L. turionifera* but retained in *L. japonica*. Only a handful of genes match this criteria, but one candidate is *HUP17*, a gene induced by hypoxia in *Arabidopsis* that promotes senescence after prolonged submergence.[50] The ortholog in *L. minor* was missing from *L. turionifera* but present in both subgenomes of *L. japonica*. The loss of HUP17 might contribute to the endurance of turions, whose high starch content and contracted intercellular air spaces promote submergence in winter. HUP17 was also missing from *S. polyrhiza* and the other turion-producing plants in this study (*Z. marina* and *C. demersum*), with the exception of *W. australiana* 8730, an isolate from subtropical New South Wales, which is unlikely to experience prolonged winters (Figure 5).

A broadly conserved, thermospermine-mediated developmental patterning regulatory module consisting of ACL5, BUD2, and the *HD-ZIP 3* family transcription factor ATHB-8 was found to be absent from Lemnaceae. Loss-of-function mutants of the thermospermine synthase gene *ACL5* and the S-adenosyl-methionine decarboxylase gene *BUD2* exhibit severe dwarfism in *Arabidopsis*, along with xylem overproliferation and defects in auxin transport influencing vein development. Mutants of their upstream transcription factor *ATHB-8* disrupt the formation of the preprocambium and procambium, as well as xylem specification and differentiation.[51] The absence of *ACL5*, which is variable in the monocots,[52] has been observed previously in *S. polyrhiza*,[53] and here we found that all Lemnaceae additionally lack *BUD2*, *ATHB-8*, and *CORONA*. These latter two genes antagonize the roles of the other *HD-ZIP 3* members *REVOLUTA* (*REV*), *PHABULOSA* (*PHB*), and *PHAVULUTA* (*PHV*) in meristem formation, organ polarity, and vascular development.[54] Duckweeds retain *REV* but have just one homolog of *PHB* or *PHV*, while at least two paralogs were found in all but one other species. The basic helix-loop-helix (bHLH) transcription factors SAC51, SACL1, and SACL2 that participate in the ACL5-auxin feedback loop are also absent.[55] While *SACL3* is retained, *PHIP1*[56] is lost, which could enhance *acl5* dwarfism, producing the "tiny-plant" phenotype found in *acl5 sacl3* mutants of *Arabidopsis*.[55] *WOX4* is also absent, which regulates cell division in the procambium,[57] and together with downstream factors

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

such as *ATHB-8*, likely contributes to the dramatic simplification of the vascular bundle in Lemnaceae.[41] Anatomical reduction, diminished vasculature, and altered leaf polarity (e.g., the presence of adaxial, rather than abaxial stomata supportive of gas exchange in a floating habitat) could be accounted for by these losses[54] (Figure 5).

Stomatal response to elevated atmospheric carbon

We found that the key flowering regulator *SUPPRESSOR OF CONSTANS 1 (SOC1)* was missing from all analyzed Lemnaceae (N4.HOG0006359), contrary to a recent study in the short day duckweed *L. aequinoctialis*[58] but in accord with prior analysis of MADS-box genes in *S. polyrhiza*.[39] In *Arabidopsis*, *SOC1* controls drought-induced flowering[59] as well as light-induced stomatal opening.[60] Neither function would be required in duckweed fronds, which typically have open stomata and are not subject to drought. The *SOC1* paralogs *XAL2*, *FYF*, and *FYF1,2* are also missing and impact various aspects of root and floral development.[45,61] The Lemnaceae, *Z. marina*, and *C. demersum* also lack orthologs of the guard-cell expressed aluminum-activated malate transporter *ALMT12* that is largely responsible for the stomatal closure response during drought stress and also involved in the closure response to CO_2.[62,63] A high-copy family of UDP-glycosyltransferases (UGTs, N4.HOG0004134) involved in the defense response accounted for the significantly enriched GO term "abscisic acid-activated signaling pathway involved in stomatal movement" was also absent from all duckweeds, *C. demersum*, and *Z. marina* (Figure 5).

Regulation of metabolic pathways

Duckweeds have simplified metabolic pathways that are reflected in both missing and uniquely paralogous OGs in each species. This is particularly true of the polyphenolic metabolism responsible for structural rigidity of cell walls in terrestrial plants. One interesting example is the architecture of the Casparian strip (CS), a lignified cell wall important for water transport. The CS has been observed in duckweeds but has substantially reduced lignin content.[41] The complete absence of dirigent protein ESB1, responsible for lignin and alkaloid biosynthesis and for organization of the CS, is consistent with this observation. The transcription factors *MYB58* and *MYB63*, which activate lignin biosynthesis during secondary cell wall formation[64] are also missing. *Lemna* and *Wolffia* have drastically reduced xylem and lack a defined shoot endodermis, likely reflecting this loss.[65] The lipid biopolymer suberin is thought to form a diffusion barrier for water, gasses, and solutes in the lamellae that surround the CS, as well as in roots, where its engineered overproduction has been proposed as an inert polymer carbon sink for carbon sequestration applications.[66] The cytochrome P450 monooxygenase *CYP86A1* is important for cutin biosynthesis in roots and seeds[67] and was found to be missing only in Lemnaceae and the rootless *C. demersum*. In another example, most duckweeds accumulate calcium oxalate crystals in calcium-rich media, which sequester CO_2[68] but can be problematic for mammalian consumption. Radiolabeling studies in *L. minor* and other plant species have demonstrated that ascorbic acid is likely to be the predominant source of oxalic acid that gives rise to crystals sequestered in idioblast cells.[69–71] *Wolffia* is an exception, making neither druses nor raphides, and we found that *Wolffia* specifically lacks the *SKS5-8* L-ascorbate oxidase OG (N4.HOG0002231), providing a possible explanation as well as a target for genetic modification. A large family (N4.HOG0000382) of germin-like proteins possessing an oxalate oxidase enzymatic domain is also missing only in *Wolffia*, consistent with the loss of this substrate (Figure 5).

RNA interference, DNA methylation, and gene silencing

The *Spirodela polyrhiza* genome has one of the lowest levels of DNA methylation found in any angiosperm.[19] Low methylation levels could be a feature of clonal reproduction as DNA methylation levels in flowering plants are typically reset in the embryo[72] and are lost during somaclonal propagation.[73] We therefore performed whole-genome bisulfite sequencing of *L. gibba*, *L. japonica*, and *W. australiana* to determine whether methylation loss was shared with *S. polyrhiza* (Figure 6A). We also sequenced the sRNA of vegetative fronds from each of the four species (Figure 6B). We profiled coverage in both datasets over protein-coding regions and interspersed repeats including LTR RTs and DNA TEs (Figure 6C).

We found that all the duckweed genomes had low genome-wide levels of CHH methylation (0.4%–1.2%), even less than in maize (2%)[75] and close to the limit of detection. However, levels of methylation were much higher in the other Lemnaceae at CG (69%–81%) and CHG (27%–43%) sites, compared with *S. polyrhiza* (8.9% CG and 2.7% CHG) (Figure 6D). Significant levels of CG methylation at LTR RTs were detected in all four duckweed genomes, including *Spirodela* (Figure 6A). This indicates that the low level of genome-wide CG methylation found in *S. polyrhiza* reflects the scarcity of intact transposons in this species. Strikingly, CG methylation in gene bodies was absent from *S. polyrhiza*, when compared with the other Lemnaceae (Figure 6A). The absence of gene body methylation in angiosperms is thought to be an indirect consequence of the loss of CHG methylation, even though CHG methylation is restricted to transposons.[76] Consistently, CHG methylation is absent from *S. polyrhiza* transposons and reduced in *Wolffia*, which also has reduced gene body methylation (Figure 6A). sRNA sequencing revealed a predominance of 21 nt microRNA (miRNA) over 24 nt small interfering RNA (siRNA) compared with other angiosperms when mapped to the whole genome (Figure 6B). When mapped to transposons, however, we found that *S. polyrhiza* had very low levels of 24 nt siRNA as previously reported,[77] but the other species had much higher levels, corresponding more or less to the number of TEs in each genome (Figure 6C).

Next, we examined gene losses in the Lemnaceae that might account for these patterns of sRNA accumulation and DNA methylation. 24 nt sRNA precursors depend on the RNA polymerase Pol 4, and the SWI2/SNF2 chromatin remodeler genes *CLSY1-4* along with the H3K9me2 reader *SHH1* are required for Pol 4 activity.[78] HOG analysis revealed that all duckweeds have lost *CLSY1*, *CLSY2*, and *SHH1*, consistent with relatively low levels of 24 nt sRNAs in vegetative fronds compared with other angiosperms (Figure 6E). However, retention of *CLSY3* in duckweeds suggests that 24 nt siRNAs may be prevalent in the germline, where *CLSY3* regulates 24 nt siRNA production in *Arabidopsis*.[78,79] Small RNA are generated from precursors by Dicer-like RNAse 3 enzymes and loaded onto Argonaute RNAseH proteins that are required for the stability and function of sRNA in silencing genes, transposons, and viruses. We found that the Lemnaceae encode Dicer-like genes from only 3 of the 5

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

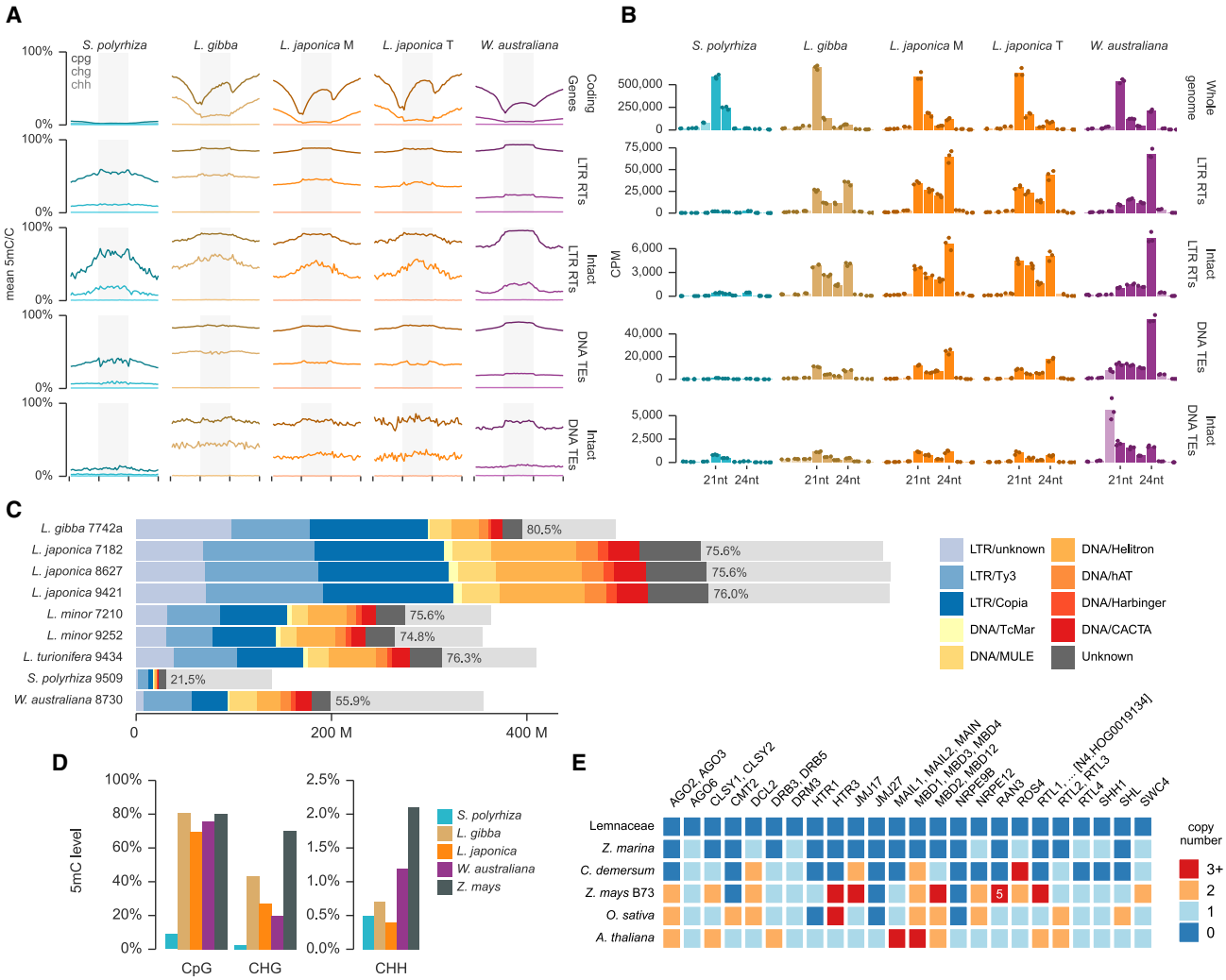


Figure 6. DNA methylation and small RNA regulation in the Lemnaceae

(A) Whole-genome bisulfite sequencing (WGBS) of Sp9509, Lg7742a, Lj8627, and Wa8730 was used to generate combined metaplots of cytosine methylation levels in the CpG, CHG, and CHH contexts over annotated regions (light gray bars) and 2 kbp upstream and downstream windows. Regions shown: coding genes, LTR retrotransposons (LTR RTs), intact LTR RTs (recently transposed), DNA transposable elements (TEs), and intact TEs, defined by intact ORFs and terminal inverted repeats (TIR).

(B) sRNA sequencing from fronds of each species was used to generate size distribution plots of 18–26 nt sRNA that mapped to the whole genome or to the transposon classes shown in part (A). Measurements from biological replicates ($n = 3$) are plotted as points, and their means, as bars. 21–24 nt sRNA are plotted in a darker shade.

(C) Transposable element repeat family content of each genome, color coded as shown and expressed as total length repeat-masked for each family in Mbp. Light gray bars show the size of each haplotype-collapsed genome assembly.

(D) Global cytosine methylation levels in each sequence context as determined by WGBS.

(E) Hierarchical ortholog groups (HOGs) for small RNA and DNA methylation gene families missing in the Lemnaceae, and compared with aquatic plants, maize (*Z. mays* B73), rice, and *Arabidopsis*. In addition to the completely missing orthogroups shown here, all Lemnaceae except Sp9509 had just one gene in the AGO4/6/9 orthogroup, resembling AGO4 (see fully annotated ortholog tables at lemna.org[74]). Color coding as in Figure 5.

See also Figure S7 and Data S2B.

angiosperm clades, as in *Spirodela*.[19] These Dicer-like genes are responsible for 21/22 nt miRNA (*DCL1*), 21 nt secondary sRNA (*DCL4*), and 24 nt siRNA (*DCL3*), but all duckweeds lack *DCL2* and *DCL5*, which are responsible for 22 and 24 nt secondary sRNA, respectively. Duckweeds also have a reduced set of only 5 Argonaute genes, compared with 10 in *Arabidopsis*, 19 in rice, and 22 in maize. They include Argonautes from each of the three major clades, which are predominantly associated with 21/22 nt and 24 nt sRNA, respectively. This is consistent with sRNA sequencing from fronds that revealed all size classes of sRNA in each species, although the relative abundance varies with the abundance of different classes of transposons (Figure 6B). Intact DNA transposons in *Wolffia* have high levels of 20–21nt siRNA, consistent with transcriptional activity and post-transcriptional silencing,[80] as do the very few DNA TE copies present in *Spirodela* (Figure 6B).

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

The Lemnaceae are specifically missing *AGO2*, *AGO3*, *AGO6*, and *AGO9* compared with maize and *Arabidopsis*. These Argonautes are highly expressed in *Arabidopsis* pollen and seeds, and their absence from the Lemnaceae could reflect their clonal growth habit.[72,81] Similarly, *DCL5* is normally expressed in the male germline of many monocots, where it is required for fertility.[82] *DCL2* is thought to be responsible for viral resistance in many angiosperms, and it is possible that Lemnaceae have less need for this particular antiviral strategy, although other aquatic plants have retained it. The lack of viral defense RNAse 3-like genes *RTL1, 2*, and *3* supports this conclusion. Intriguingly, however, the only other angiosperm in this comparison to lack *DCL2* is the African oil palm *E. guineensis*. This may reflect a role for DCL2 in reproductive isolation, as oil palm, like duckweed, is interfertile with distantly related species despite many millions of years of divergence.[83] *DCL2* has recently been found to be responsible for hybrid incompatibility and meiotic drive in maize and its relatives.[84]

DNA demethylation by ROS1 requires the histone H3K18 and H3K23 acetyltransferase gene *IDM1/ROS4* that encodes a conserved protein in the IDM complex that acts upstream of H2A.Z deposition by SWR1. In *Arabidopsis*, this mechanism rescues some euchromatic regions from promiscuous RdDM targeting.[85] Among all angiosperms analyzed here, loss of *IDM1* is only observed in duckweeds. This could reflect the very low levels of CHH methylation in vegetative fronds, making demethylation unnecessary. However, the presence of *DRM1* and *2* suggests CHH methylation is still possible, most likely in seeds and pollen grains, which have the highest levels in *Arabidopsis* but which were not examined here. Instead, the DNA methyltransferase gene *CMT2*, conserved in most other sequenced plant genomes, is responsible for high levels of heterochromatic CHH methylation in rice and *Arabidopsis* but is absent from duckweeds and maize accounting in large part for the low levels of CHH methylation (Figure 6E).[86]

Reproduction and clonal growth habits

Polyploidy is common among the Lemnaceae, and a recent meta-study of chromosome counts and genome size estimates indicate that triploidy occurs frequently in *Lemna* and *Wolffia*, with 2n ≈ 60 appearing in 9 of the 36 currently recognized Lemnaceae species, including *L. japonica* (*L. minor* × *L. turionifera*).[24] Moreover, in a companion study, we found that 29% of clones in the Landolt Duckweed Collection identified as *L. minor* had triploid *L. minor* or *L. japonica* karyotypes, with fully 70% of L. japonica hybrids being triploid.[4] Thus triploidy is prominent in wild populations. Plastome-based barcoding has demonstrated that *L. japonica* hybrids are always formed from an *L. minor* seed parent,[11] assuming maternal inheritance of plastids. Sequence-identity comparisons of the *de novo* assembled plastid and mitochondrial genomes in this study support this conclusion (Figure S6). Furthermore, the absence of RanGAP (N4.HOG0008339) genes in Lt9434 and the Lj9421 T subgenome suggests that at least some *L. turionifera* lineages might not produce viable female gametes.[87] The two major paths to polyploidy in angiosperms are somatic doubling and gametic non-reduction, with the latter being a more frequent contributor.[88] Non-reduced gamete formation resulting from abnormalities in both micro- and megasporogenesis is heritable and much more frequent in hybrids.[89] Since viable triploids with diploid contributions of either parental genome are possible (Lj7182 and Lj8627) and heterozygosity is evident in both cases (Figure 2C), unreduced gamete formation in both the *L. minor* maternal and *L. turionifera* paternal germlines is a likely explanation for the emergence of these interspecific hybrids.

Unreduced maternal gametes arise via diplospory in maize mutants of *ago104*, the ortholog of *AGO9* in *Arabidopsis*, and retain heterozygosity in unreduced gametes and their progeny.[90] *Arabidopsis ago9* mutants also have the potential to form diploid gametes via apospory, as they produce supernumerary megaspore mother cells that differentiate directly from diploid somatic cells.[91] The Lemnaceae, with the sole exception of Sp9509, only have one paralog in the *AGO4, AGO6, AGO8,* and *AGO9* clade, which appears to be related to *AGO4* (see fully annotated ortholog tables at www.lemna.org[74]). *AGO9* is highly conserved among angiosperms, and its loss from duckweeds is an unusual feature that could account for the origin of triploid hybrids such as *L. japonica* 8627, whose maternal *L. minor* parent appears to have had unreduced heterozygous gametes. By contrast, *L. japonica* 7182 has two copies of the paternal *L. turionifera* genome, indicating unreduced paternal gametes. So far, direct observations of pollen development in Lemnaceae have been limited to *L. aequinoctialis* (formerly *L. paucicostata* HEGELM.),[92] which was found to be tricellular. One candidate explanation for the production of 2n male gametes is disruption of *JASON* (*JAS*), a positive transcriptional regulator of *PARALLEL SPINDLES1 (AtPS1)* in meiotic cells that is required for pollen meiosis 2 spindle polarity but not involved in female meiosis. In *Arabidopsis*, homozygous mutations in *JAS*, as in *AtPS1*, cause heterozygous 2n pollen formation at rates up to 60%[93–95] and result in fertile haploids.[96] We found deletions and mutations in each of the two *JAS*-like loci in *Lemna* spp. in regions deeply conserved across other taxa (Figure S7A). None of the predicted *JAS* orthologs in aquatic plants possess the N-terminal Golgi-localization peptide found in terrestrial plants under hypoxia.[97]

Triploid hybrids are relatively rare among angiosperms, due to the triploid block, a prevalent form of reproductive isolation in which seeds fertilized by unreduced diploid pollen abort due to an imbalance in parental genome dosage disrupting endosperm cellularization.[98–100] This incompatibility can be described as a mismatch in endosperm balance number (EBN)—the developmentally ideal ratio of maternal-to-paternal genome dosage. EBN is thought to depend on the level of sRNA in pollen, and ecotypes lacking specific miRNA triggers, as well as mutants in several genes in the RdDM pathway, reduce or eliminate the triploid block in *Arabidopsis*.[12,13,101,102] These mutants include *ago6*, which is absent from the Lemnaceae, but completely conserved in other taxa (Figures 4C and 6D), potentially accounting for the prevalence of triploids that we observe.[4] Sexual reproduction strongly selects against triploids due to aneuploid swarms, whereby unequal segregation in triploid meiosis results in aneuploidy and severe fitness penalties.[103,104] Clonal reproduction from germinating triploid seeds avoids meiosis and enables other advantages of increased heterozygosity and gene dosage.

MSH4 and its heterodimer partner *MSH5* (MutSγ) are meiosis-specific mismatch repair proteins in the ZMM pathway required for the formation of Class 1 crossovers responsible for 80% to

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

90% of chiasmata via stabilization of double Holliday junctions.[105,106] After polyploidization in plants, meiotic recombination genes are the most rapidly lost genes in the genome,[107] and while supernumerary MutSγ copies do not increase total crossover number, reduction to a single copy per subgenome prevents interhomeolog crossovers, benefiting chromosome segregation in hybrids.[108] The *MSH4* orthologs in Lm9252 and the M subgenomes of Lj8627 and Lj9421 share a 163 residue N-terminal truncation, entirely eliminating the Holliday junction-binding MutS2 domain. In the homozygous Lm7210, this extends to 234 residues, partially encroaching on the MutS3 domain (Figure S7B). Similar N-terminal truncations of *TaMSH4D* and *TaMSH5B* resulting in pseudogenization have been noted in the subgenomes of allohexaploid wheat and its ancestral tetraploids.[109] By contrast, all other Lemnaceae orthologs, including that of Lj7182 (MTT) subgenome M, are full length.

DISCUSSION

The chromosome-level genome sequence assemblies reported here provide insight into the evolutionary history, reduced morphology, and reproductive growth habits of the Lemnaceae, the world's smallest but fastest-growing flowering plants. Our evolutionary analysis suggests that the Lemnaceae arose in the Cretaceous but diverged in the mid-Eocene, coincident with the "Azolla event," when arctic core samples suggest that huge blooms of the freshwater aquatic fern Azolla grew in the inland palaearctic sea.[110,111] These blooms are thought to be responsible for the 90% reduction in atmospheric carbon, from 3,600 to 300 ppm, in less than a million years. Although they are much harder to detect in core samples, Lemnaceae fossils have been found in shale deposits from this time and likely co-habited these warm freshwater environments.[112] Aquatic plants are uniquely adapted to high-CO_2 environments, as stomatal closure in response to elevated CO_2 has been lost in many species, and the photosynthetic rate can thus increase with rising CO_2. We found that, unlike the submerged seagrass family Zosteraceae, Lemnaceae have retained key patterning genes required for guard-cell formation, but they have lost at least 3 master regulators of stomatal closure in response to light, drought, and CO_2.

Along with genes required for lateral roots and root hairs, we found that genes for acquired and systemic disease resistance are largely missing from *Lemna* as in *Wolffia* and *Spirodela*.[19,30] This may reflect an adaptive advantage in their common floating freshwater habitat shared with waterfowl and other metazoans. In addition, a wealth of genes encoding and regulating metabolic enzymes are either missing or have unique paralogs in *Lemna*. Examples include unique and missing paralogs in long-chain fatty acid biosynthesis and in the suberin biosynthetic pathway. We have recently engineered *L. japonica* to produce and accumulate 100 times more oil (triacylglycerol) than in wild-type fronds, and long-chain lengths in this context are consistent with our findings.[3] Suberin accumulation in roots has been proposed as a strategy for carbon sequestration in terrestrial crop plants[66] but would need engineering (like oil) to be successful in Lemnaceae. An alternative carbon sink could be calcium oxalate and we identify the biosynthetic pathway found in duckweeds. Finally, Lemnaceae are a promising high-protein crop, in part due to reduced cell size relative to the number of plastids,[1] which provide most of the protein in leaves. We identify the loss of a spermine-TF module that may be responsible for the reduced stature and adaxialized polarity that underlie this key trait.

The clonal growth habit of Lemnaceae and other aquatic plants accompanies dramatic changes in chromosome biology and epigenetic regulation, consistent with prolonged clonal expansion in the absence of meiosis. Loss of transposons is thought to be a consequence of the clonal growth habit, as transposons require meiotic recombination to increase in copy number.[113–115] The more drastic loss of transposons in *Spirodela* in comparison to *Lemna* and *Wolffia* may reflect a decreased propensity for sexual reproduction in *Spirodela*,[17,116] coupled with deletion-biased somatic double-strand break repair, given the extremely high ratio of solo-LTRs to intact retrotransposons.[117,118] Consistently, *Lemna* and *Wolffia* have far greater numbers of recently active LTR RTs as evidenced by high identity LTR sequences and high levels of CpG methylation. The Lemnaceae have lost several genes encoding key components of the RdDM pathway, notably *CLSY1* and *2*, as well as *AGO6* and *AGO9*. *AGO6* is responsible for *de novo* transgene transcriptional silencing,[119] which could make duckweeds more permissive for transgenic applications. But why would it be advantageous for a clonally propagating plant to lose this aspect of gene silencing? Ectopic DNA methylation occurs spontaneously in seed plants and depends on RdDM[120] but is reprogrammed during reproductive development, which removes epigenetic variation in pollen and re-establishes parental patterns of methylation in the seed.[72,121] For thousands of asexual generations at a time, clonally propagating Lemnaceae do not undergo meiosis and therefore do not undergo reprogramming, potentially leading to clonally inherited deleterious epigenetic variation.[122] Therefore, losing at least some aspects of *de novo* methylation would mitigate these risks.

The downside of losing *AGO6* and other components of RdDM is that duckweeds may have lost the triploid block, allowing the formation of triploid hybrids when sexual reproduction does occur. These are prevalent within *L. japonica*, with 70% of surveyed hybrids found to be triploid.[4] But triploids are only problematic for sexual, not clonal, reproduction, and clonally dividing cell cultures, at least in *Arabidopsis*, also dispense with RdDM.[73] It is quite possible that some level of RdDM might be restored in seed and pollen, when cell division ceases, as observed in *Arabidopsis*.[72,73] One explanation for the prevalence of polyploidy in the Lemnaceae is the presence of defective homologs of *JASON* that could result in high frequencies of unreduced paternal gametes, and the loss of *ago9*, which could result in unreduced maternal gametes. Finally, defective orthologs of *MSH4* in the Lemnaceae would reduce homoeologous recombination in balanced polyploids, promoting fertility as in other polyploid species.[123] The presence of interhomeolog contacts in *L. japonica* Hi-C maps suggests that recombination could be potentiated if these associations also exist in meiotic cells. Our phylogenomic accounting of gene family copy-number variation provides the foundation to test these exciting hypotheses with reverse genetics.

CENH3 ChIP-seq demonstrated that at least some *Lemna* species are typified by small monocentromeres that remain at

the same position across species despite their unrelated underlying sequence. Those of *L. minor* are partially occupied by typical satellites with higher-order tandem repeats, while *L. turionifera* genomes have centromeres composed of centrophilic LTR RTs instead. These results strongly suggest the rapid gain and loss of centromeric sequence features despite the epigenetic retention of the CENH3 position. Future studies focusing on these two centromere substrates using Lemnaceae hybrids could shed light on their evolutionary dynamics.

In summary, the complete genomes of floating freshwater Lemnaceae pave the way for understanding and exploiting their regular division as novel crops, robust platforms for biomass and biotechnology applications, as well as their ancient and enormous potential for climate amelioration.

Draft versions of the genome assemblies and annotations presented here were released ahead of publication at www.lemna.org[74] and have already been utilized in several studies.[6,7,15,124–130]

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and materials should be directed to and will be fulfilled by the lead contact, Robert A. Martienssen (martiens@cshl.edu).

Materials availability

All materials generated in this study are available upon request to the lead contact.

Data and code availability

Sequencing datasets and genome assemblies generated during this study are available at NCBI (GEO: GSE238136, BioProject: PRJNA999459). Genome assemblies and annotations presented here are available along with browsing and analysis tools at www.lemna.org. Code is available at https://github.com/martienssenlab/lemnaceae-genomes-manuscript.

ACKNOWLEDGMENTS

We thank our colleagues in the duckweed community for their many contributions and enthusiastic support. This work was primarily supported by Howard Hughes Medical Institute (R.A.M.) and the U.S. Department of Energy, Office of Science, and Office of Biological and Environmental Research program under award number DE-SC0018244 (E.E., K.A., B.P., C.M.-E., U.R., E.L., and R.A.M.), as well as the Foundation for Food and Agricultural Research, Seeding Solutions grant CA21-SS-0000000100 (E.E., C.M.-E., U.R., and R.A.M). In addition, this work was supported by a Hatch project (12116), a Multi-State Capacity project (NJ12710) from the New Jersey Agricultural Experiment Station at Rutgers University (K.A., B.P., and E.L.), and the Tang Genomics Fund (N.H., B.A., K.C., A.A., and T.P.M.). Analyses were performed on the HPC cluster at Cold Spring Harbor Laboratory, supported by the U.S. National Institutes of Health (NIH) grant S10OD028632-01.

AUTHOR CONTRIBUTIONS

Conceptualization, E.E., T.P.M., I.S., E.L., and R.A.M.; investigation, E.E., B.A., K.A., P.T.N.H., C.M.-E., V.S., B.P., P.S.A., K.C., A.A., U.R., and J.A.B.; formal analysis, E.E., B.A., K.A., V.S., N.H., K.C., A.A., I.S., T.P.M., and R.A.M.; data curation, E.E.; writing – original draft, E.E., T.P.M., and R.A.M.; writing – review & editing, E.E., J.A.B., I.S., E.L., T.P.M., and R.A.M.; visualization, E.E., C.M.-E., V.S., T.P.M., and R.A.M.; project administration, E.E. and R.A.M.; and funding acquisition, E.E., E.L., T.P.M., and R.A.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Plant stocks and growth conditions
- METHOD DETAILS
  - HMW DNA extraction
  - Long read whole genome sequencing
  - Short read whole genome sequencing
  - Chromatin immunoprecipitation (ChIP)
  - Chromatin conformation capture (Hi-C)
  - Heterozygous variant calling
  - Whole genome bisulfite sequencing (WGBS)
  - Long read direct methylation analysis
  - Transcriptome sequencing
  - Small RNA sequencing and analysis
  - Genome assembly
  - Organelle genome assembly and annotation
  - Repetitive and non-coding sequence annotation
  - Gene prediction and annotation
  - Phylogenetic analysis
  - Synteny analysis
  - Genome size estimation by flow cytometry
  - GISH (Genomic *in situ* hybridization)
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2025.03.013.

Received: January 27, 2025
Revised: March 3, 2025
Accepted: March 11, 2025
Published: April 1, 2025

REFERENCES

1. Acosta, K., Appenroth, K.J., Borisjuk, L., Edelman, M., Heinig, U., Jansen, M.A.K., Oyama, T., Pasaribu, B., Schubert, I., Sorrels, S., et al. (2021). Return of the Lemnaceae: duckweed as a model plant system in the genomics and postgenomics era. Plant Cell *33*, 3207–3234. https://doi.org/10.1093/plcell/koab189.
2. Mateo-Elizalde, C., Lynn, J., Ernst, E., and Martienssen, R. (2023). Duckweeds. Curr. Biol. *33*, R89–R91. https://doi.org/10.1016/j.cub.2022.12.036.
3. Liang, Y., Yu, X.-H., Anaokar, S., Shi, H., Dahl, W.B., Cai, Y., Luo, G., Chai, J., Cai, Y., Mollá-Morales, A., et al. (2023). Engineering triacylglycerol accumulation in duckweed (Lemna japonica). Plant Biotechnol. J. *21*, 317–330. https://doi.org/10.1111/pbi.13943.
4. Michael, T.P., Hoang, P.T.N., Abramson, B.W., Mamerto, A., Pasaribu, B., Ernst, E., Allsing, N., Braglia, L., Petrus, S., Fuchs, J., et al. (2025). Triploidy is prominent in the duckweed *Lemna minor* complex. Preprint at bioRxiv. https://doi.org/10.1101/2025.02.18.638736.
5. Landolt, E. (1986). The Family of Lemnaceae – A Monographic Study (Vol. 1) (Veroffentlichungen des Geobotanischen Institutes ETH, Stiftung Rubel).

6. Braglia, L., Lauria, M., Appenroth, K.J., Bog, M., Breviario, D., Grasso, A., Gavazzi, F., and Morello, L. (2021). Duckweed Species Genotyping and Interspecific Hybrid Discovery by Tubulin-Based Polymorphism Fingerprinting. Front. Plant Sci. *12*, 625670. https://doi.org/10.3389/fpls.2021.625670.

7. Braglia, L., Breviario, D., Gianì, S., Gavazzi, F., De Gregori, J., and Morello, L. (2021). New Insights into Interspecific Hybridization in Lemna L. Sect. Lemna (Lemnaceae Martinov). Plants (Basel) *10*, 2767. https://doi.org/10.3390/plants10122767.

8. Washburn, J.D., and Birchler, J.A. (2014). Polyploids as a "model system" for the study of heterosis. Plant Reprod. *27*, 1–5. https://doi.org/10.1007/s00497-013-0237-4.

9. Bergmann, B.A., Cheng, J., Classen, J., and Stomp, A.M. (2000). In vitro selection of duckweed geographical isolates for potential use in swine lagoon effluent renovation. Bioresour. Technol. *73*, 13–20. https://doi.org/10.1016/S0960-8524(99)00137-6.

10. Bergmann, B.A., Cheng, J., Classen, J., and Stomp, A.-M. (2000). Nutrient removal from swine lagoon effluent by duckweed. Transactions of the ASAE *43*, 263–269. https://doi.org/10.13031/2013.2701.

11. Borisjuk, N., Chu, P., Gutierrez, R., Zhang, H., Acosta, K., Friesen, N., Sree, K.S., Garcia, C., Appenroth, K.J., and Lam, E. (2015). Assessment, validation and deployment strategy of a two-barcode protocol for facile genotyping of duckweed species. Plant Biol. (Stuttg) *17*, 42–49. https://doi.org/10.1111/plb.12229.

12. Martinez, G., Wolff, P., Wang, Z., Moreno-Romero, J., Santos-González, J., Conze, L.L., DeFraia, C., Slotkin, R.K., and Köhler, C. (2018). Paternal easiRNAs regulate parental genome dosage in Arabidopsis. Nat. Genet. *50*, 193–198. https://doi.org/10.1038/s41588-017-0033-4.

13. Satyaki, P.R.V., and Gehring, M. (2019). Paternally Acting Canonical RNA-Directed DNA Methylation Pathway Genes Sensitize Arabidopsis Endosperm to Paternal Genome Dosage. Plant Cell *31*, 1563–1578. https://doi.org/10.1105/tpc.19.00047.

14. Yamamoto, Y.T., Rajbhandari, N., Lin, X., Bergmann, B.A., Nishimura, Y., and Stomp, A.-M. (2001). Genetic transformation of duckweed Lemna gibba and Lemna minor. In Vitro Cell.Dev.Biol.-Plant *37*, 349–353. https://doi.org/10.1007/s11627-001-0062-6.

15. Cantó-Pastor, A., Mollá-Morales, A., Ernst, E., Dahl, W., Zhai, J., Yan, Y., Meyers, B.C., Shanklin, J., and Martienssen, R. (2015). Efficient transformation and artificial miRNA gene silencing in Lemna minor. Plant Biol. (Stuttg) *17*, 59–65. https://doi.org/10.1111/plb.12215.

16. Krasovec, M., Hoshino, M., Zheng, M., Lipinska, A.P., and Coelho, S.M. (2023). Low Spontaneous Mutation Rate in Complex Multicellular Eukaryotes with a Haploid–Diploid Life Cycle. Mol. Biol. Evol. *40*, msad105. https://doi.org/10.1093/molbev/msad105.

17. Ho, E.K.H., Bartkowska, M., Wright, S.I., and Agrawal, A.F. (2019). Population genomics of the facultatively asexual duckweed Spirodela polyrhiza. New Phytol. *224*, 1361–1371. https://doi.org/10.1111/nph.16056.

18. Xu, S., Stapley, J., Gablenz, S., Boyer, J., Appenroth, K.J., Sree, K.S., Gershenzon, J., Widmer, A., and Huber, M. (2019). Low genetic variation is associated with low mutation rate in the giant duckweed. Nat. Commun. *10*, 1243. https://doi.org/10.1038/s41467-019-09235-5.

19. Michael, T.P., Bryant, D., Gutierrez, R., Borisjuk, N., Chu, P., Zhang, H., Xia, J., Zhou, J., Peng, H., El Baidouri, M., et al. (2017). Comprehensive definition of genome features in Spirodela polyrhiza by high-depth physical mapping and short-read DNA sequencing strategies. Plant J. *89*, 617–635. https://doi.org/10.1111/tpj.13400.

20. Hoang, P.T.N., Schubert, V., Meister, A., Fuchs, J., and Schubert, I. (2019). Variation in genome size, cell and nucleus volume, chromosome number and rDNA loci among duckweeds. Sci. Rep. *9*, 3234. https://doi.org/10.1038/s41598-019-39332-w.

21. Volkov, R.A., Komarova, N.Y., and Hemleben, V. (2007). Ribosomal DNA in plant hybrids: Inheritance, rearrangement, expression. Syst. Biodivers. *5*, 261–276. https://doi.org/10.1017/S1477200007002447.

22. Naish, M., and Henderson, I.R. (2024). The structure, function, and evolution of plant centromeres. Genome Res. *34*, 161–178. https://doi.org/10.1101/gr.278409.123.

23. Hofstatter, P.G., Thangavel, G., Lux, T., Neumann, P., Vondrak, T., Novak, P., Zhang, M., Costa, L., Castellani, M., Scott, A., et al. (2022). Repeat-based holocentromeres influence genome architecture and karyotype evolution. Cell *185*, 3153–3168.e18. https://doi.org/10.1016/j.cell.2022.06.045.

24. Hoang, P.T.N., Fuchs, J., Schubert, V., Tran, T.B.N., and Schubert, I. (2022). Chromosome Numbers and Genome Sizes of All 36 Duckweed Species (Lemnaceae). Plants (Basel) *11*, 2674. https://doi.org/10.3390/plants11202674.

25. Cao, H.X., Vu, G.T.H., Wang, W., Messing, J., and Schubert, I. (2015). Chromatin organisation in duckweed interphase nuclei in relation to the nuclear DNA content. Plant Biol. (Stuttg) *17*, 120–124. https://doi.org/10.1111/plb.12194.

26. Melters, D.P., Bradnam, K.R., Young, H.A., Telis, N., May, M.R., Ruby, J.G., Sebra, R., Peluso, P., Eid, J., Rank, D., et al. (2013). Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. Genome Biol. *14*, R10. https://doi.org/10.1186/gb-2013-14-1-r10.

27. VanBuren, R., Bryant, D., Edger, P.P., Tang, H., Burgess, D., Challabathula, D., Spittle, K., Hall, R., Gu, J., Lyons, E., et al. (2015). Single-molecule sequencing of the desiccation-tolerant grass Oropetium thomaeum. Nature *527*, 508–511. https://doi.org/10.1038/nature15714.

28. Naish, M., Alonge, M., Wlodzimierz, P., Tock, A.J., Abramson, B.W., Schmücker, A., Mandáková, T., Jamge, B., Lambing, C., Kuo, P., et al. (2021). The genetic and epigenetic landscape of the Arabidopsis centromeres. Science *374*, eabi7489. https://doi.org/10.1126/science.abi7489.

29. Harkess, A., McLoughlin, F., Bilkey, N., Elliott, K., Emenecker, R., Mattoon, E., Miller, K., Czymmek, K., Vierstra, R.D., Meyers, B.C., et al. (2021). Improved Spirodela polyrhiza genome and proteomic analyses reveal a conserved chromosomal structure with high abundance of chloroplastic proteins favoring energy production. J. Exp. Bot. *72*, 2491–2500. https://doi.org/10.1093/jxb/erab006.

30. Michael, T.P., Ernst, E., Hartwick, N., Chu, P., Bryant, D., Gilbert, S., Ortleb, S., Baggs, E.L., Sree, K.S., Appenroth, K.J., et al. (2021). Genome and time-of-day transcriptome of Wolffia australiana link morphological minimization with gene loss and less growth control. Genome Res. *31*, 225–238. https://doi.org/10.1101/gr.266429.120.

31. Novák, P., Hoštáková, N., Neumann, P., and Macas, J. (2024). Dante and DANTE_LTR: lineage-centric annotation pipelines for long terminal repeat retrotransposons in plant genomes. NAR Genom. Bioinform. *6*, lqae113. https://doi.org/10.1093/nargab/lqae113.

32. Gorinsek, B., Gubensek, F., and Kordis, D. (2004). Evolutionary genomics of chromoviruses in eukaryotes. Mol. Biol. Evol. *21*, 781–798. https://doi.org/10.1093/molbev/msh057.

33. Neumann, P., Navrátilová, A., Koblížková, A., Kejnovský, E., Hřibová, E., Hobza, R., Widmer, A., Doležel, J., and Macas, J. (2011). Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mobile DNA *2*, 4. https://doi.org/10.1186/1759-8753-2-4.

34. VanBuren, R., Man Wai, C., Wang, X., Pardo, J., Yocca, A.E., Wang, H., Chaluvadi, S.R., Han, G., Bryant, D., Edger, P.P., et al. (2020). Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. Nat. Commun. *11*, 884. https://doi.org/10.1038/s41467-020-14724-z.

35. Yang, J., Wang, J., Li, Z., Li, X., He, Z., Zhang, L., Sha, T., Lyu, X., Chen, S., Gu, Y., et al. (2021). Genomic signatures of vegetable and oilseed allopolyploid Brassica juncea and genetic loci controlling the accumulation of glucosinolates. Plant Biotechnol. J. *19*, 2619–2628. https://doi.org/10.1111/pbi.13687.

36. Zhang, X., Zhang, S., Liu, Z., Zhao, W., Zhang, X., Song, J., Jia, H., Yang, W., Ma, Y., Wang, Y., et al. (2023). Characterization and acceleration of genome shuffling and ploidy reduction in synthetic allopolyploids by

genome sequencing and editing. Nucleic Acids Res. *51*, 198–217. https://doi.org/10.1093/nar/gkac1209.

37. Castellani, M., Zhang, M., Thangavel, G., Mata-Sucre, Y., Lux, T., Campoy, J.A., Marek, M., Huettel, B., Sun, H., Mayer, K.F.X., et al. (2024). Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning. Nat. Plants *10*, 423–438. https://doi.org/10.1038/s41477-024-01625-y.
38. Olsen, J.L., Rouzé, P., Verhelst, B., Lin, Y.-C., Bayer, T., Collen, J., Dattolo, E., De Paoli, E., Dittami, S., Maumus, F., et al. (2016). The genome of the seagrass Zostera marina reveals angiosperm adaptation to the sea. Nature *530*, 331–335. https://doi.org/10.1038/nature16548.
39. Wang, W., Haberer, G., Gundlach, H., Gläßer, C., Nussbaumer, T., Luo, M.C., Lomsadze, A., Borodovsky, M., Kerstetter, R.A., Shanklin, J., et al. (2014). The Spirodela polyrhiza genome reveals insights into its neotenous reduction fast growth and aquatic lifestyle. Nat. Commun. *5*, 3311. https://doi.org/10.1038/ncomms4311.
40. Abramson, B.W., Novotny, M., Hartwick, N.T., Colt, K., Aevermann, B.D., Scheuermann, R.H., and Michael, T.P. (2022). The genome and preliminary single-nuclei transcriptome of Lemna minuta reveals mechanisms of invasiveness. Plant Physiol. *188*, 879–897. https://doi.org/10.1093/plphys/kiab564.
41. Ware, A., Jones, D.H., Flis, P., Chrysanthou, E., Smith, K.E., Kümpers, B.M.C., Yant, L., Atkinson, J.A., Wells, D.M., Bhosale, R., et al. (2023). Loss of ancestral function in duckweed roots is accompanied by progressive anatomical reduction and a re-distribution of nutrient transporters. Curr. Biol. *33*, 1795–1802.e4. https://doi.org/10.1016/j.cub.2023.03.025.
42. An, D., Zhou, Y., Li, C., Xiao, Q., Wang, T., Zhang, Y., Wu, Y., Li, Y., Chao, D.-Y., Messing, J., et al. (2019). Plant evolution and environmental adaptation unveiled by long-read whole-genome sequencing of Spirodela. Proc. Natl. Acad. Sci. USA *116*, 18893–18899. https://doi.org/10.1073/pnas.1910401116.
43. Uemura, Y., Kimura, S., Ohta, T., Suzuki, T., Mase, K., Kato, H., Sakaoka, S., Uefune, M., Komine, Y., Hotta, K., et al. (2023). A very long chain fatty acid responsive transcription factor, MYB93, regulates lateral root development in Arabidopsis. Plant J. *115*, 1408–1427. https://doi.org/10.1111/tpj.16330.
44. Hazak, O., Mamon, E., Lavy, M., Sternberg, H., Behera, S., Schmitz-Thom, I., Bloch, D., Dementiev, O., Gutman, I., Danziger, T., et al. (2019). A novel Ca2+-binding protein that can rapidly transduce auxin responses during root growth. PLoS Biol. *17*, e3000085. https://doi.org/10.1371/journal.pbio.3000085.
45. Garay-Arroyo, A., Ortiz-Moreno, E., de la Paz Sánchez, M., Murphy, A.S., García-Ponce, B., Marsch-Martínez, N., de Folter, S., Corvera-Poiré, A., Jaimes-Miranda, F., Pacheco-Escobedo, M.A., et al. (2013). The MADS transcription factor XAL2/AGL14 modulates auxin transport during Arabidopsis root development by regulating PIN expression. EMBO J. *32*, 2884–2895. https://doi.org/10.1038/emboj.2013.216.
46. Park, H., Park, J.H., Lee, Y., Woo, D.U., Jeon, H.H., Sung, Y.W., Shim, S., Kim, S.H., Lee, K.O., Kim, J.-Y., et al. (2021). Genome of the world's smallest flowering plant, Wolffia australiana, helps explain its specialized physiology and unique morphology. Commun. Biol. *4*, 900. https://doi.org/10.1038/s42003-021-02422-5.
47. Ibarra-Laclette, E., Lyons, E., Hernández-Guzmán, G., Pérez-Torres, C.A., Carretero-Paulet, L., Chang, T.-H., Lan, T., Welch, A.J., Juárez, M.J.A., Simpson, J., et al. (2013). Architecture and evolution of a minute plant genome. Nature *498*, 94–98. https://doi.org/10.1038/nature12132.
48. Sun, G., Xu, Y., Liu, H., Sun, T., Zhang, J., Hettenhausen, C., Shen, G., Qi, J., Qin, Y., Li, J., et al. (2018). Large-scale gene losses underlie the genome evolution of parasitic plant Cuscuta australis. Nat. Commun. *9*, 2683. https://doi.org/10.1038/s41467-018-04721-8.
49. Pasaribu, B., Acosta, K., Aylward, A., Liang, Y., Abramson, B.W., Colt, K., Hartwick, N.T., Shanklin, J., Michael, T.P., and Lam, E. (2023). Genomics of turions from the Greater Duckweed reveal its pathways for dormancy and re-emergence strategy. New Phytol. *239*, 116–131. https://doi.org/10.1111/nph.18941.
50. Lee, S.C., Mustroph, A., Sasidharan, R., Vashisht, D., Pedersen, O., Oosumi, T., Voesenek, L.A.C.J., and Bailey-Serres, J. (2011). Molecular characterization of the submergence response of the Arabidopsis thaliana ecotype Columbia. New Phytol. *190*, 457–471. https://doi.org/10.1111/j.1469-8137.2010.03590.x.
51. Baima, S., Forte, V., Possenti, M., Peñalosa, A., Leoni, G., Salvi, S., Felici, B., Ruberti, I., and Morelli, G. (2014). Negative feedback regulation of auxin signaling by ATHB8/ACL5-BUD2 transcription module. Mol. Plant *7*, 1006–1025. https://doi.org/10.1093/mp/ssu051.
52. Roodt, D., Li, Z., Van de Peer, Y., and Mizrachi, E. (2019). Loss of Wood Formation Genes in Monocot Genomes. Genome Biol. Evol. *11*, 1986–1996. https://doi.org/10.1093/gbe/evz115.
53. Upadhyay, R.K., Shao, J., and Mattoo, A.K. (2021). Genomic analysis of the polyamine biosynthesis pathway in duckweed Spirodela polyrhiza L.: presence of the arginine decarboxylase pathway, absence of the ornithine decarboxylase pathway, and response to abiotic stresses. Planta *254*, 108. https://doi.org/10.1007/s00425-021-03755-5.
54. Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., and Clark, S.E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. Plant Cell *17*, 61–76. https://doi.org/10.1105/tpc.104.026161.
55. Cai, Q., Fukushima, H., Yamamoto, M., Ishii, N., Sakamoto, T., Kurata, T., Motose, H., and Takahashi, T. (2016). The SAC51 Family Plays a Central Role in Thermospermine Responses in Arabidopsis. Plant Cell Physiol. *57*, 1583–1592. https://doi.org/10.1093/pcp/pcw113.
56. Tanaka, T., Koyama, D., Saraumi, M., Motose, H., and Takahashi, T. (2022). RNA processing/modifying enzymes play key roles in the response to thermospermine in Arabidopsis thaliana. Preprint at bioRxiv. https://doi.org/10.1101/2022.09.19.508594.
57. Ji, J., Strable, J., Shimizu, R., Koenig, D., Sinha, N., and Scanlon, M.J. (2010). WOX4 promotes procambial development. Plant Physiol. *152*, 1346–1356. https://doi.org/10.1104/pp.109.149641.
58. Yoshida, A., Taoka, K.-I., Hosaka, A., Tanaka, K., Kobayashi, H., Muranaka, T., Toyooka, K., Oyama, T., and Tsuji, H. (2021). Characterization of frond and flower development and identification of FT and FD genes from duckweed Lemna aequinoctialis Nd. Front. Front. Plant Sci. *12*, 697206. https://doi.org/10.3389/fpls.2021.697206.
59. Hwang, K., Susila, H., Nasim, Z., Jung, J.-Y., and Ahn, J.H. (2019). Arabidopsis ABF3 and ABF4 transcription factors act with the NF-YC complex to regulate SOC1 expression and mediate drought-accelerated flowering. Mol. Plant *12*, 489–505. https://doi.org/10.1016/j.molp.2019.01.002.
60. Kimura, Y., Aoki, S., Ando, E., Kitatsuji, A., Watanabe, A., Ohnishi, M., Takahashi, K., Inoue, S.-I., Nakamichi, N., Tamada, Y., et al. (2015). A flowering integrator, SOC1, affects stomatal opening in Arabidopsis thaliana. Plant Cell Physiol. *56*, 640–649. https://doi.org/10.1093/pcp/pcu214.
61. Chen, W.-H., Lin, P.-T., Hsu, W.-H., Hsu, H.-F., Li, Y.-C., Tsao, C.-W., Hsu, M.-C., Mao, W.-T., and Yang, C.-H. (2022). Regulatory network for FOREVER YOUNG FLOWER-like genes in regulating Arabidopsis flower senescence and abscission. Commun. Biol. *5*, 662. https://doi.org/10.1038/s42003-022-03629-w.
62. Sasaki, T., Mori, I.C., Furuichi, T., Munemasa, S., Toyooka, K., Matsuoka, K., Murata, Y., and Yamamoto, Y. (2010). Closing plant stomata requires a homolog of an aluminum-activated malate transporter. Plant Cell Physiol. *51*, 354–365. https://doi.org/10.1093/pcp/pcq016.
63. Meyer, S., Mumm, P., Imes, D., Endler, A., Weder, B., Al-Rasheid, K.A., Geiger, D., Marten, I., Martinoia, E., and Hedrich, R. (2010). AtALMT12 represents an R-type anion channel required for stomatal movement in Arabidopsis guard cells. Plant J. *63*, 1054–1062. https://doi.org/10.1111/j.1365-313X.2010.04302.x.

64. Zhou, J., Lee, C., Zhong, R., and Ye, Z.-H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. Plant Cell *21*, 248–266. https://doi.org/10.1105/tpc.108.063321.

65. Landolt, E., Jäger-Zürn, I., and Schnell, R.A. (1998). Handbuch der Pflanzenanatomie. Encyclopedia of Plant Anatomy. In Traité d'anatomie végétale / Extreme Adaptations in Angiospermous Hydrophytes, H.J. Braun, S. Carlquist, P. Ozenda, and I. Roth, eds. (Borntraeger).

66. Serra, O., and Geldner, N. (2022). The making of suberin. New Phytol. *235*, 848–866. https://doi.org/10.1111/nph.18202.

67. Li, Y., Beisson, F., Koo, A.J.K., Molina, I., Pollard, M., and Ohlrogge, J. (2007). Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. Proc. Natl. Acad. Sci. USA *104*, 18339–18344. https://doi.org/10.1073/pnas.0706984104.

68. Karabourniotis, G., Horner, H.T., Bresta, P., Nikolopoulos, D., and Liakopoulos, G. (2020). New insights into the functions of carbon-calcium inclusions in plants. New Phytol. *228*, 845–854. https://doi.org/10.1111/nph.16763.

69. Franceschi, V.R. (1989). Calcium oxalate formation is a rapid and reversible process inLemna minor L. Protoplasma *148*, 130–137. https://doi.org/10.1007/BF02079332.

70. Mazen, A.M.A., Zhang, D., and Franceschi, V.R. (2004). Calcium oxalate formation in Lemna minor: physiological and ultrastructural aspects of high capacity calcium sequestration. New Phytol. *161*, 435–448. https://doi.org/10.1111/j.1469-8137.2004.00923.x.

71. Franceschi, V.R., and Nakata, P.A. (2005). CALCIUM OXALATE IN PLANTS: Formation and Function. Annu. Rev. Plant Biol. *56*, 41–71. https://doi.org/10.1146/annurev.arplant.56.032604.144106.

72. Parent, J.-S., Cahn, J., Herridge, R.P., Grimanelli, D., and Martienssen, R.A. (2021). Small RNAs guide histone methylation in Arabidopsis embryos. Genes Dev. *35*, 841–846. https://doi.org/10.1101/gad.343871.120.

73. Borges, F., Donoghue, M.T.A., LeBlanc, C., Wear, E.E., Tanurdžić, M., Berube, B., Brooks, A., Thompson, W.F., Hanley-Bowdoin, L., and Martienssen, R.A. (2021). Loss of Small-RNA-Directed DNA Methylation in the Plant Cell Cycle Promotes Germline Reprogramming and Somaclonal Variation. Curr. Biol. *31*, 591–600.e4. https://doi.org/10.1016/j.cub.2020.10.098.

74. Ernst, E., and Martienssen, R.A. (2011). Lemnaceae Genome Repository. lemna.org. http://www.lemna.org.

75. Li, Q., Eichten, S.R., Hermanson, P.J., Zaunbrecher, V.M., Song, J., Wendt, J., Rosenbaum, H., Madzima, T.F., Sloan, A.E., Huang, J., et al. (2014). Genetic perturbation of the maize methylome. Plant Cell *26*, 4602–4616. https://doi.org/10.1105/tpc.114.133140.

76. Wendte, J.M., Zhang, Y., Ji, L., Shi, X., Hazarika, R.R., Shahryary, Y., Johannes, F., and Schmitz, R.J. (2019). Epimutations are associated with CHROMOMETHYLASE 3-induced de novo DNA methylation. eLife *8*, e47891. https://doi.org/10.7554/eLife.47891.

77. Patel, P., Mathioni, S.M., Hammond, R., Harkess, A.E., Kakrana, A., Arikit, S., Dusia, A., and Meyers, B.C. (2021). Reproductive phasiRNA loci and DICER-LIKE5, but not microRNA loci, diversified in monocotyledonous plants. Plant Physiol. *185*, 1764–1782. https://doi.org/10.1093/plphys/kiab001.

78. Zhou, M., Coruh, C., Xu, G., Martins, L.M., Bourbousse, C., Lambolez, A., and Law, J.A. (2022). The CLASSY family controls tissue-specific DNA methylation patterns in Arabidopsis. Nat. Commun. *13*, 244. https://doi.org/10.1038/s41467-021-27690-x.

79. Long, J., Walker, J., She, W., Aldridge, B., Gao, H., Deans, S., Vickers, M., and Feng, X. (2021). Nurse cell–derived small RNAs define paternal epigenetic inheritance in Arabidopsis. Science *373*, eabh0556. https://doi.org/10.1126/science.abh0556.

80. Marí-Ordóñez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., and Voinnet, O. (2013). Reconstructing de novo silencing of an active plant retrotransposon. Nat. Genet. *45*, 1029–1039. https://doi.org/10.1038/ng.2703.

81. Jullien, P.E., Schröder, J.A., Bonnet, D.M.V., Pumplin, N., and Voinnet, O. (2022). Asymmetric expression of Argonautes in reproductive tissues. Plant Physiol. *188*, 38–43. https://doi.org/10.1093/plphys/kiab474.

82. Teng, C., Zhang, H., Hammond, R., Huang, K., Meyers, B.C., and Walbot, V. (2020). Dicer-like 5 deficiency confers temperature-sensitive male sterility in maize. Nat. Commun. *11*, 2912. https://doi.org/10.1038/s41467-020-16634-6.

83. Singh, R., Ong-Abdullah, M., Low, E.-T.L., Manaf, M.A.A., Rosli, R., Nookiah, R., Ooi, L.C.-L., Ooi, S.-E., Chan, K.-L., Halim, M.A., et al. (2013). Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. Nature *500*, 335–339. https://doi.org/10.1038/nature12309.

84. Berube, B., Ernst, E., Cahn, J., Roche, B., de Santis Alves, C., Lynn, J., Scheben, A., Grimanelli, D., Siepel, A., Ross-Ibarra, J., et al. (2024). Teosinte Pollen Drive guides maize diversification and domestication by RNAi. Nature *633*, 380–388. https://doi.org/10.1038/s41586-024-07788-0.

85. Li, X., Qian, W., Zhao, Y., Wang, C., Shen, J., Zhu, J.-K., and Gong, Z. (2012). Antisilencing role of the RNA-directed DNA methylation pathway and a histone acetyltransferase in Arabidopsis. Proc. Natl. Acad. Sci. USA *109*, 11425–11430. https://doi.org/10.1073/pnas.1208557109.

86. Harkess, A., Bewick, A.J., Lu, Z., Fourounjian, P., Michael, T.P., Schmitz, R.J., and Meyers, B.C. (2024). The unusual predominance of maintenance DNA methylation in Spirodela polyrhiza. G3 (Bethesda) *14*, jkae004. https://doi.org/10.1093/g3journal/jkae004.

87. Rodrigo-Peiris, T., Xu, X.M., Zhao, Q., Wang, H.-J., and Meier, I. (2011). RanGAP is required for post-meiotic mitosis in female gametophyte development in Arabidopsis thaliana. J. Exp. Bot. *62*, 2705–2714. https://doi.org/10.1093/jxb/erq448.

88. Ramsey, J., and Schemske, D.W. (1998). Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants. Annu. Rev. Ecol. Syst. *29*, 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467.

89. Tayalé, A., and Parisod, C. (2013). Natural pathways to polyploidy in plants and consequences for genome reorganization. Cytogenet. Genome Res. *140*, 79–96. https://doi.org/10.1159/000351318.

90. Singh, M., Goel, S., Meeley, R.B., Dantec, C., Parrinello, H., Michaud, C., Leblanc, O., and Grimanelli, D. (2011). Production of viable gametes without meiosis in maize deficient for an Argonaute protein. Plant Cell *23*, 443–458. https://doi.org/10.1105/tpc.110.079020.

91. Olmedo-Monfil, V., Durán-Figueroa, N., Arteaga-Vázquez, M., Demesa-Arévalo, E., Autran, D., Grimanelli, D., Slotkin, R.K., Martienssen, R.A., and Vielle-Calzada, J.-P. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature *464*, 628–632. https://doi.org/10.1038/nature08828.

92. Maheshwari, S.C., and Kapil, R.N. (1963). Morphological and embryological studies on the Lemnaceae. I. The floral structure and gametophytes of Lemna paucicostata. Am. J. Bot. *50*, 677–686. https://doi.org/10.1002/j.1537-2197.1963.tb12242.x.

93. d'Erfurth, I., Jolivet, S., Froger, N., Catrice, O., Novatchkova, M., Simon, M., Jenczewski, E., and Mercier, R. (2008). Mutations in AtPS1 (Arabidopsis thaliana parallel spindle 1) lead to the production of diploid pollen grains. PLoS Genet. *4*, e1000274. https://doi.org/10.1371/journal.pgen.1000274.

94. Erilova, A., Brownfield, L., Exner, V., Rosa, M., Twell, D., Mittelsten Scheid, O., Hennig, L., and Köhler, C. (2009). Imprinting of the polycomb group gene MEDEA serves as a ploidy sensor in Arabidopsis. PLoS Genet. *5*, e1000663. https://doi.org/10.1371/journal.pgen.1000663.

95. De Storme, N., and Geelen, D. (2011). The Arabidopsis Mutant jason Produces Unreduced First Division Restitution Male Gametes through a Parallel/Fused Spindle Mechanism in Meiosis II. Plant Physiol. *155*, 1403–1415. https://doi.org/10.1104/pp.110.170415.

96. Aboobucker, S.I., Zhou, L., and Lübberstedt, T. (2023). Haploid male fertility is restored by parallel spindle genes in Arabidopsis thaliana. Nat. Plants *9*, 214–218. https://doi.org/10.1038/s41477-022-01332-6.

97. Cabout, S., Leask, M.P., Varghese, S., Yi, J., Peters, B., Conze, L.L., Köhler, C., and Brownfield, L. (2017). The meiotic regulator JASON utilizes alternative translation initiation sites to produce differentially localized forms. J. Exp. Bot. *68*, 4205–4217. https://doi.org/10.1093/jxb/erx222.

98. Schluter, D. (2001). Ecology and the origin of species. Trends Ecol. Evol. *16*, 372–380. https://doi.org/10.1016/s0169-5347(01)02198-x.

99. Husband, B.C. (2004). The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. Biol. J. Linn. Soc. *82*, 537–546. https://doi.org/10.1111/j.1095-8312.2004.00339.x.

100. Khouider, S., and Gehring, M. (2024). Parental dialectic: Epigenetic conversations in endosperm. Curr. Opin. Plant Biol. *81*, 102591. https://doi.org/10.1016/j.pbi.2024.102591.

101. Borges, F., Parent, J.-S., van Ex, F., Wolff, P., Martínez, G., Köhler, C., and Martienssen, R.A. (2018). Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis. Nat. Genet. *50*, 186–192. https://doi.org/10.1038/s41588-017-0032-5.

102. Duszynska, D., Vilhjalmsson, B., Castillo Bravo, R., Swamidatta, S., Juenger, T.E., Donoghue, M.T.A., Comte, A., Nordborg, M., Sharbel, T.F., Brychkova, G., et al. (2019). Transgenerational effects of inter-ploidy cross direction on reproduction and F2 seed development of Arabidopsis thaliana F1 hybrid triploids. Plant Reprod. *32*, 275–289. https://doi.org/10.1007/s00497-019-00369-6.

103. Henry, I.M., Dilkes, B.P., Young, K., Watson, B., Wu, H., and Comai, L. (2005). Aneuploidy and genetic variation in the Arabidopsis thaliana triploid response. Genetics *170*, 1979–1988. https://doi.org/10.1534/genetics.104.037788.

104. Henry, I.M., Dilkes, B.P., Tyagi, A.P., Lin, H.Y., and Comai, L. (2009). Dosage and parent-of-origin effects shaping aneuploid swarms in A. thaliana. Heredity *103*, 458–468. https://doi.org/10.1038/hdy.2009.81.

105. Higgins, J.D., Armstrong, S.J., Franklin, F.C.H., and Jones, G.H. (2004). The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in Arabidopsis. Genes Dev. *18*, 2557–2570. https://doi.org/10.1101/gad.317504.

106. Higgins, J.D., Vignard, J., Mercier, R., Pugh, A.G., Franklin, F.C.H., and Jones, G.H. (2008). AtMSH5 partners AtMSH4 in the class I meiotic crossover pathway in Arabidopsis thaliana, but is not required for synapsis. Plant J. *55*, 28–39. https://doi.org/10.1111/j.1365-313X.2008.03470.x.

107. Lloyd, A.H., Ranoux, M., Vautrin, S., Glover, N., Fourment, J., Charif, D., Choulet, F., Lassalle, G., Marande, W., Tran, J., et al. (2014). Meiotic gene evolution: can you teach a new dog new tricks? Mol. Biol. Evol. *31*, 1724–1727. https://doi.org/10.1093/molbev/msu119.

108. Gonzalo, A., Lucas, M.-O., Charpentier, C., Sandmann, G., Lloyd, A., and Jenczewski, E. (2019). Reducing MSH4 copy number prevents meiotic crossovers between non-homologous chromosomes in Brassica napus. Nat. Commun. *10*, 2354. https://doi.org/10.1038/s41467-019-10010-9.

109. Desjardins, S.D., Ogle, D.E., Ayoub, M.A., Heckmann, S., Henderson, I.R., Edwards, K.J., and Higgins, J.D. (2020). MutS homologue 4 and MutS homologue 5 Maintain the Obligate Crossover in Wheat Despite Stepwise Gene Loss following Polyploidization. Plant Physiol. *183*, 1545–1558. https://doi.org/10.1104/pp.20.00534.

110. Brinkhuis, H., Schouten, S., Collinson, M.E., Sluijs, A., Sinninghe Damsté, J.S., Dickens, G.R., Huber, M., Cronin, T.M., Onodera, J., Takahashi, K., et al. (2006). Episodic fresh surface waters in the Eocene Arctic Ocean. Nature *441*, 606–609. https://doi.org/10.1038/nature04692.

111. Moran, K., Backman, J., Brinkhuis, H., Clemens, S.C., Cronin, T., Dickens, G.R., Eynaud, F., Gattacceca, J., Jakobsson, M., Jordan, R.W., et al. (2006). The Cenozoic palaeoenvironment of the Arctic Ocean. Nature *441*, 601–605. https://doi.org/10.1038/nature04800.

112. Nauheimer, L., Metzler, D., and Renner, S.S. (2012). Global history of the ancient monocot family Araceae inferred with models accounting for past continental positions and previous ranges based on fossils. New Phytol. *195*, 938–950. https://doi.org/10.1111/j.1469-8137.2012.04220.x.

113. Hickey, D.A. (1982). Selfish DNA: a sexually-transmitted nuclear parasite. Genetics *101*, 519–531. https://doi.org/10.1093/genetics/101.3-4.519.

114. Arkhipova, I., and Meselson, M. (2000). Transposable elements in sexual and ancient asexual taxa. Proc. Natl. Acad. Sci. USA *97*, 14473–14477. https://doi.org/10.1073/pnas.97.26.14473.

115. Laine, V.N., Sackton, T.B., and Meselson, M. (2022). Genomic signature of sexual reproduction in the bdelloid rotifer Macrotrachella quadricornifera. Genetics *220*, iyab221. https://doi.org/10.1093/genetics/iyab221.

116. Crawford, D.J., and Landolt, E. (1993). Allozyme Studies in Spirodela (Lemnaceae): Variation Among Conspecific Clones and Divergence Among the Species. Syst. Bot. *18*, 389–394. https://doi.org/10.2307/2419415.

117. Dombey, R., Buendía-Ávila, D., Barragán-Borrero, V., Diezma-Navas, L., Ponce-Mañe, A., Vargas-Guerrero, J.M., Elias, R., and Marí-Ordóñez, A. (2025). Atypical epigenetic and small RNA control of degenerated transposons and their fragments in clonally reproducing Spirodela polyrhiza. Genome Res. *35*, 522–544.

118. Schubert, I., and Vu, G.T.H. (2016). Genome stability and evolution: Attempting a holistic view. Trends Plant Sci. *21*, 749–757. https://doi.org/10.1016/j.tplants.2016.06.003.

119. McCue, A.D., Panda, K., Nuthikattu, S., Choudury, S.G., Thomas, E.N., and Slotkin, R.K. (2015). ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. EMBO J. *34*, 20–35. https://doi.org/10.15252/embj.201489499.

120. Becker, C., Hagmann, J., Müller, J., Koenig, D., Stegle, O., Borgwardt, K., and Weigel, D. (2011). Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. Nature *480*, 245–249. https://doi.org/10.1038/nature10555.

121. Calarco, J.P., Borges, F., Donoghue, M.T.A., Van Ex, F., Jullien, P.E., Lopes, T., Gardner, R., Berger, F., Feijó, J.A., Becker, J.D., et al. (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell *151*, 194–205. https://doi.org/10.1016/j.cell.2012.09.001.

122. Van Antro, M., Prelovsek, S., Ivanovic, S., Gawehns, F., Wagemaker, N.C.A.M., Mysara, M., Horemans, N., Vergeer, P., and Verhoeven, K.J.F. (2023). DNA methylation in clonal duckweed (Lemna minor L.) lineages reflects current and historical environmental exposures. Mol. Ecol. *32*, 428–443. https://doi.org/10.1111/mec.16757.

123. Bomblies, K. (2023). Learning to tango with four (or more): the molecular basis of adaptation to polyploid meiosis. Plant Reprod. *36*, 107–124. https://doi.org/10.1007/s00497-022-00448-1.

124. Yan, Y., Candreva, J., Shi, H., Ernst, E., Martienssen, R., Schwender, J., and Shanklin, J. (2013). Survey of the total fatty acid and triacylglycerol composition and content of 30 duckweed species and cloning of a Δ6-desaturase responsible for the production of γ-linolenic and stearidonic acids in Lemna gibba. BMC Plant Biol. *13*, 201. https://doi.org/10.1186/1471-2229-13-201.

125. Loll, A., Reinwald, H., Ayobahan, S.U., Göckener, B., Salinas, G., Schäfers, C., Schlich, K., Hamscher, G., and Eilebrecht, S. (2022). Short-Term Test for Toxicogenomic Analysis of Ecotoxic Modes of Action in Lemna minor. Environ. Sci. Technol. *56*, 11504–11515. https://doi.org/10.1021/acs.est.2c01777.

126. Zhou, Y., Kishchenko, O., Stepanenko, A., Chen, G., Wang, W., Zhou, J., Pan, C., and Borisjuk, N. (2021). The Dynamics of NO3− and NH4+ Uptake in Duckweed Are Coordinated with the Expression of Major Nitrogen Assimilation Genes. Plants (Basel) *11*, 11. https://doi.org/10.3390/plants11010011.

127. Braglia, L., Ceschin, S., Iannelli, M.A., Bog, M., Fabriani, M., Frugis, G., Gavazzi, F., Gianì, S., Mariani, F., Muzzi, M., et al. (2024). Characterization of the cryptic interspecific hybrid Lemna×mediterranea

by an integrated approach provides new insights into duckweed diversity. J. Exp. Bot. *75*, 3092–3110. https://doi.org/10.1093/jxb/erae059.

128. Acosta, K., Sree, K.S., Okamoto, N., Koseki, K., Sorrels, S., Jahreis, G., Watanabe, F., Appenroth, K.-J., and Lam, E. (2024). Source of Vitamin B12 in plants of the Lemnaceae family and its production by duckweed-associated bacteria. J. Food Compos. Anal. *135*, 106603. https://doi.org/10.1016/j.jfca.2024.106603.

129. Schmid, M.W., Moradi, A., Leigh, D.M., Schuman, M.C., and van Moorsel, S.J. (2024). Covering the bases: Population genomic structure of Lemna minor and the cryptic species L. japonica in Switzerland. Ecol. Evol. *14*, e11599. https://doi.org/10.1002/ece3.11599.

130. Kerstetter, J., Reid, A., Armstrong, J., Zallek, T., Hobble, T., and Turcotte, M. (2023). Characterization of microsatellite markers for the duckweed Spirodela polyrhiza and Lemna minor tested on samples from Europe and the United States of America: Spirodela polyrhiza and Lemna minor microsatellites. GenResJ. *4*, 46–55. https://doi.org/10.46265/genresj.ALFV3636.

131. Wang, W., and Messing, J. (2011). High-throughput sequencing of three Lemnoideae (duckweeds) chloroplast genomes from total DNA. PLoS One *6*, e24670. https://doi.org/10.1371/journal.pone.0024670.

132. Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., et al. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. *40*, D1178–D1186. https://doi.org/10.1093/nar/gkr944.

133. Neumann, P., Novák, P., Hoštáková, N., and Macas, J. (2019). Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mobile DNA *10*, 1. https://doi.org/10.1186/s13100-018-0144-1.

134. Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. *39*, D152–D157. https://doi.org/10.1093/nar/gkq1027.

135. Kriventseva, E.V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F.A., and Zdobnov, E.M. (2019). OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res. *47*, D807–D811. https://doi.org/10.1093/nar/gky1053.

136. Kalvari, I., Nawrocki, E.P., Ontiveros-Palacios, N., Argasinska, J., Lamkiewicz, K., Marz, M., Griffiths-Jones, S., Toffano-Nioche, C., Gautheret, D., Weinberg, Z., et al. (2021). Rfam 14: expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Res. *49*, D192–D200. https://doi.org/10.1093/nar/gkaa1047.

137. Vasimuddin, M., Misra, S., Li, H., and Aluru, S. (2019). Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS), 314–324.

138. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.

139. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

140. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

141. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst. *3*, 95–98. https://doi.org/10.1016/j.cels.2016.07.002.

142. Abdennur, N., and Mirny, L.A. (2020). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics *36*, 311–316. https://doi.org/10.1093/bioinformatics/btz540.

143. Kerpedjiev, P., Abdennur, N., Lekschas, F., McCallum, C., Dinkla, K., Strobelt, H., Luber, J.M., Ouellette, S.B., Azhir, A., Kumar, N., et al. (2018). HiGlass: web-based visual exploration and analysis of genome interaction maps. Genome Biol. *19*, 125. https://doi.org/10.1186/s13059-018-1486-1.

144. Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

145. Zheng, Z., Li, S., Su, J., Leung, A.W.-S., Lam, T.-W., and Luo, R. (2022). Symphonizing pileup and full-alignment for deep learning-based long-read variant calling. Nat. Comput. Sci. *2*, 797–803. https://doi.org/10.1038/s43588-022-00387-x.

146. Smolka, M., Paulin, L.F., Grochowski, C.M., Horner, D.W., Mahmoud, M., Behera, S., Kalef-Ezra, E., Gandhi, M., Hong, K., Pehlivan, D., et al. (2024). Detection of mosaic and population-level structural variants with Sniffles2. Nat. Biotechnol. *42*, 1571–1580. https://doi.org/10.1038/s41587-023-02024-y.

147. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

148. Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. Bioinformatics *27*, 1571–1572. https://doi.org/10.1093/bioinformatics/btr167.

149. Jiang, H., Lei, R., Ding, S.-W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics *15*, 182. https://doi.org/10.1186/1471-2105-15-182.

150. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

151. Johnson, N.R., Yeoh, J.M., Coruh, C., and Axtell, M.J. (2016). Improved Placement of Multi-mapping Small RNAs. G3 (Bethesda) *6*, 2103–2111. https://doi.org/10.1534/g3.116.030452.

152. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. *37*, 540–546. https://doi.org/10.1038/s41587-019-0072-8.

153. Shafin, K., Pesout, T., Chang, P.-C., Nattestad, M., Kolesnikov, A., Goel, S., Baid, G., Kolmogorov, M., Eizenga, J.M., Miga, K.H., et al. (2021). Haplotype-aware variant calling with PEPPER-Margin-DeepVariant enables high accuracy in nanopore long-reads. Nat. Methods *18*, 1322–1332. https://doi.org/10.1038/s41592-021-01299-w.

154. Hu, J., Fan, J., Sun, Z., and Liu, S. (2020). NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics *36*, 2253–2255. https://doi.org/10.1093/bioinformatics/btz891.

155. Guan, D., McCarthy, S.A., Wood, J., Howe, K., Wang, Y., and Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics *36*, 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025.

156. Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. J. Comput. Biol. *7*, 203–214. https://doi.org/10.1089/10665270050081478.

157. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., et al. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95. https://doi.org/10.1126/science.aal3327.

158. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst. *3*, 99–101. https://doi.org/10.1016/j.cels.2015.07.012.

159. Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., Wang, X., Lippman, Z.B., Schatz, M.C., and Soyk, S. (2022). Automated assembly scaffolding using RagTag elevates a new tomato

system for high-throughput genome editing. Genome Biol. *23*, 258. https://doi.org/10.1186/s13059-022-02823-7.

160. Aury, J.-M., and Istace, B. (2021). Hapo-G, haplotype-aware polishing of genome assemblies with accurate reads. NAR Genom. Bioinform. *3*, lqab034. https://doi.org/10.1093/nargab/lqab034.

161. Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLoS One *11*, e0163962. https://doi.org/10.1371/journal.pone.0163962.

162. Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. *5*, R12. https://doi.org/10.1186/gb-2004-5-2-r12.

163. Poorten, T. (2017). dotPlotly: Generate an interactive dot plot from mummer or minimap alignments. https://github.com/tpoorten/dotPlotly.

164. Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., and Greiner, S. (2017). GeSeq - versatile and accurate annotation of organelle genomes. Nucleic Acids Res. *45*, W6–W11. https://doi.org/10.1093/nar/gkx391.

165. Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J.R.A., Hellinga, A.J., Lugo, C.S.B., Elliott, T.A., Ware, D., Peterson, T., et al. (2019). Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. Genome Biol. *20*, 275. https://doi.org/10.1186/s13059-019-1905-y.

166. Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. *27*, 573–580. https://doi.org/10.1093/nar/27.2.573.

167. Nawrocki, E.P., and Eddy, S.R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics *29*, 2933–2935. https://doi.org/10.1093/bioinformatics/btt509.

168. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915. https://doi.org/10.1038/s41587-019-0201-4.

169. Brůna, T., Hoff, K.J., Lomsadze, A., Stanke, M., and Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genom. Bioinform. *3*, lqaa108. https://doi.org/10.1093/nargab/lqaa108.

170. Gabriel, L., Hoff, K.J., Brůna, T., Borodovsky, M., and Stanke, M. (2021). TSEBRA: transcript selector for BRAKER. BMC Bioinformatics *22*, 566. https://doi.org/10.1186/s12859-021-04482-0.

171. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

172. Song, L., Sabunciyan, S., Yang, G., and Florea, L. (2019). A multi-sample approach increases the accuracy of transcript assembly. Nat. Commun. *10*, 5000. https://doi.org/10.1038/s41467-019-12990-0.

173. Kovaka, S., Zimin, A.V., Pertea, G.M., Razaghi, R., Salzberg, S.L., and Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol. *20*, 278. https://doi.org/10.1186/s13059-019-1910-1.

174. Sahlin, K., and Mäkinen, V. (2021). Accurate spliced alignment of long RNA sequencing reads. Bioinformatics *37*, 4643–4651. https://doi.org/10.1093/bioinformatics/btab540.

175. Mapleson, D., Venturini, L., Kaithakottil, G., and Swarbreck, D. (2018). Efficient and accurate detection of splice junctions from RNA-seq with Portcullis. GigaScience *7*, giy131. https://doi.org/10.1093/gigascience/giy131.

176. Venturini, L., Caim, S., Kaithakottil, G.G., Mapleson, D.L., and Swarbreck, D. (2018). Leveraging multiple transcriptome assembly methods for improved gene structure annotation. GigaScience *7*, giy093. https://doi.org/10.1093/gigascience/giy093.

177. Haas, B., and Papanicolaou, A.T.D. (2016). TransDecoder. https://github.com/TransDecoder.

178. Keilwagen, J., Wenk, M., Erickson, J.L., Schattat, M.H., Grau, J., and Hartung, F. (2016). Using intron position conservation for homology-based gene prediction. Nucleic Acids Res. *44*, e89. https://doi.org/10.1093/nar/gkw092.

179. Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., and Zdobnov, E.M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol. Biol. Evol. *38*, 4647–4654. https://doi.org/10.1093/molbev/msab199.

180. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

181. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

182. Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I.T., Novak, A.M., Deran, A., Fang, Q., Xie, D., Feng, S., Stiller, J., et al. (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature *587*, 246–251. https://doi.org/10.1038/s41586-020-2871-y.

183. König, S., Romoth, L.W., Gerischer, L., and Stanke, M. (2016). Simultaneous gene finding in multiple genomes. Bioinformatics *32*, 3388–3395. https://doi.org/10.1093/bioinformatics/btw494.

184. Dainat, J., Hereñú, D., and Pucholt, P. (2020) AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format. Version v0. https://doi.org/10.5281/zenodo.3549547.

185. Holt, C., and Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics *12*, 491. https://doi.org/10.1186/1471-2105-12-491.

186. Zhang, R.-G., Li, G.-Y., Wang, X.-L., Dainat, J., Wang, Z.-X., Ou, S., and Ma, Y. (2022). TEsorter: an accurate and fast method to classify LTR-retrotransposons in plant genomes. Hortic. Res. *9*, uhac017. https://doi.org/10.1093/hr/uhac017.

187. Haas, B.J., Delcher, A.L., Mount, S.M., Wortman, J.R., Smith, R.K., Jr., Hannick, L.I., Maiti, R., Ronning, C.M., Rusch, D.B., Town, C.D., et al. (2003). Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. *31*, 5654–5666. https://doi.org/10.1093/nar/gkg770.

188. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. *20*, 238. https://doi.org/10.1186/s13059-019-1832-y.

189. Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods *18*, 366–368. https://doi.org/10.1038/s41592-021-01101-x.

190. Piñeiro, C., Abuín, J.M., and Pichel, J.C. (2020). Very Fast Tree: speeding up the estimation of phylogenies for large alignments through parallelization and vectorization strategies. Bioinformatics *36*, 4658–4659. https://doi.org/10.1093/bioinformatics/btaa582.

191. McMurdie, P.J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One *8*, e61217. https://doi.org/10.1371/journal.pone.0061217.

192. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y. (2017). Ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. *8*, 28–36. https://doi.org/10.1111/2041-210X.12628.

193. Gearty, W. (2021). Deeptime: plotting tools for anyone working in deep time. R package version 0.0. https://cran.r-project.org/web/packages/deeptime/index.html.

194. Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Mol. Biol. Evol. *34*, 2115–2122. https://doi.org/10.1093/molbev/msx148.

195. Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature *485*, 635–641. https://doi.org/10.1038/nature11119.

196. Alexa, A., and Rahnenfuhrer, J. (2009). Gene set enrichment analysis with topGO. https://bioconductor.statistik.tu-dortmund.de/packages/3.3/bioc/vignettes/topGO/inst/doc/topGO.pdf.

197. Sayols, S. (2023). rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. MicroPubl. Biol. *2023*. https://doi.org/10.17912/micropub.biology.000811.

198. Lovell, J.T., Sreedasyam, A., Schranz, M.E., Wilson, M., Carlson, J.W., Harkess, A., Emms, D., Goodstein, D.M., and Schmutz, J. (2022). GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. eLife *11*, e78526. https://doi.org/10.7554/eLife.78526.

199. Schenk, R.U., and Hildebrandt, A.C. (1972). Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. Can. J. Bot. *50*, 199–204. https://doi.org/10.1139/b72-026.

200. Barbier, F.F., Chabikwa, T.G., Ahsan, M.U., Cook, S.E., Powell, R., Tanurdzic, M., and Beveridge, C.A. (2019). A phenol/chloroform-free method to extract nucleic acids from recalcitrant, woody tropical species for gene expression and sequencing. Plant Methods *15*, 62. https://doi.org/10.1186/s13007-019-0447-3.

201. Lutz, K.A., Wang, W., Zdepski, A., and Michael, T.P. (2011). Isolation and analysis of high quality nuclear DNA with reduced organellar DNA for plant genome sequencing and resequencing. BMC Biotechnol. *11*, 54. https://doi.org/10.1186/1472-6750-11-54.

202. Oberacker, P., Stepper, P., Bond, D.M., Höhn, S., Focken, J., Meyer, V., Schelle, L., Sugrue, V.J., Jeunen, G.-J., Moser, T., et al. (2019). Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. PLoS Biol. *17*, e3000107. https://doi.org/10.1371/journal.pbio.3000107.

203. Schalamun, M., Nagar, R., and Kainer, D. (2018). Harnessing the Min ION: An example of how to establish long-read sequencing in a laboratory using challenging plant tissue from Eucalyptus pauciflora. Mol. Ecol. Resour *19*, 77–89. https://doi.org/10.1111/1755-0998.12938.

204. Zhao, H., Li, H., Jia, Y., Wen, X., Guo, H., Xu, H., and Wang, Y. (2020). Building a Robust Chromatin Immunoprecipitation Method with Substantially Improved Efficiency. Plant Physiol. *183*, 1026–1034. https://doi.org/10.1104/pp.20.00392.

205. Cahn, J., Regulski, M., Lynn, J., Ernst, E., de Santis Alves, C., Ramakrishnan, S., Chougule, K., Wei, S., Lu, Z., Xu, X., et al. (2024). MaizeCODE reveals bi-directionally expressed enhancers that harbor molecular signatures of maize domestication. Nat. Commun. *15*, 10854. https://doi.org/10.1038/s41467-024-55195-w.

206. Schultz, M.D., Schmitz, R.J., and Ecker, J.R. (2012). "Leveling"the playing field for analyses of single-base resolution DNA methylomes. Trends Genet. *28*, 583–585. https://doi.org/10.1016/j.tig.2012.10.012.

207. Van Hoeck, A., Horemans, N., Monsieurs, P., Cao, H.X., Vandenhove, H., and Blust, R. (2015). The first draft genome of the aquatic model plant Lemna minor opens the route for future stress physiology research and biotechnological applications. Biotechnol. Biofuels *8*, 188. https://doi.org/10.1186/s13068-015-0381-1.

208. Sayers, E.W., Beck, J., Bolton, E.E., Bourexis, D., Brister, J.R., Canese, K., Comeau, D.C., Funk, K., Kim, S., Klimke, W., et al. (2021). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. *49*, D10–D17. https://doi.org/10.1093/nar/gkaa892.

209. Mardanov, A.V., Ravin, N.V., Kuznetsov, B.B., Samigullin, T.H., Antonov, A.S., Kolganova, T.V., and Skyabin, K.G. (2008). Complete sequence of the duckweed (Lemna minor) chloroplast genome: structural organization and phylogenetic relationships to other angiosperms. J. Mol. Evol. *66*, 555–564. https://doi.org/10.1007/s00239-008-9091-7.

210. Wang, W., and Lanfear, R. (2019). Long-Reads Reveal That the Chloroplast Genome Exists in Two Distinct Versions in Most Plants. Genome Biol. Evol. *11*, 3372–3381. https://doi.org/10.1093/gbe/evz256.

211. Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., Brover, S., Schoch, C.L., Kimchi, A., and DiCuccio, M. (2018). Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int. J. Syst. Evol. Microbiol. *68*, 2386–2392. https://doi.org/10.1099/ijsem.0.002809.

212. Campbell, M.S., Law, M., Holt, C., Stein, J.C., Moghe, G.D., Hufnagel, D.E., Lei, J., Achawanantakun, R., Jiao, D., Lawrence, C.J., et al. (2014). MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. Plant Physiol. *164*, 513–524. https://doi.org/10.1104/pp.113.230144.

213. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. *9*, R7. https://doi.org/10.1186/gb-2008-9-1-r7.

214. Emms, D.M., and Kelly, S. (2017). STRIDE: Species Tree Root Inference from Gene Duplication Events. Mol. Biol. Evol. *34*, 3267–3278. https://doi.org/10.1093/molbev/msx259.

215. Emms, D.M., and Kelly, S. (2018). STAG: Species Tree Inference from All Genes. Preprint at bioRxiv, 267914. https://doi.org/10.1101/267914.

216. He, J., Alonge, M., Ramakrishnan, S., Benoit, M., Soyk, S., Reem, N.T., Hendelman, A., Van Eck, J., Schatz, M.C., and Lippman, Z.B. (2023). Establishing Physalis as a Solanaceae model system enables genetic reevaluation of the inflated calyx syndrome. Plant Cell *35*, 351–368. https://doi.org/10.1093/plcell/koac305.

217. Galbraith, D.W., Harkins, K.R., Maddox, J.M., Ayres, N.M., Sharma, D.P., and Firoozabady, E. (1983). Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science *220*, 1049–1051. https://doi.org/10.1126/science.220.4601.1049.

218. Appenroth, K.-J., Teller, S., and Horn, M. (1996). Photophysiology of turion formation and germination inSpirodela polyrhiza. Biol. Plant. *38*, 95–106.

219. Hoang, P.T.N., and Schubert, I. (2017). Reconstruction of chromosome rearrangements between the two most ancestral duckweed species Spirodela polyrhiza and S. intermedia. Chromosoma *126*, 729–739. https://doi.org/10.1007/s00412-017-0636-7.

220. Weisshart, K., Fuchs, J., and Schubert, V. (2016). Structured illumination microscopy (SIM) and photoactivated localization microscopy (PALM) to analyze the abundance and distribution of RNA polymerase II molecules on flow-sorted Arabidopsis nuclei. BIO-PROTOCOL *6*, e1725. https://doi.org/10.21769/BioProtoc.1725.

221. Mandáková, T., and Lysak, M.A. (2008). Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). Plant Cell *20*, 2559–2570. https://doi.org/10.1105/tpc.108.062166.

222. Hoang, P.N.T., Michael, T.P., Gilbert, S., Chu, P., Motley, S.T., Appenroth, K.J., Schubert, I., and Lam, E. (2018). Generating a high-confidence reference genome map of the Greater Duckweed by integration of cytogenomic, optical mapping, and Oxford Nanopore technologies. Plant J. *96*, 670–684. https://doi.org/10.1111/tpj.14049.

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-LjCENH3	GenScript, this study	N/A
Rabbit polyclonal anti-H3	Sigma-Aldrich	Cat# H0164; RRID: AB_532248
Rabbit polyclonal anti-H3K4me3	Millipore	Cat# 07-473; RRID: AB_1977252
Chemicals, peptides, and recombinant proteins		
Lemna japonica CENH3 N-terminal synthetic peptide (sequence: RTKHISGKRRRTE)	GenScript, this paper	N/A
Critical commercial assays		
ChIP DNA clean and concentrator kit	Zymo Research	Cat# D5205
NEBNext Ultra II DNA Library Prep Kit for Illumina	New England Biolabs	Cat# E7645S
Ligation Sequencing Kit SQK-LSK108	Oxford Nanopore Technologies	Cat# SQK-LSK108
Ligation Sequencing Kit SQK-LSK109	Oxford Nanopore Technologies	Cat# SQK-LSK109
Rapid Barcoding Kit SQK-RBK004	Oxford Nanopore Technologies	Cat# SQK-RBK004
Illumina TruSeq DNA PCR-Free Kit	Illumina	Cat# 20015962
NEBNext Ultra II DNA Library Prep Kit for Illumina	New England Biolabs	Cat# E7645L
NEBNext Multiplex Oligos for Illumina	New England Biolabs	Cat# E6609S
Dynabeads™ Protein A for Immunoprecipitation	Invitrogen	Cat# 10001D
Proximo Hi-C (Plant) Kit	Phase Genomics	Cat# KT3040
EZ DNA Methylation-Gold	Zymo Research	Cat# D5005
NEBNext® Multiplex Oligos for Illumina® (Methylated Adaptor, Index Primers Set 1)	New England Biolabs	Cat# E7535
AMPure XP Reagent	Beckman Coulter	Cat# A63880
EZ DNA Methylation-Lightning Kit	Zymo Research	Cat# D5030
KAPA HiFi Uracil+ Kit	Roche	Cat# KK2801
DNeasy Plant Mini Kit	Qiagen	Cat# 69104
Quick-DNA/RNA Miniprep Kit	Zymo Research	Cat# D7001
RNA Clean & Concentrator kit	Zymo Research	Cat# R1017
NEXTflex Small RNA-Seq Kit v3	Bioo Scientific	Cat# 5132-06
Dynabeads mRNA Purification Kit	Invitrogen	Cat# 61006
Ribo-Zero rRNA Removal Kit (Plant Leaf)	Illumina	Cat# MRZPL116
ScriptSeq v2 RNA-Seq Library Preparation Kit	Illumina	Cat# SSV21106
PCR cDNA Sequencing Kit	Oxford Nanopore Technologies	Cat# SQK-PCS108
Deposited data		
Genome assemblies	This paper	BioProject: PRJNA999459
Raw and processed data	This paper	GEO: GSE238136
Raw data	Michael et al.[30]	SRA: SRX8008794
Raw data	Michael et al.[30]	SRA: SRX8008795
Raw and processed data	This paper	GEO: GSE238136
Wa7733 chloroplast genome	Wang et al.[131]	GenBank: JN160605.1
Phytozome 13 reference proteomes	Goodstein et al.[132]	https://phytozome-next.jgi.doe.gov

(*Continued on next page*)

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NCBI Elaeis guineensis Annotation Release 102	Orion Genomics/NCBI RefSeq	https://www.ncbi.nlm.nih.gov/refseq/annotation_euk/Elaeis_guineensis/102
REXdb: a reference database of transposable element protein domains	Neumann et al.[133]	http://repeatexplorer.org
miRBase 22.1	Kozomara and Griffiths-Jones[134]	https://www.mirbase.org
OrthoDB v10	Kriventseva et al.[135]	https://v10-1.orthodb.org
Rfam 14.1	Kalvari et al.[136]	https://rfam.org
Experimental models: Organisms/strains		
Spirodela polyrhiza 9509	Rutgers Duckweed Stock Cooperative	9509
Lemna gibba 7742a	Rutgers Duckweed Stock Cooperative	7742a
Lemna minor 7210	Rutgers Duckweed Stock Cooperative	7210
Lemna minor 9252	Rutgers Duckweed Stock Cooperative	9252
Lemna japonica 7182	Rutgers Duckweed Stock Cooperative	7182
Lemna japonica 8627	Rutgers Duckweed Stock Cooperative	8627
Lemna japonica 9421	Rutgers Duckweed Stock Cooperative	9421
Lemna turionifera 8133	Rutgers Duckweed Stock Cooperative	8133
Lemna turionifera 9434	Rutgers Duckweed Stock Cooperative	9434
Wolffia australiana 8730	Rutgers Duckweed Stock Cooperative	8730
Software and algorithms		
Guppy v5.0.7	Oxford Nanopore Technologies	https://nanoporetech.com/software/other/guppy
bwa-mem2 v2.2.1	Vasimuddin et al.[137]	https://github.com/bwa-mem2/bwa-mem2
deeptools	Ramírez et al.[138]	https://github.com/deeptools/deeptools
MACS2	Zhang et al.[139]	https://github.com/macs3-project/MACS/tree/macs_v2
bedtools	Quinlan and Hall[140]	https://github.com/arq5x/bedtools2
DANTE v0.2.5	Novák et al.[31]	https://github.com/kavonrtep/dante
Juicer pipeline	Durand et al.[141]	https://github.com/aidenlab/juicer
hic-straw v0.0.8	Durand et al.[141]	https://github.com/aidenlab/straw
Cooler v0.9.2	Abdennur and Mirny[142]	https://github.com/open2c/cooler
HiGlass v1.11	Kerpedjiev et al.[143]	https://github.com/higlass/higlass
minimap2	Li[144]	https://github.com/lh3/minimap2
Clair3 v0.1-r12	Zheng et al.[145]	https://github.com/HKU-BAL/Clair3
Sniffles2 v2.0.7	Smolka et al.[146]	https://github.com/fritzsedlazeck/Sniffles
RTG Tools v3.12.1	Real Time Genomics	https://github.com/RealTimeGenomics/rtg-tools
Trimmomatic v0.35	Bolger et al.[147]	https://github.com/timflutre/trimmomatic
Bismark v0.23.1	Krueger and Andrews[148]	https://github.com/FelixKrueger/Bismark
Megalodon v2.5.0	Oxford Nanopore Technologies	https://github.com/nanoporetech/megalodon
skewer v0.2.2	Jiang et al.[149]	https://github.com/relipmoc/skewer
fastp v0.20.1	Chen et al.[150]	https://github.com/OpenGene/fastp
ShortStack v4.0.0	Johnson et al.[151]	https://github.com/MikeAxtell/ShortStack
Flye v2.8.3	Kolmogorov et al.[152]	https://github.com/mikolmogorov/Flye
PEPPER-Margin-DeepVariant v0.4	Shafin et al.[153]	https://github.com/kishwarshafin/pepper/tree/r0.4
NextPolish v1.3.1	Hu et al.[154]	https://github.com/Nextomics/NextPolish
purge_dups v1.2.5	Guan et al.[155]	https://github.com/dfguan/purge_dups
BLAST+ v2.11.0	Zhang et al.[156]	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+

(*Continued on next page*)

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
3D-DNA	Dudchenko et al.[157]	https://github.com/aidenlab/3d-dna
JBAT	Durand et al.[158]	https://github.com/aidenlab/Juicebox
RagTag v2.0.1	Alonge et al.[159]	https://github.com/malonge/RagTag
Hapo-G v1.2	Aury and Istace[160]	https://github.com/institut-de-genomique/HAPO-G
SeqKit	Shen et al.[161]	https://github.com/shenwei356/seqkit
nucmer v3.1	Kurtz et al.[162]	https://github.com/garviz/MUMmer
dotPlotly	Poorten[163]	https://github.com/tpoorten/dotPlotly
GeSeq v2.03	Tillich et al.[164]	https://chlorobox.mpimp-golm.mpg.de/geseq.html
EDTA v1.9.6	Ou et al.[165]	https://github.com/oushujun/EDTA
Tandem Repeats Finder v4.09	Benson[166]	https://github.com/Benson-Genomics-Lab/TRF
Infernal v1.1.4	Nawrocki and Eddy[167]	https://github.com/EddyRivasLab/infernal
HISAT2 v2.2.1	Kim et al.[168]	https://github.com/DaehwanKimLab/hisat2
BRAKER v2.1.6	Brůna et al.[169]	https://github.com/Gaius-Augustus/BRAKER
TSEBRA v1.0.3	Gabriel et al.[170]	https://github.com/Gaius-Augustus/TSEBRA
STAR v2.7.9a	Dobin et al.[171]	https://github.com/alexdobin/STAR
PsiCLASS v1.0.2	Song et al.[172]	https://github.com/splicebox/PsiCLASS
StringTie v2.2.0	Kovaka et al.[173]	https://github.com/gpertea/stringtie
uLTRA v0.0.4	Sahlin and Mäkinen[174]	https://github.com/ksahlin/ultra
Portcullis v1.2.0	Mapleson et al.[175]	https://github.com/EI-CoreBioinformatics/portcullis
Mikado v2.3.2	Venturini et al.[176]	https://github.com/EI-CoreBioinformatics/mikado
TransDecoder v5.5.0	Haas and Papanicolaou[177]	https://github.com/TransDecoder/TransDecoder
GeMoMa v1.8	Keilwagen et al.[178]	https://www.jstacs.de/index.php/GeMoMa
BUSCO v5.1.3	Manni et al.[179]	https://gitlab.com/ezlab/busco
MAFFT v7.487	Katoh and Standley[180]	https://mafft.cbrc.jp/alignment/software/
IQ-TREE v2.1.4	Minh et al.[181]	https://github.com/iqtree/iqtree2
Cactus v2.0.4	Armstrong et al.[182]	https://github.com/ComparativeGenomicsToolkit/cactus
Augustus-CGP v3.4.0	König et al.[183]	https://github.com/Gaius-Augustus/Augustus
AGAT	Dainat et al.[184]	https://github.com/NBISweden/AGAT
MAKER v3.01.04	Holt and Yandell[185]	https://github.com/Yandell-Lab/maker
TEsorter v1.3.0	Zhang et al.[186]	https://github.com/zhangrengang/TEsorter
PASA v2.5.1	Haas et al.[187]	https://github.com/PASApipeline/PASApipeline
OrthoFinder2 v2.5.4	Emms and Kelly[188]	https://github.com/davidemms/OrthoFinder
diamond	Buchfink et al.[189]	https://github.com/bbuchfink/diamond
VeryFastTree v3.1.0	Piñeiro et al.[190]	https://github.com/citiususc/veryfasttree
phyloseq	McMurdie and Holmes[191]	https://github.com/joey711/phyloseq
ggtree	Yu et al.[192]	https://github.com/YuLab-SMU/ggtree
deeptime	Gearty[193]	https://github.com/willgearty/deeptime

(*Continued on next page*)

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
eggnog-mapper v2.1.6	Huerta-Cepas et al.[194]	https://github.com/eggnogdb/eggnog-mapper
AHRD v3.11	Sato et al.[195]	https://github.com/groupschoof/AHRD
TopGO	Alex and Rahnenfuhrer[196]	https://bioconductor.org/packages/topGO/
rrvgo	Sayols[197]	https://github.com/ssayols/rrvgo
GENESPACE v1.1.4	Lovell et al.[198]	https://github.com/jtlovell/GENESPACE

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Plant stocks and growth conditions

Sterile cultures of the 9 accessions in Figure 1 were obtained from the Rutgers Duckweed Stock Cooperative and cultivated in 50 mL Schenk and Hildebrandt (SH) medium with 1% sucrose at pH 5.6[199] at 23°C under a 16 hour photoperiod of approximately 30 μmol/m2/s per second of white fluorescent light. For HMW DNA extraction, culture flasks were covered in foil and grown in the dark for up to 2 weeks prior to harvest to deplete excess carbohydrates before flash-freezing in liquid N_2 and storage at -80°C.

METHOD DETAILS

HMW DNA extraction

High molecular weight (HMW) DNA extractions were performed for Lg7742a, Lj8627, and Wa8730 using a modified CTAB prep followed by a high-salt low-ethanol starch cleanup (Pacific Biosciences) as described previously,[84] with the omission of the sorbitol wash. Four grams of dark-treated frozen duckweed tissue was ground under LN_2, transferred to 20 ml pre-warmed lysis buffer (100 mM Tris-HCl (pH 9.0), 2% w/v CTAB, 1.4 M NaCl, 20 mM EDTA, 2% PVP-10, 1% 2-mercaptoethanol, 0.1% sarkosyl and 100 μg ml−1 proteinase K), pulse-vortexed briefly x3, and incubated for 1 h at 65 °C. Serial organic extractions were carried out in phase-lock gel tubes using 1 vol phenol:chloroform:isoamyl alcohol (25:24:1), followed by 1 vol chloroform:isoamyl alcohol. DNA was precipitated by adding 0.1 vol 3 M NaOAc (pH 5.2) followed by 0.7 vol isopropanol. HMW DNA was hooked out with a pasteur pipette and washed with 70% EtOH, air dried for 2 min and resuspended in 200 μl Tris-HCl (pH 8.5; EB). The resuspension was treated with 2 μl 20 mg ml−1 RNase A at 37 °C for 20 min followed by 2 μl 50 mg ml−1 proteinase K at 50 °C for 30 min. 194 μl EB, 100 μl NaCl and 2 μl 0.5 M EDTA were added, and organic extractions were performed as before. DNA was precipitated with 1.7 vol EtOH, hooked out of solution with a pasteur pipette, washed with 70% EtOH and resuspended in 50 μl EB. HMW DNA was isolated for Lm9252 and Lj7182 using a CTAB/PVP protocol as previously described.[200] Frozen tissue was ground under LN_2, and lysed in a 2% CTAB buffer (CTAB 2%, NaCl 1.4 M, EDTA 20 mM, Tris–HCl 100 mM, PVP40 2%) with RNase A for 1 h at 37 °C with periodic vortexing. SDS 10% was added to a final concentration of 0.7% before thorough vortexing. The solution was centrifuged (20,000 x g, 15 min, RT), and the liquid phase was transferred to a new tube. DNA was precipitated with 1 x vol isopropanol at -20 °C, centrifuged (20,000 x g, 1 h, 4 °C), and washed with 70% ethanol. DNA pellets were dried for 10 min and resuspended in RNase-free water. For Lj9421 and Lt9434, a modified nuclear extraction was used as previously described.[201] Ten grams of flash-frozen duckweed tissue was ground under LN_2 and transferred to SEB extraction buffer, placed on ice for 30 min, and filtered through two layers of cheesecloth. Triton-X was added drop-by-drop to a final concentration of 0.5%, and after 10 minutes resting on ice, centrifuged (650 x g, 15 min, 4 °C). The nuclei pellet was washed once in SEB, centrifuged again, and resuspended in TE buffer. Resuspended nuclei were treated with RNase A and digested with Proteinase K, and then precipitated with 1 vol isopropanol. For Lm7210, frozen tissue was ground under LN_2 and HMW DNA was isolated using a modified Bomb protocol.[202]

Long read whole genome sequencing

Long-read sequencing libraries for Lg7742a, Lj8627, and Wa8730 were prepared as follows: HMW DNA was gently sheared by 20x passage through a P1000 pipette. Ligation Sequencing gDNA kits (Oxford Nanopore Technologies Cat#SQK-LSK108, Cat#SQK-LSK109) were performed as previously described[84] following manufacturer's instructions with modifications: (1) DNA repair, end-prep and ligation reactions were extended to 20 min; (2) 0.8× vol of a custom SPRI bead solution was substituted for all cleanups[203]; (3) bead elution was carried out at 50 °C for 5 min. Completed libraries were loaded onto R9 or R9.4.1 flow cells and sequenced on the MinION instrument. Libraries for Lj7182, Lj9421, Lm7210, Lm9252, and Lt9434 were prepared from 1.5 μg of unsheared HMW DNA using the Rapid Barcoding Sequencing Kit (Oxford Nanopore Technologies Cat#SQK-RBK004) following manufacturer's instructions in the "Rapid sequencing gDNA" protocol. The resulting libraries were sequenced on R9.4.1 flow cells on the MinION and PromethION platforms. Offline base calling of all ONT reads was performed with Guppy v5.0.7 (Oxford Nanopore Technologies) and the R9.4.1 450bps SUP model on NVIDIA Tesla V100 GPUs.

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

Short read whole genome sequencing

Short read WGS libraries for Lg7742a and Lj8627 were prepared from 2μg of HMW gDNA using the Illumina TruSeq DNA PCR-Free kit (Illumina Cat#20015962) and sequenced on an Illumina MiSeq (PE 250bp, PE 300bp) or HiSeq 2500 instrument (PE 150bp). Libraries for other *Lemna* accessions were prepared Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs Cat#E7645L) and indexed with NEBNext Multiplex Oligos for Illumina (New England Biolabs Cat#E6609S). Upstream of library preparation, samples were fragmented via Covaris mechanical shearing. Previously published libraries were used for Wa8730 (SRA accessions SRX8008794, SRX8008795).[30] Illumina gDNA reads were aligned to the reference assemblies with bwa-mem2 v2.2.1[137] and coverage was calculated over 1bp bins with deeptools v3.5.2[138] "bamCoverage –samFlagExclude 2304 –ignoreDuplicates –binSize 1 –normalizeUsing CPM".

Chromatin immunoprecipitation (ChIP)

CENH3 orthologues in *S. polyrhiza*, *W. australiana*, *L. gibba*, and *L. japonica* were identified by sequence similarity search with *O. sativa* and *A. thaliana* sequences. A rabbit polyclonal antibody was raised against the synthesized polypeptide sequence "RTKHISGKRRRTE" unique to the *L. minor* and *L. turionifera* CENH3 N-terminal tail (LjCENH3), and then validated by Western blot. Chromatin extraction and immunoprecipitation was performed as previously described[72] with minor modifications. Briefly, two biological replicates of 750 mg whole frond tissue per clone were crosslinked in 1% formaldehyde for 10 min, and ground to a fine powder under liquid nitrogen. Nuclei were extracted in an SDS-Tris buffer, lysed, and sonicated with a Bioruptor Pico (Diagenode) for 12 cycles at 30 Hz, 30 s on, 30 s off. Chromatin was cleared with Protein A magnetic beads (Invitrogen Cat#10001D) and incubated overnight with following antibodies: anti-LjCENH3 (GenScript, this study), anti-H3 (Sigma-Aldrich Cat#H0164, RRID:AB_532248), and anti-H3K4me3 (Millipore Cat#07-473, RRID:AB_1977252). Immune complexes were eluted, and crosslinks were reversed with Proteinase K digestion for 2 hrs at 55C with gentle shaking. Overnight NaCl incubation was omitted as in.[204] DNA was purified with the ChIP Purification Kit (Zymo Cat#D5205), and the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs Cat#E7645S) was used to construct sequencing libraries. *L. turionifera* samples in these libraries were from the clone 8133, rather than 9434. Initial data analysis including adapter trimming, alignment, and peak calling, was carried out as described in [205]. After visualizing CENH3/H3 enrichment, centromere regions were computationally estimated as follows: 1) $\log_2$(CENH3/H3) was computed across 10 Kbp genome bins, smoothed by 30 Kbp with deepTools v3.5.5[138] bamCompare; 2) The bin with the highest enrichment per chromosome was selected, and the intersecting MACS2[139] broad peak call with largest ($\log_2$(CENH3/H3) * peak_width) as computed by the MaizeCODE pipeline was chosen as the primary peak with bedtools v2.31.1[140] intersect and awk; 3) all MACS2 peaks within 5 Kbp of each other were merged, and the merged peaks overlapping the primary peak per chromosome were taken as the predicted centromere regions. Predicted transposable elements overlapping these regions were annotated with DANTE v0.2.5.[31]

Chromatin conformation capture (Hi-C)

For Lg7742a, Lj8627 and Wa8730, approximately 10g of tissue per sample was sent to Dovetail Genomics and Hi-C libraries were prepared using their in-house protocol with the DpnII enzyme. For Lm7210 and Lt9434, the Proximo Hi-C (Plant) Kit (Phase Genomics Cat#KT3040) was used to prepare libraries.[40] The resulting libraries were sequenced on an Illumina NextSeq 500 (PE 150bp). The Juicer pipeline v2.20[141] with options "-s DpnII" was used to align reads back to the final assemblies and construct initial contact maps (MAPQ $\geq$ 30) before multi-resolution cool file conversion with hic-straw v0.0.8,[141] ICE balancing with Cooler v0.9.2,[142] and visualization with HiGlass v1.11.[143]

Heterozygous variant calling

ONT reads were aligned to their respective reference assemblies with minimap2 v2.24-r1122[144] "-x map-ont –MD". Single nucleotide variants and indels shorter than 50 bp (SNVs) were called using Clair3 v0.1-r12[145] with options "–min_contig_size=1000000 –platform ont –model_path /opt/models/r941_prom_sup_g5014". Structural variants (SVs) were called with Sniffles2 v2.0.7[146] with default options. SNVs and SVs were filtered using RTG Tools v3.12.1 (https://github.com/RealTimeGenomics/rtg-tools) with options "vcffilter –min-read-depth=10 –min-genotype-quality=20", and homozygous calls were removed with "vcffilter –remove-hom".

Whole genome bisulfite sequencing (WGBS)

WGBS libraries were prepared for Sp9509 and Wa8730 as previously described.[19] Briefly, 5ug genomic DNA was treated with EZ DNA Methylation-Gold (Zymo Research Cat#D5005) and libraries were prepared from converted and unconverted samples with the EpiGnome Methyl-Seq Kit (Illumina Cat#EGMK81312). Lg7742a and Lj8627 libraries were prepared as previously described[72] for *Arabidopsis* embryos, except that DNA from the HMW extraction methods described above was used. DNA (150 ng) was sheared on a Bioruptor device, ligated to NEBNext methylated adapters (New England Biolabs Cat#E7535), purified with AMPure XP beads (Beckman Coulter Cat#A63880), and treated with the EZ DNA Methylation-Lightning Kit (Zymo Research Cat#D5030). Converted DNA was eluted with 20 μL of buffer, and 10 μL was amplified with KAPA HiFi uracil-tolerant polymerase (Roche Cat#KK2801). Libraries were sequenced on a NextSeq 500 (PE 151) or a HiSeq 2500 (PE 108). Adapter sequences were removed and reads were hard-trimmed with Trimmomatic v0.35[147] with options "HEADCROP:5 TRAILING:3 MINLEN:25". Technical and biological replicates were merged, and reads were aligned to the genomes using Bismark v0.23.1[148] with options "-N 1 -L 20 –maxins 1200". Reads were

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

deduplicated and methylated cytosines were called using "bismark_methylation_extractor –CX –bedGraph –ignore_r2 2 –comprehensive" and a genome-wide cytosine report was generated with coverage2cytosine, and separate bigWig coverage tracks were derived for CpG, CHG, and CHH contexts. Profiles over genomic regions were calculated with deepTools 3.5.1[138] computeMatrix with options "scale-regions –skipZeros -bs 100 -m 2000 -b 2000 -a 2000" and plotted in R. Genome-wide methylation levels were determined by calculating the weighted methylation level (#C/(#C+#T))[206] for nuclear genome cytosine positions with a coverage of at least 5 reads. For reference, approximate genome-wide methylation levels for *Z. mays* B73 were derived from.[75]

Long read direct methylation analysis

Direct 5-methylcytosine modification calling in all contexts from the Oxford Nanopore WGS reads was performed with Megalodon v2.5.0 (https://github.com/nanoporetech/megalodon) with the dna_r9.4.1_450bps_modbases_5mc_hac model and Guppy v5.0.7 on NVIDIA Tesla V100 GPUs. Calls for Cs with fewer than 5 supporting reads were discarded, and the resulting bedMethyl files were split by cytosine context (CpG, CHG, CHH) using bedtools v2.30.0[140] intersect. Fractional methylation calls at each cytosine were adjusted to a [0..1] scale, and the bedMethyl files were converted to bigWig format. Profiles over genomic regions were calculated as for the WGBS libraries.

Transcriptome sequencing

To provide broad transcriptional evidence for annotation, RNA samples were collected from fronds of Lg7742a and Lj8627 grown under a diverse set of conditions: variable daylength, nutrient stress, temperature stress, high NaCl, high pH, UV damage, and exogenous hormone exposure. Samples from all conditions were pooled, polyA enriched with Dynabeads mRNA Purification Kit (Invitrogen Cat#61006) or rRNA depleted with the Ribo-Zero rRNA Removal Kit (Plant Leaf) (Illumina Cat#MRZPL116). Strand-specific cDNA libraries were prepared for Illumina sequencing using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Illumina Cat#SSV21106), and for nanopore sequencing using the PCR cDNA Sequencing Kit (Oxford Nanopore Technologies Cat #SQK-PCS108) following manufacturer instructions.

Small RNA sequencing and analysis

RNA was extracted from 3 biological replicates of 100mg of frozen tissue of each accession using the Quick-DNA/RNA Miniprep kit (Zymo Research Cat#D7001) following manufacturer's instructions with the following modifications: frozen tissue was ground under LN_2 with a mortar and pestle and resuspended in the Shield solution. It was refrozen in LN_2, thawed, and treated with Proteinase K, and frozen again after the addition of lysis buffer. After thawing, samples were centrifuged at full speed, RT for 3 min. to remove debris, and the manufacturer's protocol was followed as described. Afterwards, samples were DNase treated, enriched for RNAs 17-200 nt in length, and concentrated using the RNA Clean & Concentrator kit (Zymo Research Cat#R1017). Libraries were prepared with the NEXTflex Small RNA-Seq Kit v3 (Bioo Scientific Cat#5132-06) following manufacturer instructions and sequenced on a NextSeq 500 (SE 75). Sequencing adapters were trimmed using skewer v0.2.2[149] and 4nt randomer sequences were removed from both ends of each read with fastp v0.20.1[150] with options "–trim_front1 4 –trim_tail1 4 –length_required 10 –length_limit 35 –disable_adapter_trimming –trim_poly_g -q 20 –unqualified_percent_limit 10". Reads mapping perfectly to annotated structural RNAs and organelles were removed, and those remaining were aligned to the genomes with ShortStack v4.0.0[151] using options "–mincov 0.5 –mmap f –dn_mirna –knownRNAs <*miRBase 22.1 mature plant sequences*>". Separate ShortStack runs with merged biological replicates were performed for microRNA discovery, and microRNAs identified by ShortStack were annotated by top BLAST hit of the mature sequence to miRBase 22.1.[134] Alignments were split by size class and bigWig coverage tracks computed with deepTools "bamCoverage –binSize 1 –normalizeUsing CPM". Genomic region profiles were computed as described for the bisulfite libraries and plotted in R.

Genome assembly

Reads longer than 1 Kbp were assembled into contigs with Flye v2.8.3-b1722[152] with options "–extra-params max_bubble_length=2000000 -m 20000 –plasmids -t 48 –nano-raw". The same reads were then aligned to the assembly using minimap2 2.20-r1061,[144] and these alignments were passed to the PEPPER-Margin-DeepVariant v0.4 pipeline[153] to polish the initial consensus with default options. To correct remaining SNVs and small indels, Illumina gDNA libraries were mapped to the long read polished consensus with bwa-mem2 2.2.1[137] for further polishing with NextPolish v1.3.1.[154] To reduce occurrences of uncollapsed haplotigs and heterozygous overlaps in the assemblies, purge_dups v1.2.5[155] was run with options "-a 80 -2". For the hybrids, contigs were first assigned to either the *L. minor* or *L. turionifera* subgenomes by performing an initial pseudomolecule scaffolding with Hi-C reads as described below, followed by sequence similarity ranking of pseudomolecules using MegaBLAST v2.11.0+[156] alignment of *L. minor* 5500 contigs[207] against the target scaffolds. Target contigs comprising each parental pseudomolecule set were then independently treated with purge_dups. Next, Hi-C reads (PE150) were mapped to the polished, heterozygosity-purged contigs with the Juicer pipeline v1.6[141] UGER scripts with options "-s DpnII". The resulting "merged_nodups.txt" alignments were passed to the 3D-DNA pipeline[157] to iteratively order and orient the input contigs and correct misjoins. The initial automatic scaffolding was followed by manual review with JBAT.[158] No Hi-C data were available for accessions Lj7182, Lj9421, and Lj9252, and instead pseudomolecules were constructed using RagTag v2.0.1[159] correct and scaffold steps with default options and the final Lj8627 assembly as a reference. For all accessions, A final haplotype-aware short read polishing step was performed with Hapo-G v1.2[160] using default options. Assembled pseudomolecules for Lm7210 and Wa8730 were named chr{1..N} according to length. All other *Lemna*

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

pseudomolecules were oriented and named according to homology with Lm7210 chromosomes. Unintegrated contigs were screened for viral and bacterial contaminants by MegaBLAST search against assembled target accession pseudomolecules, organelles, and the NCBI nt[208] databases simultaneously.

Organelle genome assembly and annotation

Reference *Lemna minor* and *Wolffia australiana* plastid (CP) and *Spirodela polyrhiza* mitochondrial (MT) genomes were downloaded from NCBI RefSeq (GenBank: NC_010109.1, NC_015899.1, NC_017840.1)[39,131,209], and 180° rotations were generated using SeqKit 2.2.0[161] sliding "-C <reference> -s *<ref length / 2>* -W *<ref length>*". ONT reads at least 40 Kbp long were aligned with minimap2 2.22[144] to the original and rotated versions of CP and MT references simultaneously. Reads aligning to each reference were extracted for separate *de novo* assembly with Flye 2.8.3-b1722[152] with options "-m 20000 –asm-coverage 100 –nano-raw" (CP genomes) or "-m 20000 –meta –nano-raw" (MT genomes). A single CP contig of length ±10 Kbp relative to the reference was assembled for all accessions, with ~700-800x downsampled coverage. In the MT case where numerous contigs > 150 Kbp were assembled for each accession, only the contig with the highest coverage (~13x-104x) was retained. PEPPER-Margin-DeepVariant and then NextPolish with downsampled short reads were used to polish the assemblies, which were then manually oriented and rotated for consistency with psbA(-), ndhF(-), ccsA(+).[210] Assemblies were compared by calculating average nucleotide identities (ANI)[211] and constructing dot-plots from pairwise genome alignments using nucmer v3.1[162] and an R script from the dotPlotly package[163] (https://github.com/tpoorten/dotPlotly/blob/master/mummerCoordsDotPlotly.R).

Organelle genomes were annotated using the web application GeSeq v2.03.[164] The following non-default settings were used to annotate plastomes: circular; Sequence source = Plastid (land plants)/Mitochondrial; Annotation revision = Keep all annotations; HMMER profile search: [enabled]; tRNAscan-SE v2.0.7: [enabled]; Chloe v0.1.0: Annotate = CDS+tRNA+rRNA. Inverted repeats were annotated using self-pairwise genome alignment with nucmer. Gene annotations were manually reviewed. blatN annotations were used for rRNA genes and Chloe annotations were selected for tRNA genes. For protein-coding genes, Chloe annotations were used for most protein-coding genes but in cases where Chloe annotations did not result in correct open reading frames (proper start and end codon), then blatX or HMMER annotations with the correct open reading frames were selected. To ensure the correct annotation of the Wa8730 plastome, the Wa7733 plastome assembly (GenBank: JN160605.1) was downloaded, annotated as above, and compared to the Wa8730 reference.

Repetitive and non-coding sequence annotation

De novo repeat libraries were constructed for each accession using EDTA v1.9.6[165] with options "–anno 1 –cds <Sp7498 CDS sequences> –sensitive 1". A softmasked version of each assembly was generated with the EDTA make_masked.pl script with options "-minlength 80". Tandem repeats were identified in the genome assemblies using Tandem Repeats Finder v4.09[166] with options "1 1 2 80 5 200 2000 -d -h". The resulting *dat files were reformed and each repeat length was summarized to identify putative centromere and telomere arrays as described previously.[30] Repeats of a specific length were summed and plotted as a function of repeat length revealing potential centromere arrays (Figure 3A). Ribosomal DNA loci were identified (along with other conserved non-coding genes) using Infernal v1.1.4[167] cmscan with Rfam 14.1[136] with options "-Z <effective genome size> –cut_ga –rfam –nohmmonly". Lower-scoring overlapping hits were removed. Exact occurrences of telomere sequence were identified on both strands using SeqKit v2.2.0[161] locate with options "–bed –ignore-case -p TTTAGGG".

Gene prediction and annotation

Protein coding gene annotation for each accession was performed with a combination of *ab initio* gene prediction, RNA transcript evidence, homologous protein evidence, and comparative gene prediction approaches. Softmasked versions of the assemblies were used for all steps.

First, all available short-read (SR) cDNA evidence was aligned with HISAT2 v2.2.1[168] with options "–dta –max-intronlen 10000 –rna-strandness <orientation>". An initial *ab initio* prediction set was built with BRAKER v2.1.6[169] and TSEBRA v1.0.3,[170] incorporating the HISAT2 alignments and plant protein evidence from OrthoDB v10[135] (*ab_initio_preds*).

Alternative SR alignments were made with STAR v2.7.9a[171] with no reference annotation with options "–twopassMode Basic –alignIntronMin 20 –alignIntronMax 10000 –alignMatesGapMax 10000 –outFilterMismatchNmax 4 –outFilterIntronMotifs RemoveNoncanonical –outSAMstrandField intronMotif –outFilterMultimapNmax 50". HISAT2 and STAR alignments of SR cDNA libraries were assembled independently with PsiCLASS v1.0.2[172] with default options, and StringTie v2.2.0[173] with options "-G <ab_initio_preds> –conservative <–rf or –fr>". Long-read (LR) cDNA libraries were aligned with uLTRA v0.0.4[174] with parameters "–ont or –isoseq, –max_intron 10000, –use_NAM_seeds" and exon hints from *ab_initio_preds*. LR cDNA alignments were cleaned and collapsed with StringTie with options "-G <ab_initio_preds>, -L -R". A set of high-confidence splice junctions was selected by running Portcullis v1.2.0[175] with default options on combined HISAT2 and STAR alignments. Mikado v2.3.2[176] was used to generate an annotation set from only transcript evidence with options "config –mode permissive –scoring plant.yaml –copy-scoring plant.yaml –junctions *<portcullis junctions>*", TransDecoder v5.5.0[177] for ORF prediction, and "serialize –no-start-adjustment" (*rna_preds*).

Due to the scarcity of duckweed RNA-seq data from varied tissues, developmental stages, and growth conditions, we anticipated that a protein homology-based annotation approach would recover more accurate gene models for a large number of genes. The GeMoMa v1.8[178] pipeline was used for this purpose. Reference genomes and proteomes from Phytozome 13[132] were gathered

for *A. comosus*, *A. thaliana*, *A. trichopoda*, *B. distachyon*, *N. colorata*, *O. sativa*, and *Z. marina*, and from NCBI RefSeq for *E. guineensis*. Independent GeMoMaPipeline runs were performed for each reference with default parameters except "AnnotationFinalizer.r=NO", and then combined with GAF with the default of 10 maximum predictions per locus (*protein_preds*).

Common single-copy BUSCO v5.1.3[179] sequences were determined for all novel assemblies in this study, Sp9509, and Phytozome 13 assemblies for *A. americanus*, *A. comosus*, *Z. marina* and *S. polyrhiza* 7498. MAFFT v7.487[180] "–auto" was used to build protein MSAs for each BUSCO across all accessions, which were then concatenated. IQ-TREE v2.1.4[181] with options "-B 1000 –mset LG,WAG,JTT" was used to build a guide tree, and a multiple whole-genome alignment was constructed with Cactus v2.0.4.[182] Using this alignment, a combined hints file from *ab_initio_preds*, and Lm7210 as the reference, Augustus-CGP v3.4.0[183] with options "softmasking=1 –allow_hinted_splicesites=gcag,atac" was used to generate comparative gene predictions (*cgp_preds*).

To privilege the transfer of annotations from reference proteomes, while also ensuring the retention of novel gene loci predicted by other methods, the agat_sp_complement_annotations.pl script from AGAT[184] was run with *protein_preds* as the reference, filling in predictions at unannotated loci first with *rna_preds* and then with *cgp_preds*. MAKER v3.01.04[185] was then used to evaluate the evidence supporting these complemented gene models (model_gff) both from transcript assemblies (*rna_preds*) and protein homology (*protein_preds*). A MAKER-P "standard" build was created as described in [212] to retain only models with evidence support or a Pfam domain. These models were further filtered to remove likely transposon sequences by screening with TEsorter v1.3.0[186] against the REXdb plant TE database.[133] If at least 25% of the amino acid sequence of any gene prediction was covered by transposase matches, and supported by fewer than 2 reference proteomes in the GeMoMa annotation, it was removed. Finally, the PASA v2.5.1 pipeline[187] was used as previously described[213] to update the gene models, using SR and LR transcript assemblies to add UTRs and alternative splice forms (*final_preds*). A subsequent round of filtering out TE and organelle-derived gene models was carried out by DC-MegaBLAST v2.13.0+ against the accession-specific organellar gene CDS and EDTA TE libraries. If more than 50% of the *final_preds* CDS sequence was covered in the top hit to either database, that gene model was removed (*final_preds_filt*).

Phylogenetic analysis

We used OrthoFinder2 v2.5.4[188,214,215] to infer a species tree and phylogenetic relationships among the reference and novel proteomes presented in this study (Figure 4A). A complete listing and details are provided in Data S2. *G. montanum* was used as the outgroup. The proteomes of each subgenome of the *L. japonica* hybrids were treated separately for this analysis. All vs. all alignments were computed with diamond[189] "–iterate –ultra-sensitive -e 0.001". OrthoFinder was run in MSA mode "-M msa" using MAFFT v7.487[180] "–localpair –maxiterate 1000" for alignments with fewer than 1,000 sequences, and default options otherwise. Trees were constructed with either IQ-TREE v2.1.4[181] or VeryFastTree v3.1.0[190] conditionally as follows: (species tree) "iqtree2 –alrt 1000 -T 48 -m MFP –mset Q.plant,LG,WAG,JTT"; (> 5,000 sequences) "VeryFastTree -ext AVX2 -threads 8 -double-precision"; (> 1,000 sequences) "iqtree2 -fast –alrt 1000 -T 24 -m MFP –mset Q.plant,LG,WAG,JTT"; (> 2 sequences) "iqtree2 –alrt 1000 -T 8 -m MFP –mset Q.plant,LG,WAG,JTT". The species tree was transformed into a time tree using the make_ultrametric.py script distributed with OrthoFinder, and plotted using the phyloseq,[191] ggtree,[192] and deeptime[193] packages in R.

Beyond the individual proteomes, groupings of multiple taxa were constructed according to phylogeny (e.g. monocots, Lemnaceae, *Lemna* spp., etc.) or ecology (e.g. aquatic-floating). For each grouping, HOGs (hierarchical orthogroups) that were missing exclusively from all members of the group but not other taxa (missing HOGs), and HOGs unique to the taxa in the grouping (unique paralogs) were tabulated using R scripts. Five tables were produced in this manner reflecting different phylogenetic constraints: the "all_angiosperms" table shows missing HOGs and unique paralogs at the N1 (angiosperm MRCA) level in the species tree, with the subgenomes of the *L. japonica* hybrid accessions merged; the "ath_1mono" table shows HOGs missing from each grouping that were present in *A. thaliana* and at least one other monocot outside of the target grouping at the N4 (monocot-eudicot MRCA) level; the "ath_osa" table shows missing HOGs and unique paralogs from each grouping relative to *A. thaliana* and rice; the "intra_lemnaceae" table considers only variation within duckweeds; the "hybrids" table examines HOG variation among groupings of the *L. japonica* genomes and subgenomes. eggnog-mapper v2.1.6[194] and AHRD v3.11[195] were used to assign functional and Gene Ontology term annotations to the sequences in all proteomes independently, and a merged annotation record was generated for each HOG. If the HOG contained *A. thaliana*, rice, or maize sequences, symbols (from TAIR10, IGRSP-1.0, for those genes were added to the annotation. GO-term enrichment analysis, reduction, and treemap plotting were performed for each list of missing HOGs or unique paralogs under each constraint using the TopGO[196] and rrvgo[197] R packages. Significantly enriched GO terms annotated to the gene members of each HOG were added to the merged annotation (Data S2).

Synteny analysis

Syntenic relationships between the 9 chromosome resolved Lemnaceae assemblies annotated in this study were determined and plotted using GENESPACE v1.1.4[198] with "onewayBlast = TRUE". *Z. marina* was used as an outgroup for this independent OrthoFinder run within GENESPACE, but was not used in the subsequent analysis. Lm7210 was used as the reference for riparian plots, and chromosomes of Sp9509 and Wa8730 were ordered and oriented to emphasize syntenic relationships with *Lemna* species in the plots.

Please cite this article in press as: Ernst et al., Duckweed genomes and epigenomes underlie triploid hybridization and clonal reproduction, Current Biology (2025), https://doi.org/10.1016/j.cub.2025.03.013

Genome size estimation by flow cytometry

Nuclear DNA content of duckweed accessions was measured by flow cytometry in triplicate, using *Spirodela polyrhiza* 7498[39] and *Physalis grisea*[216] as controls. For each sample, two duckweed colonies from the target accession, two *S. polyrhiza* 7498 colonies and approximately 1 cm^2 of *P. grisea* leaf were chopped with a razor blade in a 60x15mm petri dish containing 1 ml of cold Galbraith buffer (45mM $MgCl_2$, 30mM sodium citrate, 20mM MOPS, 0.1% (v/v) triton X-100, pH 7.0)[217] for two minutes. Samples were then passed through a 30 μm CellTrics disposable filter and stained with 50 μg of Propidium iodide. Fluorescence was measured on an LSR Dual Fortessa Cell Analyzer (Becton Dickinson).

GISH (Genomic *in situ* hybridization)

Chromosome preparation: The fronds were grown in liquid nutrient medium[218] under 16 h white light of 100 μmol m-2 s-1 at 24°C. The mitotic chromosome spreading was carried out according to our previously published protocol.[219] Fronds were treated in 2 mM 8-hydroxyquinoline, fixed in fresh 3:1 absolute ethanol: acetic acid, softened in PC enzyme mixture [1% pectinase and 1% cellulase in Na-citrate buffer, pH 4.6], macerated and squashed in 45% acetic acid. After freezing in liquid nitrogen, chromosome spreads were treated with pepsin [50 μg/ml in 0.01 N HCl], post-fixed in 4% formaldehyde in 2 × SSC [300 mM Na-citrate, 30 mM NaCl, pH 7.0], rinsed twice in 2 × SSC, 5 min each, dehydrated in an ethanol series (70, 90 and 96%, 2 min each) air-dried and inspected using spatial super-resolution structured illumination microscopy (3D-SIM).[220]

DNA isolation: For each sample, 0.3 g of fresh and healthy fronds were harvested and cleaned in distilled water, put into a 2 ml Eppendorf tube with two metal balls, frozen in liquid nitrogen, and ground by a ball mill mixer (Retsch MM400). The genomic DNA of the studied species was isolated using the DNeasy Plant Mini Kit (Qiagen Cat#69104). DNA was eluted by 200 μl buffer AE and quality checked by electrophoresis. Genomic DNA was sonicated before labeling.

Probe preparation: Sonicated genomic DNA (1 μg) was labeled with Cy3-dUTP (GE Healthcare Life Science) or Alexa Fluor 488-5-dUTP (Life Technologies) by nick-translation, then precipitated in ethanol[221] with sonicated unlabeled DNA of the other presumed parental species as carrier DNA in excess. Probe pellets from 10 μL nick translation product for GISH probes were dissolved in 100 μL hybridization buffer [50% (v/v) formamide, 20% (w/v) dextran sulfate in 2 × SSC, pH 7] at 37°C for at least 1 h. The ready-to-use probes were stored at -20°C.

GISH: Probes were pre-denatured at 95°C for 5 min and chilled on ice for 10 min before adding 20 μL probe per slide. Two-rounds of GISH with alternatively labeled genomic probes of the presumed parental species were performed to investigate the distribution of the corresponding probe signals on the chromosome complement of the tested clones as described.[222]

Microscopy and image processing: Fluorescence microscopy for signal detection followed.[220] To analyze the ultrastructure and spatial arrangement of signals and chromatin at a lateral resolution of ~120 nm (super-resolution, achieved with a 488 nm laser), 3D-SIM was applied using a Plan-Apochromat 63×/1.4 oil objective of an Elyra PS.1 microscope system and the software ZENblack (Carl Zeiss GmbH).

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical details of analysis applied in this paper are provided alongside in the figure legends.

ADDITIONAL RESOURCES

Genome assemblies, annotations, machine-readable tables, and browsing and analysis tools: www.lemna.org