Genome-Wide Allele-Specific Expression in Obligately Asexual *Daphnia pulex* and the Implications for the Genetic Basis of Asexuality

Zhiqiang Ye[1,*], Xiaoqian Jiang[2], Michael E. Pfrender[3] and Michael Lynch[1]

[1]Center for Mechanisms of Evolution, Arizona State University, Tempe, Arizona

[2]Systems Biology, CrownBio Inc., Suzhou, Jiangsu, China

[3]Department of Biological Sciences and Environmental Change Initiative, University of Notre Dame, Notre Dame, Indiana

*Corresponding author: E-mail: zhiqiang.yes@gmail.com.

Accepted: 24 October 2021

Abstract

Although obligately asexual lineages are thought to experience selective disadvantages associated with reduced efficiency of fixing beneficial mutations and purging deleterious mutations, such lineages are phylogenetically and geographically widespread. However, despite several genome-wide association studies, little is known about the genetic elements underlying the origin of obligate asexuality and how they spread. Because many obligately asexual lineages have hybrid origins, it has been suggested that asexuality is caused by the unbalanced expression of alleles from the hybridizing species. Here, we investigate this idea by identifying genes with allele-specific expression (ASE) in a *Daphnia pulex* population, in which obligate parthenogens (OP) and cyclical parthenogens (CP) coexist, with the OP clones having been originally derived from hybridization between CP *D. pulex* and its sister species, *Daphnia pulicaria*. OP *D. pulex* have significantly more ASE genes (ASEGs) than do CP *D. pulex*. Whole-genomic comparison of OP and CP clones revealed ~15,000 OP-specific markers and 42 consistent ASEGs enriched in marker-defined regions. Ten of the 42 ASEGs have alleles coding for different protein sequences, suggesting functional differences between the products of the two parental alleles. Atleast three ofthese ten genes appear to be directly involved in meiosis-related processes, for example, RanBP2 can cause abnormal chromosome segregation in anaphase I, and the presence of Wee1 in immature oocytes leads to failure to enter meiosis II. These results provide a guide for future molecular resolution of the genetic basis of the transition to ameiotic parthenogenesis.

Key words: *Daphnia*, asexuality, allele-specific expression, hybridization.

Significance Statement

Many obligately asexual lineages have hybrid origins, and it has been suggested that asexuality may be caused by the unbalanced expression of alleles from the hybridizing species. We test this hypothesis in *Daphnia pulex*, in which obligate parthenogens (OP) and cyclical parthenogens (CP) coexist. We found that OP clones have significantly more genes experiencing allele-specific expression (ASE) than that in CP clones, and several of the ASE genes (ASEGs) in OP clones have known functions in meiosis-related processes. Our results suggest that ASEGs originating from hybridization, through direct effects on meiosis, may underly the origin of asexuality in *D. pulex*.

Introduction

Most obligately asexual lineages do not have meiotic recombination and are thought to suffer from selective disadvantages resulting from the reduced efficiency of fixing beneficial mutations and purging deleterious mutations (Muller 1932; Maynard Smith 1978; Bell 1982). In the absence of recombination, deleterious-mutation accumulation is expected to progressively reduce fitness (Maynard Smith 1978; Bell 1982; Barton and Charlesworth 1998). However, despite the long-term evolutionary cost of asexual reproduction, transitions from an ancestrally meiotic condition to obligate asexuality have occurred in many taxa (Bell 1982; Lynch 1984; Barton and Charlesworth 1998; Burke and Bonduriansky 2019). Although the contrast between the expected reduced efficiency of natural selection and the phylogenetically and geographically wide distributions of obligate asexuals has prompted considerable theoretical investigation, the genetic/cellular mechanisms underlying the origin of asexual lineages are still largely unknown, except in a few cases such as *Wolbachia*-induced asexuality (Weeks and Breeuwer 2001; Charlat, et al. 2003).

Many obligately asexual lineages have hybrid origins (Lynch 1984; Simon et al. 2003; Neaves and Baumann 2011; Neiman et al. 2014; Lamelza et al. 2019). However, it is still unclear whether hybridization causes genetic changes that induce asexuality or whether natural selection simply favors hybrid asexuals (Lynch 1984). Consistent with the view that asexuality results from incompatibilities between hybridizing species that disrupt meiosis, hybrids from distantly related species are more likely to produce unreduced eggs (Moritz et al. 1992). Wang et al. (2010) found that unreduced gametes result from erroneous homologous chromosome pairing in plants, which may be caused by the asynchronous movement of paired homologous chromosomes during meiosis. Such asynchronous movement is thought to be caused by the unequal expression of the parental alleles in polyploids (Carman 1997). Owing to differences in *cis*-regulation (Pastinen and Hudson 2004) or genomic imprinting (Lawson et al. 2013), the two parental alleles within hybrids may have unequal expression (Springer and Stupar 2007; Shao et al. 2019), in some cases resulting in novel phenotypes (Springer and Stupar 2007; Bell et al. 2013; Corral et al. 2013; Shao et al. 2019). However, although it has been suggested that the production of unreduced gametes in hybrids may be determined by a single gene with ASE (Corral et al. 2013), a general link between ASE and asexuality remains to be established.

Here, we explore the latter issue by analyzing genome-wide ASE in the freshwater microcrustacean *Daphnia pulex*. Most *D. pulex* reproduce by cyclical parthenogenesis (CP), with females reproducing via parthenogenesis under favorable environmental conditions, but male production and haploid resting-egg production (i.e., ephippia) are induced under unfavorable conditions. However, some lineages have lost the ability to engage in meiosis and have become OP (Hebert and Crease 1980; Paland et al. 2005). OP *D. pulex* still produce embryos through parthenogenesis in favorable conditions, and in unfavorable conditions, produce diploid resting eggs requiring no fertilization. Moreover, in some cases, the OP females retain the ability to produce males capable of producing functional haploid sperm (Innes and Hebert 1988; Xu et al. 2015). As a consequence, the transition from CP to OP in *D. pulex* is thought to result from the proliferation of meiosis-suppressing genetic elements with sex-limited expression, transmitted by OP-produced males backcrossing with CP females (Lynch et al. 2008; Tucker et al. 2013).

Previous analyses have shown that all OP *D. pulex* share common haplotypes on chromosomes 5, 8, and 9, which arose by introgression from CP *Daphnia pulicaria* (Lynch et al. 2008; Tucker et al. 2013; Xu et al. 2015). However, crossing experiments in the laboratory using *D. pulex* females (because most OP *D. pulex* clones have a *D. pulex* maternal genome, based on mitochondrial genome sequence) and *D. pulicaria* males from multiple locations have only generated CP hybrids (Heier and Dudycha 2009), suggesting that hybridization between CP *D. pulex* and CP *D. pulicaria* does not always create OP hybrids. Thus, it has been suggested that OP *D. pulex* initially originated from a single ancestral hybridization event, with a diverse set of asexual haplotypes being generated subsequently by the backcrossing of OP-produced males with CP females (Lynch et al. 2008; Tucker et al. 2013; Xu et al. 2015). Because OP males carry full sets of *D. pulicaria*-derived alleles in regions of OP-shared haplotypes (Tucker et al. 2013; Xu et al. 2015) and CP females carry *D. pulex* alleles, the asexuality in OP *D. pulex* is likely caused by the sequence or expression-level divergence of alleles from two hybridizing species (i.e., *D. pulex* and *D. pulicaria*). Thus, it follows that genes in regions of OP-shared haplotypes must underlie the origin of asexuality in *D. pulex*, possibly via ASE. To establish a reliable gene set to pursue this analysis, we first generated a high-quality genome assembly of a CP *D. pulex* using 80x coverage PacBio reads, and predicted gene models from RNA-seq evidence. Second, we generated high-coverage RNA-seq data for four OP and four CP *D. pulex* clones to reveal ASEGs in each clone. Third, we identified genomic regions with markers unambiguously associated with obligate asexuality and tested for their enrichment with ASEGs. Finally, based on patterns of gene regulation in candidate regions, we identified ten candidate ASEGs with potential involvement in the origin of asexuality in *D. pulex*, with three of them having known functions in meiosis-related processes.

Results

Daphnia pulex Genome Assembly and Annotation

Although a reference genome assembly from Illumina reads is available for a clone of *D. pulex* (Ye et al. 2017), it is still

Table 1.
Comparison of the PacBio Assembly (PA42 4.2) and the Previous Assembly (PA42 3.0)

Assembly	PA42 4.2	PA42 3.0
Total size of scaffolds (bp)	155,146,858	156,418,198
Number of scaffolds	320	1,822
Longest scaffold (bp)	7,007,577	1,661,524
Mean scaffold size (bp)	484,834	85,850
Median scaffold size (bp)	188,098	5,441
N50 scaffold length (bp)	1,294,056	494,773
L50 scaffold count	27	96
Total size of gaps (bp)	3,046,365	13,443,613
Final gene models[a]	18,449	18,440
BUSCO completeness	97.8%	94.4%

[a]Gene models with >200 bp region covered by unique RNA-seq reads.

fragmented (1,822 scaffolds) and contains ~13 million gaps. To reduce potential limitations from the old assembly, we generated an improved genome assembly of a CP *D. pulex* using 80x coverage of PacBio reads from an RSII platform, with a subread N50 of 8.5 kb. The draft genome presented here, PA42 4.2, consists of 155 Mb of sequences located on 320 scaffolds. The N50 scaffold size of 1.3 Mb represents a 1.3x increase over the previous assembly (table 1). The N50 scaffold number is 27, the largest scaffold is 7.0 Mb (table 1), and mean and median scaffold sizes are 5.6x and 34.6x larger than in the previous assembly, indicating high integrity of the assembly. The predicted genome size is almost the same as the previous assembly (155 Mb vs. 156 Mb), while the total gap size has decreased by 10.4 Mb, likely due to the resolution of repetitive genomic regions by the long reads. PA42 4.2 is estimated to have a completeness of 97.8% by BUSCO (Sima~o et al. 2015), 3.4% higher than the previous version PA42 3.0 (Ye et al. 2017). Results from KAT (Mapleson et al. 2017) indicate that PA42 4.2 has fewer duplicate regions compared with PA42 3.0 (Supplementary fig. S1).

Gene prediction is vital to our analysis, which requires access to key genes related to asexuality. To obtain a reliable gene set, we predicted gene models using direct RNA-seq evidence. Because gene expression is environment-dependent, to obtain a relatively complete gene profile, we used previously published *D. pulex* RNA-seq data from various conditions, including six environmental stressors and three developmental stages (Ye et al. 2017), as well as transcriptomes induced by the predator *Chaoborus* larvae (An et al. 2018), and induced by juvenile hormone (Toyota et al. 2015). A total of 18,449 gene models predicted based on unique RNA transcripts were used in the subsequent analyses.

Identification of genes with ASE

To identify genes with allele-specific expression (ASEGs) (i.e., genes with imbalanced expression levels for the two alleles), we generated both DNA and RNA-seq data from four obligately parthenogenetic (OP) clones and four cyclically parthenogenetic (CP) clones from a *D. pulex* population (BRG) collected from Iowa, USA, where OP and CP clones coexist. To determine the DNA-level divergence among BRG clones to gain insight into their degree of independence, we constructed a phylogenetic tree using phased haplotypes from previously defined hybrid regions (i.e., chromosome 8 and 9), revealing that asexual haplotypes for OP BRG clones have the same origin as those collected from previous papers (Tucker at al. 2013; Xu et al. 2015; Supplementary fig. S2). The mean silent-site divergence in the hybrid regions among asexual haplotypes of OP BRG clones is estimated to be d_s¼ 0.00017, which is comparable to an average of 0.00023 obtained from OP clones in a broader geographic range (Supplementary fig. S2). Our results suggest that the asexual haplotypes of the BRG clones did not fully develop endogenously, but represent multiple colonizations of this population. Under the assumption of silent-site neutrality, the mean divergence time of the asexual haplotypes in BRG clones can be estimated as $d_s/(2-l)$, where l¼ 5.69 x 10^{-9}/site/generation is the DNA mutation rate estimated in Keith et al. (2016). Assuming approximately five generations per year based on our understanding of the biology of these species, we estimate that asexual haplotypes in OP BRG clones started to diverge 2,988 years ago. This estimate is similar to the radiation time of the asexual haplotypes in OP clones from a broader geographic range (4,000 years, based on pairwise synonymous divergence of all asexual haplotypes in Tucker et al. [2013] and Xu et al. [2015]), thus further ruling out the possibility of OP BRG clones being close relatives. In addition, the sexual haplotypes for the OP BRG clones have an average silent-site divergence of 0.016 6 0.0002, which is similar to that for CP BRG clones (0.017 6 0.0002), again indicating that these OP clones have different genetic backgrounds, at the levels of both OP- and CP-inherited chromosomal regions.

In an attempt to reveal why OP clones produce unreduced eggs, RNA-seq data were generated at the stage just before resting-egg (i.e., ephippium) reproduction, with each clone having ~220 million reads after filtering (Supplementary table S1). By mapping filtered reads to the 18,449 predicted gene models, we found: 15,751 (85%) genes to be expressed in at least one clone; 14,754 and 14,737 expressed in every OP and CP clone, respectively; and 14,157 expressed in every clone. To distinguish alternative alleles in each gene, we used both RNA and DNA data to reveal 10,233 genes with heterozygous SNPs in every OP clone and 7,576 genes with heterozygous SNPs in every CP clone. These genes provide the basis for ASE analysis in each phenotype.

Among the 10,233 genes that are potentially suitable for ASE analysis in OP clones, only 3,239 (32%) can be phased by short reads; in CP clones, 1,217 (16%) out of the 7,576 genes can be phased. To avoid the massive loss of data, we used GeneiASE (Edsg€ard et al. 2016) to predict genes experiencing ASE, which does not require phasing. GeneiASE records read

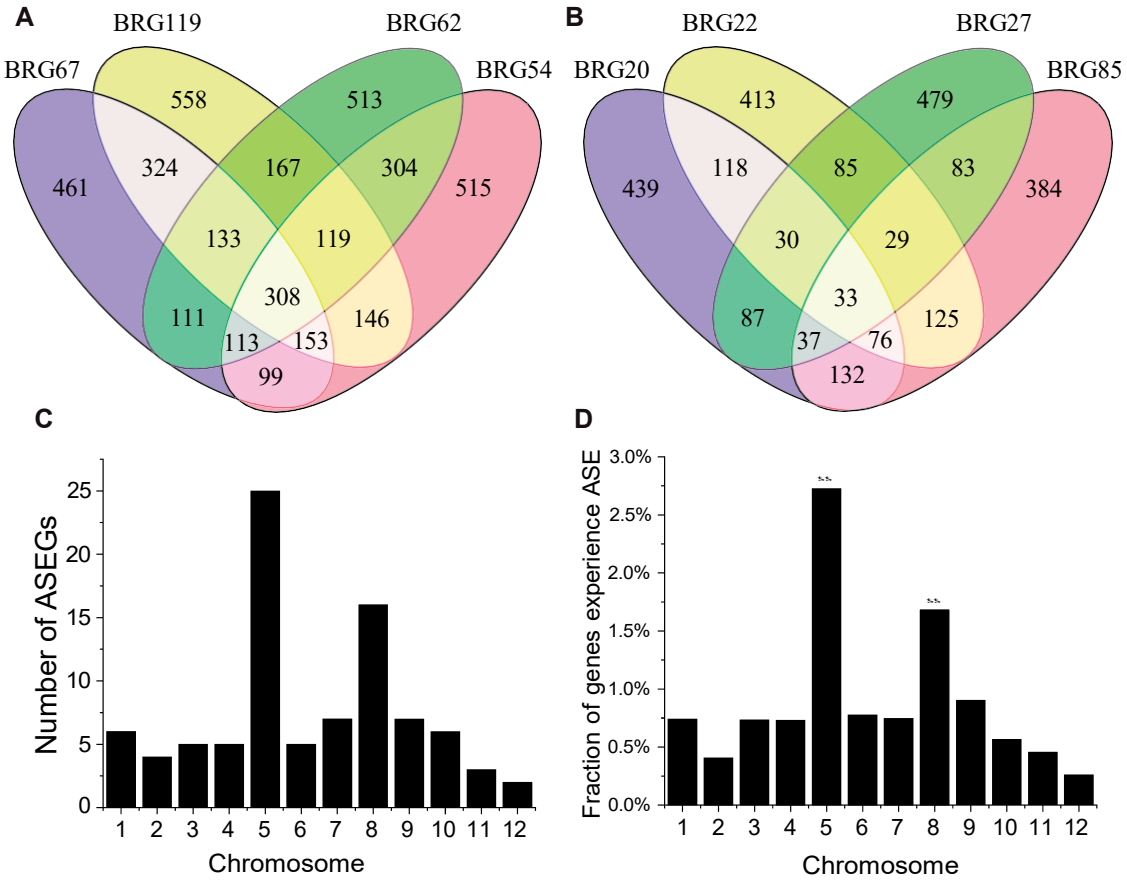

FIG. 1.—Genes experiencing ASE in each (*A*) OP and (*B*) CP *D. pulex* clone—the numbers following BRG are clone designations. (*C*) Chromosome distribution of 91 consistent OP ASEGs and (*D*) fraction of genes on each chromosome experience ASE (number of ASEGs divided by total analyzed genes on each chromosome) (**$P < 0.01$).

counts for each SNP in a gene and converts the count data to a test statistic for each SNP. SNP effects are combined for each gene to obtain a gene effect, and the gene effect is then compared with a null distribution obtained from resampling (Edsg€ard et al. 2016). Here, ASEGs were identified as those with significant gene effects from GeneiASE after correction for multiple testing.

To avoid false mapping of reads from paralogs, only uniquely mapped reads were used, and we further selected reads that are equally good at mapping to the two parental alleles (see the "Methods" section). On average, 17.4% of the 10,233 genes from each OP clone exhibit ASE, while 12.0% of the 7,576 genes in CP clones do so ($P < 0.0001$; v^2 test; fig. 1). The number of shared ASEGs in OP clones is greater than that in CP clones—308 (3.01% of the 10,233) and 33 (0.44% of the 7,576) genes experience ASE in every OP and every CP clone ($P < 0.0001$; v^2 test; figure 1*A* and *B*), respectively. Only 12 ASEGs overlap between the two reproductive types. We found that the number of CP-shared ASEGs is not significantly different from the random expectation (33 vs. 35, $P¼0.33$), while the number of OP-shared ASEGs is significantly higher than the random expectation (308 vs. 14, $P < 0.0001$). After excluding the 12 shared OP and CP ASEGs, we found that 91 of the remaining 296 OP ASEGs showed consistent directional patterns of ASE (biased toward the same alleles) in all OP clones.

Forty-one (45%) of these 91 consistent ASEGs locate to chromosomes 5 and 8 ($P < 0.01$; fig. 1*C*). This result is not simply due to chromosomes 5 and 8 containing more annotated gene models, as the fractions of genes exhibiting ASE on each of these chromosomes are significantly higher than on other chromosomes ($P < 0.01$; fig. 1*D*). Intriguingly, these two chromosomes are known to carry markers associated with hybridization between *D. pulex* and *D. pulicaria*, implicating their involvement in the origin of obligate asexuality in *D. pulex* (Lynch et al. 2008; Xu et al. 2015). Thus, our results suggest that hybridization played an important role in generating ASEGs in OP clones.

ASEGs in the Hybrid Chromosomal Regions

Because OP *D. pulex* originated from hybridization between CP *D. pulex* and CP *D. pulicaria*, chromosomal regions in OP

Table 2.
Genome-Wide Distribution of 13,039 OP-Specific SNPs and 1,813 OP-Specific INDELs

Chr	Scaffold	#SNP	#INDEL	ASEGs	Size (Mb)	Marker density (/Mb)
9	13	3,320	335	0	2.81	1,300.7
9	9	2,044	370	1	3.35	720.6
9	39	1,007	107	0	1.11	1,003.6
9	23	279	82	0	1.74	207.5
9	134	293	40	2	0.39	853.8
9	80	199	43	0	0.69	350.7
9	107	67	17	1	0.47	178.7
9	76	49	15	0	0.71	90.1
9	71	46	9	0	0.74	74.3
8	5	4,440	667	12	4.96	1,029.6
8	26	1,079	101	2	1.60	737.5
5	3	81	12	24	6.54	14.2
Total		12,904	1,798	42	25.11	585.5

Note: Scaffolds with <30 markers are not shown (135 SNP and 15 INDEL markers).

D. pulex that uniquely and consistently differ from those in CP clones are likely remnant products of introgression from *D. pulicaria*. Owing to the recurrent production of novel asexual lineages by backcrossing of OP-produced males to sexual *D. pulex*, only the chromosomal regions in such lineages associated with OP are expected to retain *D. pulicaria* DNA (Lynch et al. 2008). Thus, to determine if the 91 OP ASEGs are associated with historical hybridization, as suggested by their excess abundance on chromosomes 5 and 8, we performed a genome-wide search for regions suggesting historical introgression of *D. pulicaria* into *D. pulex*.

By comparing whole-genomic sequences of 11 CP and 29 OP *D. pulex* clones from North America, we identified 13,039 SNP markers and 1,813 indel markers present in all OP *D. pulex* but absent from all CP *D. pulex* (Supplementary file S1). All of the sites containing OP-specific SNP markers are heterozygous, with the alternative allele being identical to that in CP *D. pulex*. Ninety-nine percent of these markers are contained within 12 scaffolds, 9 of which locate to chromosome 9, 2 to chromosome 8, and 1 to chromosome 5 (table 2). Although these 12 scaffolds constitute 20.0% of the *D. pulex* genome, they contain 42 (46%) of the 91 OP ASEGs ($P < 0.001$; table 2 and Supplementary file S2), indicating that ASEGs are enriched in regions associated with hybridization. Among the 42 ASEGs, 38 locate to chromosomes 8 and 5. Although the marker-enriched scaffolds on chromosome 9 cover 48% of the entire length of identified hybrid regions, they contain only four ASEGs.

To gain a better understanding of the nature of the 42 ASEGs in the hybrid regions, we first checked their directional bias, revealing that the *D. pulex* allele had higher expression in 41 of them (fig. 2). Because the alleles of such genes reside in the same *trans*-acting regulatory environment, the allelic expression differences must be associated with *cis*-regulatory elements. To potentially identify such elements, we searched for OP-specific SNP/Indels in the likely *cis*-regulatory regions (i.e., UTRs, introns, and within 1 kb upstream of the transcription start sites). All 18 ASEGs on chromosomes 8 and 9 (table 2) contain OP-specific SNP/Indel markers (~8 markers per gene) in these regions, while none of the 24 ASEGs on chromosome 5 contain OP-specific SNP/Indels in potential regulatory regions. Because the closest OP-specific SNP/Indels are >100 kb away from the ASEGs on chromosome 5, it is unlikely that ASEGs on this linkage group are regulated by OP-specific SNP/Indels. Thus, our results suggest that ASEGs on chromosomes 8 and 9 are controlled by *cis*-regulatory elements derived from hybridization. As this appears not to be the case for the highlighted genes on chromosome 5, ASE in this case might result from differential genomic imprinting of the two alleles caused by epigenetic modifications of DNA (Lawson et al. 2013), from regulatory differences encoded within coding regions, or from mRNA stability differences.

Impacts and Functions of the 42 ASEGs in the Hybrid Regions

To further examine the potential significance of the 42 ASEGs in the hybrid regions, we investigated whether the average combined expression level of the two alleles between OP and CP clones are different. Only one (dp_gene2903) out of the 42 ASEGs is found to be differentially expressed between OP and CP clones, which is not different from random expectation (v^2 test, $P < 0.632$). Our results suggest that although the two alleles in the 42 ASEGs experience unbalanced expression, their combined expression levels are not significantly different between OP and CP clones.

To understand how the combined expression of the two alleles in OP is left the same as in CP while maintaining ASE for the two alleles, we checked the expression levels of the two alleles for the 42 ASEGs in OP clones by comparing them with the combined expression level of the two alleles in CP clones. After excluding the outlier DE gene (dp_gene2903), the average expression levels for the *D. pulicaria* and *D. pulex* alleles in OP clones are, respectively, 44.5% and 63.2% of the combined expression in CP clones. Although the expression-level difference between the two alleles in OP clones is 18.7% (63.2–44.5%), the combined expression level for the two alleles in OP clones is only 7.7% (63.2% þ 44.5–100%) higher than that that in CP clones. Our results suggest a compensatory effect between the two alleles in OP clones, that is, one goes up, and the other goes down. The higher expression for *D. pulex* alleles may be due to OP clones having *trans*-regulatory environments more similar to that in *D. pulex*, as only 20% of the genome in OP clones is associated with *D. pulicaria* introgression.

Comparison of sequence variation between the two parental alleles for the 42 ASEGs revealed the heterozygosity

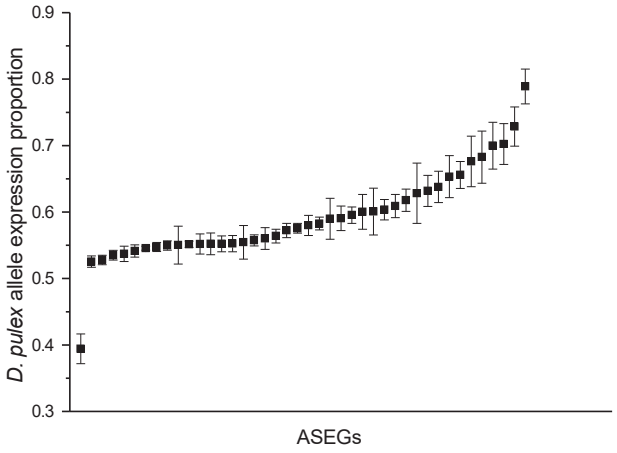

FIG. 2.—Expression proportion of the *D. pulex* alleles for the 42 loci showing consistent ASE in all OP clones. Each gene in OP clones has two alleles (*D. pulex* allele and *D. pulicaria* allele). The proportion is calculated by dividing the number of reads mapped to the *D. pulex* allele to the total mapped reads for the two alleles. Error bars are standard errors of the means. A value of >0.5 at the *Y*-axis suggests the *D. pulex* allele has a higher expression level.

level for ASEGs to be higher than that in non-ASEGs (genes in the hybrid regions that do not show ASE), although not greatly so (0.0064 vs. 0.0050; *t*-test; $P < 0.001$). To understand if these 42 ASEGs are simply an outcome of hybridization, or if they are maintained by natural selection, we calculated neutrality index (NI; Rand and Kann 1996) for these genes. NI is defined to be the ratio of p_n/p_s to d_n/d_s, where p_n and p_s are nonsynonymous and synonymous polymorphisms within OP clones, and d_n and d_s are nonsynonymous and synonymous substitutions compared with the outgroup *D. obutsa*. NI < 1 indicates an excess of nonsynonymous substitution caused by positive selection and NI > 1 indicates negative selection. We found that NI for these ASEGs are all >1, and not different from those for the non-ASEGs, indicating a lack of positive selection in the coding region during the divergence of the two alleles.

To understand whether the sequence-level difference impacts function, we searched for shared SNPs/Indels in OP clones with inferred moderate to high impacts on function (i.e., causing nonsynonymous changes, protein truncation, or loss of function) defined by SnpEff (Cingolani et al. 2012). Ten of the 42 ASEGs contain SNPs/Indels with moderate predicted impacts (table 3 and Supplementary fig. S2), which is more than in the expectation based on non-ASEGs (23.8% versus 15.3%, $P < 0.001$). Nine of the above 10 ASEGs have upwardly biased expression toward the *D. pulex* allele. However, we found one gene (Wee1) with moderately lower expression level for the *D. pulex* allele than the *D. pulicaria* allele (39% vs. 61%). Among the 10 ASEGs with predicted functional diversity, Wee1 is located on chromosome 9,

Table 3.

Ten ASEGs with Nonsynonymous Substitutions that Impact Function of the Two Parental Alleles

Gene symbol	Gene ID	Description	Chr.
RanBP2	dp_gene11888	Ran-binding protein 2	8
Nrt	dp_gene4347	Neurotactin	8
Mfsd11	dp_gene4379	UNC93-like protein	8
Ggnbp2	dp_gene4428	Gametogenetin-binding protein 2	8
Alix	dp_gene4578	ALG-2 interacting protein X	8
	dp_gene4606	Hypothetical protein	8
Upf2	dp_gene4654	Regulator of nonsense transcripts 2	8
Ace	dp_gene4684	Angiotensin-converting enzyme	8
Arrdc3	dp_gene4780	Arrestin domain-containing 3	8
Wee1	dp_gene6398	Wee1 kinase	9

Notes: Gene symbol and descriptions came from the best hits in NCBI when performing BLASTp search with minimum *e*-value of 10^{-5} and query coverage >50% are required to be considered as true candidate.

while the remaining nine ASEGs are located in a 3.8-Mb region on chromosome 8 (fig. 3). Transposable elements (TEs) unique to OP clones are thought to be involved in asexuality in *D. pulex* (Eads et al. 2012), but none of the above ASEGs have OP-specific TEs within flanking regions of 10 kb from both ends of the gene.

To understand whether and how the above ten ASEGs might be involved in asexuality, we performed functional annotation based on homology at the level of sequence and/or protein-domain structures in other model species, which suggested orthologs for nine of the ten, three of which (RanBP2, Ggnbp2, and Wee1) are inferred to be involved in meiotic cell cycles. For example, Ran-binding protein 2 (RanBP2), is a nucleoporin with SUMO E3 ligase activity that plays an important role in chromosome segregation during both mitotic and meiotic cell cycles (Pichler et al. 2002; Klein et al. 2009; Kim et al. 2017). During mouse oocyte development, the loss of RanBP2 leads to abnormal chromosome segregation during anaphase I (Kim et al. 2017). Ggnbp2 is a zinc-finger protein expressed abundantly in spermatocytes and spermatids in mice, loss of which affects prophase I and inhibits the formation of haploid cells in vitro (Guo et al. 2018). Finally, Wee1 is a universal mitotic inhibitor, known to be down-regulated during oogenesis in *Xenopus*, with ectopic expression in immature oocytes leading to failure to enter meiosis II (Nakajo et al. 2000). Although indirect, these results suggest that several of the ten candidate ASEGs in chromosomal regions inferred to be associated with asexuality are likely involved in meiosis-related processes.

Discussion

Here, we have sought candidate genes underlying asexuality by searching for those showing ASE specifically in OP *D. pulex*. Although it has been shown that unreduced eggs in hybrids

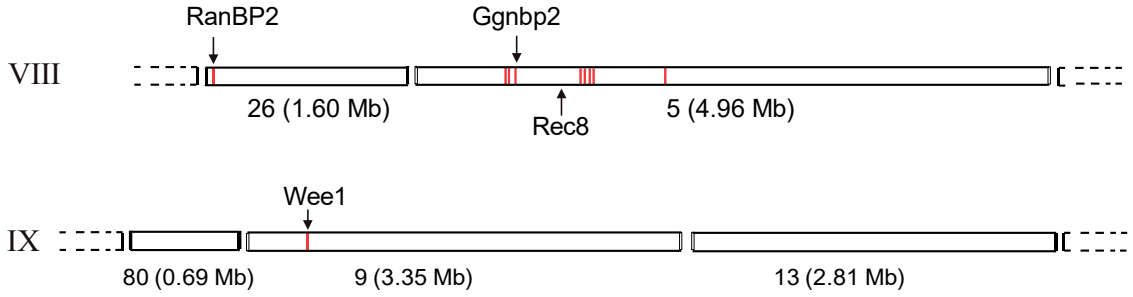

FIG. 3.—Chromosome distributions of the ten ASEGs with SNPs/Indels with moderate predicted impacts on functions. Each scaffold is scaled to its length and red bars indicate the locations of the ten candidate genes on scaffolds. Numbers preceding the parentheses are scaffold numbers and those within parentheses are scaffold lengths. Rec8 is a previously identified candidate gene underlying asexuality (Lynch et al. 2008).

can be caused by genes showing ASE in plants (Sharbel et al. 2009; Corral et al. 2013), identification of genes underlying asexuality via genome-wide study is still very challenging, as it is difficult to separate causal elements from background noise. For example, more than 50% of human genes experience ASE and while such genes are thought to be involved in many biological processes (Lo et al. 2003; Palacios et al. 2009), only a small fraction of them have been linked to phenotypic changes (De La Chapelle 2009; Wei et al. 2013; Liu et al. 2018). To control for background noise, we focused on ASEGs in regions containing markers perfectly associated with the OP trait, almost all of which are located on chromosomes 5, 8, and 9 (table 2), consistent with previous analyses (Lynch et al. 2008; Xu et al. 2015).

For causal ASEGs underlying asexuality, regulation is expected to occur in a consistent manner with functional consequences in all OP clones. Only nine genes on chromosome 8 and one on chromosome 9 satisfy these conditions. All ten genes contain nonsynonymous substitutions/indels that impact the coding sequences, and at least three of them have functions directly involved in meiosis-related processes. Aside from the above ten candidate genes on chromosome 8 and 9, 24 ASEGs on chromosome 5 are regulated in a consistent manner but the two alleles do not differ in the *cis*-regulatory regions. It is still unclear how ASEGs on chromosome 5 are regulated, but possible explanations include genomic imprinting, with the expression difference of the two alleles caused by epigenetic modifications of DNA (Lawson et al. 2013); and/or mRNA stability differences.

Although it has been postulated that asexuality in *D. pulex* may be controlled by a single dominant mutation (Hebert 1981), our results are more supportive of a multilocus mode for asexuality. The ten ASEGs implicated to be underlying asexuality in our study span at least a 3.8-Mb region on chromosome 8 and one locus on 9 (fig. 3). Because chromosomes 8 and 9 can freely segregate and crossover events occur within chromosome 8 during OP male spermatogenesis (Xu et al. 2015), the probability of a single sperm acquiring all paternal alleles for the ten ASEGs should be :::0.5. Assuming all ten paternal alleles are required for OP, and at least one recombination event within the region of ASEGs on chromosome 8, then the transmission rate of obligate asexuality by sperm would be no more than 0.25. Consistent with this, Lynch et al. (2008) found that only 2 of the 31 hybrids between 6 OP and 6 CP *D. pulex* are obligate asexual clones. Also, Xu et al. (2015) showed that in crosses of 5 OP and 1 CP *D. pulex*, only 3 out of the 122 F_1s are OP clones, while most of the F_1s appear to suffer from reproductive deficiency that prevents resting-egg (ephippial) production. Although Innes and Hebert (1988) obtained close to a 1:1 ratio (6 CP: 4 OP) for F_1 hybrids, their experiments suffered from a low success rate with only 10 out of the 178 F_1 hybrids producing ephippia. If ephippial production requires all paternal alleles from the ten ASEGs, the low success rate of ephippial reproduction in this study is also consistent with a multilocus mode of asexuality. That being said, we cannot entirely rule out a one gene model for asexuality as none of the above studies have tested whether each OP male produces viable sperm, and the low efficiency of the transmission rate may be simply due to defective sperm.

It has been hypothesized that meiosis suppression in resting-egg reproduction is the same as that during parthenogenetic reproduction (Hiruta et al. 2010) because meiosis-related genes have similar expression patterns during sexual reproduction and parthenogenesis (Schurko et al. 2009). Parthenogenesis in *D. pulex* proceeds by arresting cell division at early anaphase I, followed by a division similar to meiosis II (Hiruta et al. 2010). The result is similar to central fusion known in automictic parthenogenesis in ants and earthworms (Suomalainen et al. 1987), except that in the latter case the two daughter cells from meiosis I fuse. Our results reveal one ASEG (RanBP2) that may play an important role in chromosome segregation during anaphase I in mouse (Kim et al. 2017) and might be the cause of anaphase I abortion. We also identified one ASEG (wee1) that controls entrance to meiosis II (Nakajo et al. 2000). In *Xenopus*, wee1 needs to be silenced for the cell to enter meiosis II (Nakajo et al. 2000). Nonetheless, the function for those ASEGs is predicted from

species with large evolutionary distance with *D. pulex* and the conservation of the function requires further verification.

Although meiosis-suppression in OP *Daphnia* is female limited, with males sometimes still producing functional haploid sperm (Innes and Hebert 1988), it is thought that male production will eventually be reduced or eliminated by mutation accumulation in OP *D. pulex*, because unreduced diapausing eggs do not require fertilization. Consistent with this hypothesis, field and laboratory experiments show that most OP *D. pulex* clones have much reduced or no male-producing ability, while most CP *D. pulex* clones are capable of producing males (Hebert et al. 1989; Innes et al. 2000). In our study, all four CP clones produce males, while only two of the four OP clones can do so.

It has been shown that some OP clones produce males with functional haploid sperm, providing a route to the spread of the OP trait via backcrossing to CP clones (Xu et al. 2015). The spread of asexuality into a CP population under such a system can be very rapid, even though it involves multiple independent loci (Lynch et al. 2008). Assuming a four-locus model and an undetectable frequency (0.0001) of a male-producing OP clone, it takes only 300 sexual generations to completely displace a CP population (Lynch et al. 2008). If, as proposed here, the OP trait is controlled by only two loci, one on chromosome 8 and the other on chromosome 9, provided functional males can be produced by OP clones, only ~50 sexual generations (assuming one generation per year, it is 50 years) are required to essentially entirely displace a CP population via backcrossing starting from a very low OP frequency (Lynch et al. 2008). Moreover, Mergeay et al. (2006) found that native CP *D. pulex* in Africa have coexisted for ~60 years with the invasive OP *D. pulex* before they were completely displaced by those OP clones. Although it is still unclear whether such displacement is caused by competition or hybridization, the observation suggests that the displacement of CP by OP was a rapid process.

Although obligately asexual lineages often have hybrid origins (Lynch 1984; Simon et al. 2003; Neaves and Baumann 2011; Neiman et al. 2014; Lamelza et al. 2019), it is still unclear whether and how hybridization causes asexuality. Here, we have brought things closer to a mechanistic understanding by showing that hybridization results in ASE associated with genes involved in meiosis-related processes that may result in asexuality. In particular, hybridization leads to unbalanced expression of the two alleles for genes involving meiosis that may interfere with homologous chromosome segregation during anaphase I (Kim et al. 2017) and stop the cell from entering meiosis II (Nakajo et al. 2000). Our observations are consistent with the others showing that hybrids from distantly related species are more likely to produce unreduced eggs (Moritz et al. 1992). As the phylogenetic distance increases, the two hybridizing species are more likely to have divergent *cis*- and *trans*-regulatory environments, resulting in unbalanced expression of paternal and maternal alleles for meiosis-related genes (as well as other processes).

Materials and Methods

Daphnia pulex Genome Assembly and Annotation

A sexual isolate of *D. pulex* (PA42) was used for whole genome sequencing. DNA of the PA42 isolate was extracted from the mass culture of clonally reproducing individuals using MagAttract HMW DNA kit. To minimize bacterial contamination, animals were starved 24 h in COMBO media before DNA extraction. The DNA was sent to the University of California, Irvine, genomics high-throughput facility for PacBio sequencing. Reads were generated from 20 PacBio SMRT cells and sequenced using the RSII platform. We assembled the raw reads using Canu 1.3 (Koren et al. 2017) with default parameters. Then the initial assembly was further scaffolded with dovetail reads. Bacteria contamination from the assembly was removed following Ye et al. (2017). The duplication level for the assembly was estimated using the comp tool from KAT kit 2.4.2 (Mapleson et al. 2017) with default kmer length, and the duplicate regions were then purged using the program purge_dups (Guan et al. 2020). We estimated the completeness for the genome assembly using BUSCO V5.1.2 (Sim~ao et al. 2015), with the lineage parameter set to arthropoda_odb10. The genome-wide heterozygosity level for the PA42 4.2 assembly was estimated using GenomeScope 2.0 (Ranallo-Benavidez et al. 2020).

We annotated the PA42 draft genome using RNA-seq-guided gene discovery. To perform annotation using gene expression data, we collected RNA-seq data for *D. pulex* from various conditions, including *D. pulex* (PA42) under six environmental stressors and three developmental stages (Ye et al. 2017), predator-induced expression (An et al. 2018), and expression induced by juvenile hormones (Toyota et al. 2015). The transcriptome for all life stages/conditions were pooled and assembled de novo using the software Trinity (Haas et al. 2013) and assembled transcripts were then used as seed for evidence-based gene annotation using Maker 2.0 (Holt and Yandell 2011). Furthermore, we performed ab initio gene prediction on the draft assembly using the software SNAP (Korf 2004) and gene models were trained from 800 genes with complete gene structure (i.e., 3^0-UTR, 5^0-UTR, and more than three exons). From the ab initio gene prediction, we kept those supported by >200 bp transcripts from RNA-seq data.

Daphnia Culture and RNA Sequencing

Daphnia pulex clones used for ASE analysis were collected from O'Brien Conservation Area (BRG), IA (latitude 40.084; longitude −93.6896) in the spring of 2014. To maximize the likelihood that each individual would represent a unique genotype, we collected adult individuals soon after hatchlings and clonally expanded the individual clone in the laboratory. For each clone, we removed males from the beaker, which can be visually distinguished from females based on enlarged

antennules and a flattened ventral carapace margin. To determine whether a particular clone is a sexual or obligately asexual clone, we tested their ability to produce viable embryos in the ephippia in the absence of males.

Consistent results of no ephippial embryos (empty ephippia) from at least ten independent tests indicate the status of CP, otherwise, the status of OP would be determined.

Adult females from each clone were isolated and placed into 150 mL beakers and kept at 20°C with 120,000 cells/mL single-celled alga *Scenedesmus obliquus*. To obtain sexually reproducing (pre-ephippial) females, we selected ~200 newborn females from the mass culturing beakers and put them into ten new beakers, with each beaker containing ~20 individuals. We fed the clones on day 1 with 120,000 cells/mL single-celled alga *S. obliquus* and a 12/12 light cycle. No more food was added. These clones will reach maturity and food shortage will promote sexual production. Under favorable conditions, both CP and OP *D. pulex* directly produce embryos through parthenogenesis; in unfavorable conditions, CP *D. pulex* switch to producing haploid resting eggs (i.e., ephippium) and OP *D. pulex* produce diploid resting eggs. We used the color of *D. pulex* ovaries to distinguish females that will produce live embryos or ephippial eggs in the next clutch. Females reproducing asexually have green and bulbous ovaries, while reproducing sexually have black and smoother ovaries. Pre-ephippial females from both OP and CP clones were collected from each beaker and sacrificed for RNA and DNA extraction within 15 min. RNA for each genotype was extracted using ISOLATE II RNA Mini Kit from ~20 individuals. RNA-seq libraries were then constructed using TruSeq Stranded Total RNA Kits (Illumina) by the Genomics core facility at Arizona State University. The 150-bp paired-end reads were sequenced using Highseq 2500 platform at Genomics core facility at Arizona State University. DNA for each genotype was extracted using a cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987) and libraries constructed using a Nextera kit, followed by tagging a unique oligomer barcode for each sample. A total of 250 bp paired-end reads were sequenced using the Highseq 2500 platform at Center for Genomics and Bioinformatics at Indiana University.

Identification of OP Specific Markers

Illumina reads from 11 CP and 29 diploid OP *D. pulex* (Tucker et al. 2013; Xu et al. 2015) were mapped to the *D. pulex* reference genome (PA42 4.2), using Novoalign (http://www.novocraft.com) with option "r None" to exclude reads that mapped to more than one location. We then applied Samtools (Li et al. 2009) to generate a mpileup file for each sample (samtools mpileup -f reference.fasta sample.bam > sample.mpileup).

To avoid false-positive genotyping, we removed sites covered by less than five reads and sites with minor-allele frequencies < 0.05. A pro file is generated by pooling all mpileup files using the *proview* program in MAPGD (Ackerman et al. 2017), followed by calling genotype using the *genotype* program from MAPGD. After obtaining genotypes for each sample, we searched for alleles that existed in every OP clone but not found in any CP clone, which were defined as OP-specific SNPs. The *mpileup2indel* program from VarScan (Koboldt et al. 2009) was used to identify OP-specific INDELs from mpileup files with default parameters. Likewise, OP-specific INDELs indicate insertions and deletions exist in every OP clone but not found in any CP clone.

Identification of Heterozygous SNPs and Allelic Expression

To identify heterozygous SNPs, we mapped RNA and DNA reads for each clone to the assembly PA42 4.2 using Tophat2 (Kim et al. 2013) with "-g 1." SNPs from each clone were identified using Samtools (Li et al. 2009) with command "samtools mpileup -uf ref.fasta sorted.bam j bcftools call -mv > raw.vcf" and "bcftools filter -s LowQual -e '%QUAL < 20' raw.vcf >flt.vcf."

Analyses of ASE are known to suffer from allelic mapping bias, genotyping errors, and technical artifacts (Castel et al. 2015). To minimize such effects in our estimation, the following filters were applied: 1) Only uniquely mapped reads were used, which is controlled by using the "-g 1" option in Tophat2 (Kim et al. 2013); 2) Remove duplicate reads; 3) we masked regions with >3 SNPs within 100 bp because additional differentiating sites will bias read alignment (Castel et al. 2015); and 4) the read mapping Bam files from Tophat2 (Kim et al. 2013) and SNPs identified from both RNA-seq and genomic reads are provided to WASP (Van de Geijn et al. 2015) to select reads that have equal chances of mapping to paternal and maternal alleles. Reads passed WASP selection are counted for each allele using the ASEReadCounter module from GATK v3.8 with parameters -mmq 50 and -mbq 25 (Castel et al. 2015). Then, ASEGs for each sample are detected by GeneiASE (Edsg€ard et al. 2016) static model. Because GeneiASE does not require phasing of variants to estimate expression differences, to determine the directions of bias (paternal or maternal ASEGs) for genes, we used two strategies. First, for genes with OP-specific SNPs, all OP-specific SNPs were assigned into the same haplotype and the other allele into the background haplotype. Second, for the remaining genes, we searched for shared SNPs across clones. If >80% of the shared SNPs across clones have the same direction of bias, then it would be considered as consistent ASEG.

Silent-Site Nucleotide Diversity and Phasing Asexual Haplotypes

For each site in the coding region of *D. pulex* assembly PA42 4.2, we first defined each site as 0-fold, 2-fold, 3-fold, and 4-fold synonymous depending on whether potential

substitutions will change the amino acid. For example, if none of the potential substitutions change the amino acid, then a site will be defined as 4-fold synonymous (silent site). Then, DNA reads for the eight BRG clones were mapped to the reference assembly using Novoalign (http://www.novocraft.com) with option "r None." SNPs from each BRG clone were identified using Samtools (Li et al. 2009) with command "samtools mpileup -uf ref.fasta sorted.bam j bcftools call -mv > raw.vcf" and "bcftools filter -s LowQual -e '%QUAL < 20' raw.vcf >flt.vcf." For each defined silent site, pairwise comparison was performed between all pairs of clones to search for nucleotide differences. The silent-site nucleotide diversity was calculated by the total number of nucleotide differences divided by the number of sites compared.

We followed the method in Tucker et al. (2013) to phase the hybrid regions in OP clones, that is, all OP-specific SNPs and SNPs within the same reads goes to the asexual haplotype (because the asexual haplotype are thought to be non-recombining; Tucker et al. 2013) and the other allele into the background haplotype. Coding sequences were extracted from the phased haplotypes and provided to MEGA-X (Kumar et al. 2018) for phylogenetic inference using the neighbor-joining (NJ) method. Maximum composite likelihood model with uniform rate and homogeneous pattern among sites was used for NJ tree construction.

NI Analyses

For each of the 42 ASEGs in OP clones, we first phased them into two haplotypes (see the "phasing asexual haplotypes" section for details). Phased haplotypes from all OP clones were pooled for each gene and provided to MEGA-X (Kumar et al. 2018) to calculate within-population diversity at nonsynonymous and synonymous sites, p_n and p_s. When running MEGA-X, maximum composite likelihood model with uniform rate and homogeneous pattern among sites was used for the calculations. To estimate synonymous and nonsynonymous substitutions per nonsynonymous and synonymous sites, d_n and d_s, a *Daphnia obtusa* clone (accession number: JAACYE000000000) was used as an outgroup. Orthologous genes between *D. pulex* and *D. obtusa* were obtained by reciprocal BLAST and aligned using the *Muscle* program in MEGA X. d_n and d_s were calculated for each haplotype of the 42 ASEGs using MEGA-X and then averaged between the two haplotypes in a gene. The NI (Rand and Kann 1996) was calculated for each gene as the ratio of p_n/p_s to d_n/d_s.

Differential Expression

To detect differentially expressed genes between OP and CP *D. pulex*, we first mapped RNA-seq reads to the *D. pulex* reference genome with Tophat2 (Trapnell et al. 2012) using the "-g 1" option. Then the number of uniquely mapped reads on each gene was counted using the python script countexpression.py. DEseq2 (Anders and Huber 2010) was used to detect differentially expressed genes between OP and CP clones with FDR 0.05.

Data Availability

The *D. pulex* genome assembly PA42 v4.2 is available at www.ebi.ac.uk under accession PRJEB46221 and the corresponding annotation file can be accessed at https://osf.io/y57ct. PacBio reads for PA42 v4.2 were uploaded to NCBI SRA under accession SAMN12816806. Transcriptomes for OP and CP *D. pulex* generated in this study can be accessed under SAMN14096463-SAMN14096470. Genomic data for OP and CP *D. pulex* generated in this study can be accessed under SAMN16056572-SAMN16056579. Genomic data used for identifying OP-specific markers can be accessed at accession numbers SAMN03964756–SAMN03964769 (Xu et al. 2015) and accession numbers SAMN02252729–SAMN02252752 (Tucker et al. 2013). Genome sequence of the *D. obtusa* clone is available at the NCBI GenBank (accession number JAACYE000000000).

Acknowledgments

We thank Wei-chin Ho and Wen Wei for helpful discussions and Emily Williams for technical support.

Funding

This work was supported by the National Institutes of Health grants R01-GM101672 and R35-GM122566 to M.L. The computational work was supported by the National Center for Genome Analysis Support, funded by the National Science Foundation grant DBI-1458641 to Indiana University, by Indiana University Research Technology's computational resources.

Disclosure Declaration

The authors confirm that there are no known conflicts of interest associated with this publication.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Literature Cited

Ackerman MS, et al. 2017. Estimating seven coefficients of pairwise relatedness using population-genomic data. Genetics 206(1):105–118.

An H, Do TD, Jung G, Karagozlu MZ, Kim CB. 2018. Comparative transcriptome analysis for understanding predator-induced polyphenism in the water flea *Daphnia pulex*. IJMS 19(7):2110.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome Biol. 11(10):R106.

Barton NH, Charlesworth B. 1998. Why sex and recombination? Science 281(5385):1986–1990.
Bell G. 1982. The masterpiece of nature. Berkeley: University of California Press.
Bell GD, Kane NC, Rieseberg LH, Adams KL. 2013. RNA-seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. Genome Biol Evol. 5(7):1309–1323.
Burke NW, Bonduriansky R. 2019. The paradox of obligate sex: the roles of sexual conflict and mate scarcity in transitions to facultative and obligate asexuality. J Evol Biol. 32(11):1230–1241.
Carman JG. 1997. Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. Biol J Linn Soc. 61(1):51–94.
Castel SE, Levy-Moonshine A, Mohammadi P, Banks E, Lappalainen T. 2015. Tools and best practices for data processing in allelic expression analysis. Genome Biol. 16(1):195.
Cingolani P, et al. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) 6(2):80–92.
Charlat S, Hurst GD, Merc¸ot H. 2003. Evolutionary consequences of Wolbachia infections. Trends Genet. 19(4):217–223.
Corral JM, et al. 2013. A conserved apomixis-specific polymorphism is correlated with exclusive exonuclease expression in premeiotic ovules of apomictic *Boechera* species. Plant Physiol. 163(4):1660–1672.
De La Chapelle A. 2009. Genetic predisposition to human disease: allele-specific expression and low-penetrance regulatory loci. Oncogene. 28(38):3345–3348.
Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.
Eads BD, Tsuchiya D, Andrews J, Lynch M, Zolan ME. 2012. The spread of a transposon insertion in *Rec8* is associated with obligate asexuality in *Daphnia*. Proc Natl Acad Sci U S A. 109(3):858–863.
Edsg€ard D, et al. 2016. GeneiASE: detection of condition-dependent and static allele-specific expression from RNA-seq data without haplotype information. Sci Rep. 6:21134.
Guan D, et al. 2020. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 36(9):2896–2898.
Guo K, et al. 2018. Ablation of Ggnbp2 impairs meiotic DNA double-strand break repair during spermatogenesis in mice. J Cell Mol Med. 22(10):4863–4874.
Haas BJ, et al. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8(8):1494–1512.
Hebert PD, Beaton MJ, Schwartz SS, Stanton DJ. 1989. Polyphyletic origins of asexuality in *Daphnia pulex*. I. Breeding-system variation and levels of clonal diversity. Evolution. 43(5):1004–1015.
Hebert PD, Crease TJ. 1980. Clonal coexistence in *Daphnia pulex* (Leydig): Another planktonic paradox. Science. 207(4437):1363–1365.
Hebert PD. 1981. Obligate asexuality in *Daphnia*. Am Nat. 117(5):784–789.
Heier CR, Dudycha JL. 2009. Ecological speciation in a cyclic parthenogen: Sexual capability of experimental hybrids between *Daphnia pulex* and *Daphnia pulicaria*. Limnol Oceanogr. 54(2):492–502.
Hiruta C, Nishida C, Tochinai S. 2010. Abortive meiosis in the oogenesis of parthenogenetic *Daphnia pulex*. Chromosome Res. 18(7):833–840.
Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12(1):491.
Innes DJ, Hebert PD. 1988. The origin and genetic basis of obligate parthenogenesis in *Daphnia pulex*. Evolution 42(5):1024–1035.
Innes DJ, Fox CJ, Winsor GL. 2000. Avoiding the cost of males in obligately asexual *Daphnia pulex* (Leydig). Proc Biol Sci. 267(1447):991–997.
Keith N, et al. 2016. High mutational rates of large-scale duplication and deletion in *Daphnia pulex*. Genome Res. 26(1):60–69.
Kim D, et al. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14(4):R36.
Kim HJ, et al. 2017. Zap70 and downstream RanBP2 are required for the exact timing of the meiotic cell cycle in oocytes. Cell Cycle 16(16):1534–1546.
Klein UR, Haindl M, Nigg EA, Muller S. 2009. RanBP2 and SENP3 function in a mitotic SUMO2/3 conjugation-deconjugation cycle on Borealin. Mol Biol Cell. 20(1):410–418.
Koboldt DC, et al. 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25(17):2283–2285.
Koren S, et al. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27(5):722–736.
Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics 5(1):59.
Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35(6):1547–1549.
Lamelza P, et al. 2019. Hybridization promotes asexual reproduction in *Caenorhabditis* nematodes. PLoS Genet. 15(12):e1008520.
Lawson HA, Cheverud JM, Wolf JB. 2013. Genomic imprinting and parent-of-origin effects on complex traits. Nat Rev Genet. 14(9):609–617.
Li H, et al.; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079.
Liu Z, Dong X, Li Y. 2018. A genome-wide study of allele-specific expression in colorectal cancer. Front Genet. 9:570.
Lo HS, et al. 2003. Allelic variation in gene expression is common in the human genome. Genome Res. 13(8):1855–1862.
Lynch M. 1984. Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. Q Rev Biol. 59(3):257–290.
Lynch M, Seyfert A, Eads B, Williams E. 2008. Localization of the genetic determinants of meiosis suppression in *Daphnia pulex*. Genetics 180(1):317–327.
Mapleson D, Garcia Accinelli G, Kettleborough G, Wright J, Clavijo BJ. 2017. KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics 33(4):574–576.
Mergeay J, Verschuren D, Meester LD. 2006. Invasion of an asexual American water flea clone throughout Africa and rapid displacement of a native sibling species. Proc R Soc Lond B. 273(1603):2839–2844.
Moritz C, Wright J, Brown W. 1992. Mitochondrial DNA analyses and the origin and relative age of parthenogenetic *Cnemidophorus*: phylogenetic constraints on hybrid origins. Evolution 46(1):184–192.
Muller HJ. 1932. Some genetic aspects of sex. Am Nat. 66(703):118–138.
Nakajo N, et al. 2000. Absence of Wee1 ensures the meiotic cell cycle in *Xenopus* oocytes. Genes Dev. 14(3):328–338.
Neaves WB, Baumann P. 2011. Unisexual reproduction among vertebrates. Trends Genet. 27(3):81–88.
Neiman M, Sharbel T, Schwander T. 2014. Genetic causes of transitions from sexual reproduction to asexuality in plants and animals. J Evol Biol. 27(7):1346–1359.
Palacios R, et al. 2009. Allele-specific gene expression is widespread across the genome and biological processes. PLoS ONE. 4(1):e4150.
Paland S, Colbourne JK, Lynch M. 2005. Evolutionary history of contagious asexuality in *Daphnia pulex*. Evolution 59(4):800–813.
Pastinen T, Hudson TJ. 2004. Cis-acting regulatory variation in the human genome. Science 306(5696):647–650.
Pichler A, Gast A, Seeler JS, Dejean A, Melchior F. 2002. The nucleoporin RanBP2 has SUMO1 E3 ligase activity. Cell 108(1):109–120.

Ranallo-Benavidez TR, Jaron KS, Schatz MC. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 11(1):1432.
Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: Contrasts among genes from *Drosophila*, mice, and humans. Mol Biol Evol. 13(6):735–748.
Schurko AM, Logsdon JM, Eads BD. 2009. Meiosis genes in *Daphnia pulex* and the role of parthenogenesis in genome evolution. BMC Evol Biol. 9(1):78.
Shao L, et al. 2019. Patterns of genome-wide allele-specific expression in hybrid rice and the implications on the genetic basis of heterosis. Proc Natl Acad Sci U S A. 116(12):5653–5658.
Sharbel TF, et al. 2009. Molecular signatures of apomictic and sexual ovules in the *Boechera holboellii* complex. Plant J. 58(5):870–882.
Sim~ao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31(19):3210–3212.
Simon JC, Delmotte F, Rispe C, Crease T. 2003. Phylogenetic relationships between parthenogens and their sexual relatives: The possible routes to parthenogenesis in animals. Biol J Linn Soc. 79(1):151–163.
Maynard Smith J. 1978. The evolution of sex. Vol. 4. Cambridge: Cambridge University Press.
Springer NM, Stupar RM. 2007. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. Plant Cell. 19(8):2391–2402.
Suomalainen E, Saura A, Lokki J. 1987. Cytology and evolution in parthenogenesis. Boca Raton (FL): CRC Press.
Toyota K, et al. 2015. NMDA receptor activation upstream of methyl farnesoate signaling for short day-induced male offspring production in the water flea, *Daphnia pulex*. BMC Genomics 16(1):186.
Trapnell C, et al. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 7(3):562–578.
Tucker AE, Ackerman MS, Eads BD, Xu S, Lynch M. 2013. Population-genomic insights into the evolutionary origin and fate of obligately asexual *Daphnia pulex*. Proc Natl Acad Sci U S A. 110(39):15740–15745.
Van de Geijn B, McVicker G, Gilad Y, Pritchard JK. 2015. WASP: allele-specific software for robust molecular quantitative trait locus discovery. Nat Methods 12(11):1061–1063.
Wang CJ, et al. 2010. Formation of unreduced gametes is impeded by homologous chromosome pairing in tetraploid *Triticum turgidum* x *Aegilops tauschii* hybrids. Euphytica 175(3):323–329.
Weeks A, Breeuwer J. 2001. Wolbachia–induced parthenogenesis in a genus of phytophagous mites. Proc R Soc Lond B. 268(1482):2245–2251.
Wei QX, et al. 2013. Germline allele-specific expression of DAPK1 in chronic lymphocytic leukemia. PloS ONE 8(1):e55261.
Xu S, et al. 2015. Hybridization and the origin of contagious asexuality in *Daphnia pulex*. Mol Biol Evol. 32(12):3215–3225.
Ye Z, et al. 2017. A new reference genome assembly for the microcrustacean *Daphnia pulex*. G3 (Bethesda). 7(5):1405–1416.

Associate editor: Andrea Betancourt