Genomic Signatures of Speciation in Sympatric and Allopatric Hawaiian Picture-Winged *Drosophila*

Lin Kang[1], Robert Settlage[1], Wyatt McMahon[2], Katarzyna Michalak[1], Hongseok Tae[1], Harold R. Garner[3], Elizabeth A. Stacy[4], Donald K. Price[4,†], and Pawel Michalak[1,*,†]

[1]Biocomplexity Institute, Virginia Tech, Blacksburg, Virginia

[2]Howard Hughes Medical Institute, Johns Hopkins Medical Institutes, Baltimore, Maryland

[3]Primary Care Research Network and the Center for Bioinformatics and Genetics, Edward Via College of Osteopathic Medicine, Blacksburg, Virginia

[4]Tropical Conservation Biology and Environmental Science Graduate Program, University of Hawaiì at Hilo

*Corresponding author: E-mail: pawel@vbi.vt.edu.

[†]These senior authors contributed equally to this work.

Accepted: April 18, 2016

Data deposition: The sequence data are available in the NCBI's Sequence Read Archive (SRA) under the accession number SRP063110.

Abstract

The Hawaiian archipelago provides a natural arena for understanding adaptive radiation and speciation. The Hawaiian *Drosophila* are one of the most diverse endemic groups in Hawaiì with up to 1,000 species. We sequenced and analyzed entire genomes of recently diverged species of Hawaiian picture-winged *Drosophila*, *Drosophila silvestris* and *Drosophila heteroneura* from Hawaiì Island, in comparison with *Drosophila planitibia*, their sister species from Maui, a neighboring island where a common ancestor of all three had likely occurred. Genome-wide single nucleotide polymorphism patterns suggest the more recent origin of *D. silvestris* and *D. heteroneura*, as well as a pervasive influence of positive selection on divergence of the three species, with the signatures of positive selection more prominent in sympatry than allopatry. Positively selected genes were significantly enriched for functional terms related to sensory detection and mating, suggesting that sexual selection played an important role in speciation of these species. In particular, sequence variation in *Olfactory receptor* and *Gustatory receptor* genes seems to play a major role in adaptive radiation in Hawaiian pictured-winged *Drosophila*.

Key words: Hawaiian *Drosophila*, speciation, genome analysis, allopatry, sympatry.

Introduction

The advancement of genomics and bioinformatics provides new approaches to elucidate the relationships between evolutionary processes and genomic divergence patterns, as well as between genomic properties and speciation processes (Seehausen et al. 2014). The mode of speciation should have profound impacts on the genomic architecture and patterns of reproductive isolation of new species. Instances of speciation in sympatry with gene flow can promote the rapid evolution of reproductive isolating barriers that generally build up much more slowly in allopatry (Noor 1995; Coyne and Orr 2004; Seehausen et al. 2014). Although divergence at both the phenotypic and genomic levels associated with speciation are documented, combined analyses of genomic changes and empirical-based measures of reproductive isolation within young radiations comprising both sympatric and allopatric speciation are lacking (Feder et al. 2012). To investigate the genomic changes associated with both sympatric and allopatric settings, we have sequenced and assembled three new genomes from a recently diverged and well-studied clade of Hawaiian picture-winged *Drosophila*.

The Hawaiian *Drosophila* are a spectacular example of sequential colonization, adaptive radiation, and speciation in the islands with nearly 1,000 estimated species, of which more than 500 have been described to date (O'Grady et al. 2011). Within the system, the Hawaiian picture-winged *Drosophila* are a charismatic subgroup of approximately 120 species (Magnacca and Price 2015). Sexual selection involving mating and aggressive behaviors and their associated morphological ornaments has been proposed as an important

© The Author 2016. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

driver of reproductive isolation and speciation in the Hawaiian picture-winged *Drosophila* (Hoikkala and Kaneshiro 1993), especially for closely related species sharing the same island (Carson and Clague 1995, Carson 1997). For closely related species occurring on different islands, geographic isolation coupled with genetic drift or differential adaptation between island colonists and their source populations may be more important for speciation.

Drosophila heteroneura and *Drosophila silvestris* are two iconic Hawaiian picture-winged *Drosophila* species endemic to midaltitude rainforests of the Big Island (Hawai'i), the youngest (< 0.5 Myr old) of the Hawaiian Islands (Carson 1982). Their closest relative, *Drosophila planitibia*, presumably directly derived from the ancestral lineage that colonized the Big Island, is endemic to a similar habitat on Maui (Carson and Kaneshiro 1976; DeSalle and Giddings 1986; Bonacum et al. 2005). All three species are morphologically distinct, particularly *D. heteroneura*, which possesses a novel stalk-eyed head shape (fig. 1*A*) and distinct male-male aggressive behaviors (Kaneshiro 1976; Price and Boake 1995). Over the past several decades, the species have experienced population declines, with *D. heteroneura* currently classified as federally endangered.

Interspecies mate discrimination is greater in experimental crosses between sympatric species relative to those between allopatric species, with *D. silvestris* and *D. heteroneura* females being less discriminatory against allopatric *D. planitibia* males (Ahearn et al. 1974). Despite their greater mate discrimination, *D. silvestris* and *D. heteroneura* hybridize in nature (Carson et al. 1989) with asymmetrical mating between the species determined at an early stage of courtship, indicating the importance of species-recognition factors in the behavioral reproductive isolation and maintenance of species boundaries (Kaneshiro 1976; Price and Boake 1995; Boake 2005; Price et al. 2014). Patterns of fertility of F_1 hybrids also differ between the allopatric and sympatric species pairs. F_1 hybrid males are sterile in allopatric crosses, that is, *D. planitibia* bred with either *D. silvestris* or *D. heteroneura* (Ahearn et al. 1974; Brill et al. 2016), but F_1 hybrid females are fertile, as is

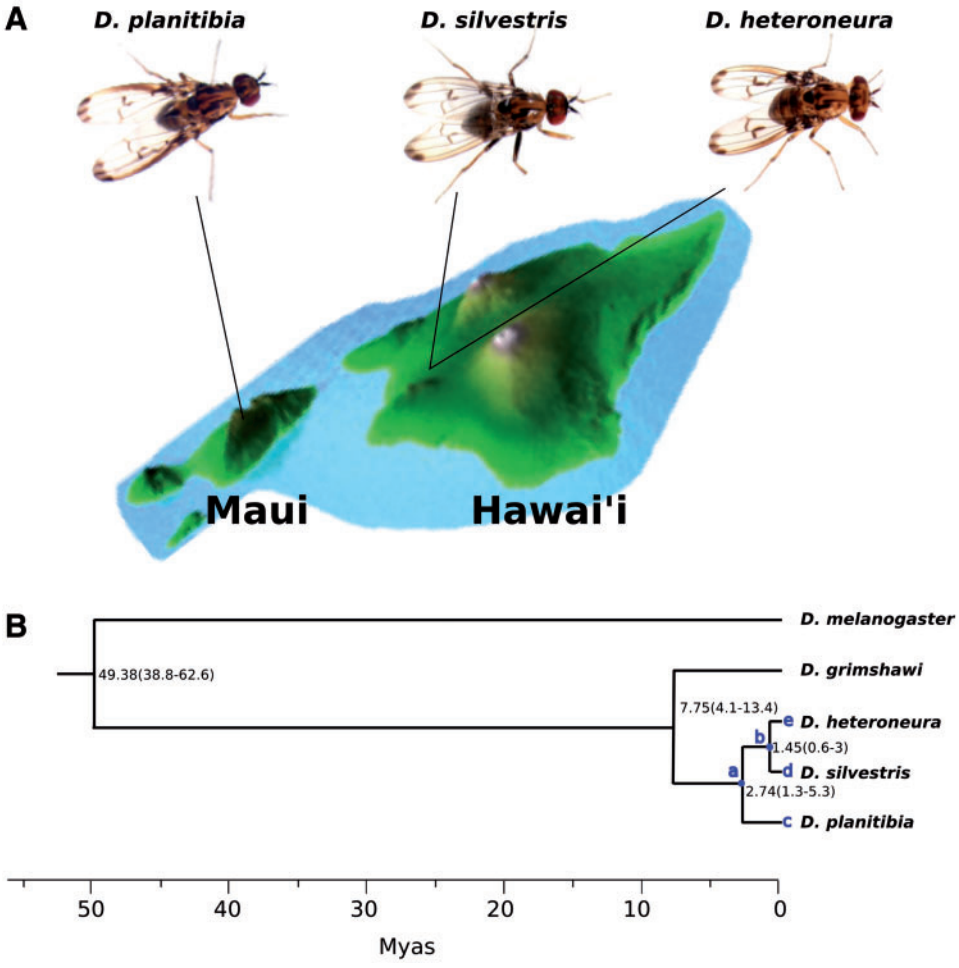

Fig. 1.—(*A*) Recent speciation in Hawaiian *D. silvestris*, *D. heteroneura*, and *D. planitibia*. (*B*) A phylogenetic tree based on 100 homologs from mcmctree (Yang 2007). Divergence time in Ma, with intervals in parentheses.

common in crosses between closely related *Drosophila* species (Coyne and Orr 1997). Both F_1 hybrid males and females produced through crosses between sympatric *D. silvestris* and *D. heteroneura* are fully fertile and possess distinct combinations of morphological and behavioral traits consistent with both dominant and additive genetic factors that differ between the species (Boake et al. 1998). These contrasts suggest very different patterns of genome divergence in sympatry versus allopatry.

Due to the sequential geological formation of islands in the Hawaiian archipelago, founder events within the *Drosophila* appear to have occurred in sequential order from the oldest (northwest) to the youngest (southeast) islands (Price and Clague 2002). Therefore, *D. silvestris* and *D. heteroneura* endemic to the slopes of the volcanoes on Hawaiì, the newest island in the chain (~0.5 Myr), are not only the youngest species in the group, but their ancestry can be traced back to lineages from neighboring islands. *Drosophila planitibia* living on the geologically older island of Maui is the closest sister species to both Hawaiì Island species on the basis of morphological, behavioral, chromosomal, and genetic characteristics (Hunt et al. 1984; DeSalle and Giddings 1986). An alternative hypothesis, however, based on morphological and behavioral data, places *D. planitibia* closer to *D. silvestris* and a separate species from Molokai, *Drosophila differens*, situates closer to *D. heteroneura*, implying two independent ancestral lineages from different islands. Here, we report the sequencing and analysis of entire genomes of *D. planitibia*, *D. heteroneura*, and *D. silvestris*, three recently diverged picture-winged species for which pre- and postzygotic reproductive isolating barriers are well-characterized. We contrast the two leading phylogenetic hypotheses using the three new genomes and analyze the signatures of adaptive evolution that recent speciation may have left on genomes of these species.

Materials and Methods

Flies

Genomic DNA was extracted (Gentra Puregene Tissue Kit, Qiagen) from 10 *D. heteroneura*, 10 *D. silvestris*, and 10 *D. planitibia* noninbred males and pooled within species. DNA pooling enabled us to compare allele frequencies per SNP and estimate differentiation between populations (*F*st), while keeping sequencing costs relatively low (Kofler, Pandey, et al. 2011). *Drosophila heteroneura* and *D. silvestris* individuals were from populations initiated with wild-caught individuals collected at the same location in the rainforest at 1,400 m elevation in the Kukuiopa'e section of South Kona Forest Reserve from 2009 to 2011. *Drosophila planitibia* originated from Waikamoi Preserve, east Maui, and were collected in December 2012. All flies were raised at the University of Hawaiì at Hilo.

Genome Sequencing

Illumina paired-end HiSeq (2 × 100 bp, 500 bp inserts), paired-end Miseq (2 × 300 bp, 500 and 800 bp inserts), and Nextera Mate Pair libraries were sequenced at a total sequence coverage greater than 80×.

Genome Assembly

Adapters were removed from the raw sequencing reads, and low quality and duplicated reads were discarded using FastqMcf (Aronesty 2013), error corrections were performed by SOAPec from SOAPdenovo2 package (Luo et al. 2012). The 300 bp pair-end reads from Miseq were merged into long reads using mergepairs from ABYSS package (Simpson et al. 2009). Mate-pair sequences were processed with nextclip (Leggett et al. 2014) with default parameters, and reads from categories A, B, and C were used to make the assembly. To exclude possible contamination, all reads were aligned to bacterial database downloaded from NCBI (http://www.ncbi.nlm.nih.gov/, last accessed April 30, 2016), and unmapped reads were used for the assembly. Processed reads were assembled with Spades (Bankevich et al. 2012) and duplicated reads were removed with picard (https://github.com/broadinstitute/picard, last accessed April 30, 2016) according to the alignments which mapped the processed reads against the first assembly. Deduplicated reads were then assembled with Spades combined with scaffolding step using SSPACE (Boetzer et al. 2011) (default parameters) and contigs with length less than 500 bp were discarded from further analyses (table 2).

Genome Completeness

The completeness of assembly was estimated using CEGMA by examining 248 core eukaryotic genes (Parra et al. 2007). Completeness estimates were in the range of 93–98% ("complete") and 98–99% ("partial").

Gene Prediction and Annotation

Protein-coding genes were predicted using MAKER2 (Holt and Yandell 2011), which used *Drosophila melanogaster* protein sequences from FlyBase (r6.02, http://flybase.org) as protein homology evidence and integrated with prediction methods including BLASTX and SNAP. Predicted genes were subsequently used as query sequences in a BLASTx database search of NR database (nonredundant database, http://www.ncbi.nlm.nih.gov/). BLASTx alignments with e-value greater than 1e-10 were discarded, and the top hit (or top hit from *Drosophila* species if existed) was used to annotate the query genes.

Functional Enrichment

All functional enrichment analyses were performed by importing the appropriate gene list into DAVID (Huang da et al. 2009)

and using annotated genes (HET, SIL, PLA) or *D. melanogaster* as background. GO terms with a Benjamini–Hoschberg-adjusted *P* value of <0.05 were considered significant.

Ka/Ks Ratio

To reduce the possible impact of Ka/Ks ratio by wrong annotation, we used only annotations against Swissprot (http://www.ebi.ac.uk/uniprot), BLASTx alignments with e-value greater than 1e-40 or identity less than 40% were discarded. Sequences with same annotation were grouped together, and Clustal-omega (Sievers et al. 2011) was used to conduct the multiple sequence alignments. Nucleotide sequences were parsed to amino acid sequences before carrying multiple-sequence alignments to avoid possible frameshift, and the amino acid sequences of alignment were changed back to nucleotide sequences for Ka/Ks calculations. PAML (Yang 2007; version 4.7) was used to calculate the Ka/Ks ratio values, setting the model = 0 in the control file of codeml. To further minimize the possible effect by the wrong annotation and grouping, Ks values greater than 2 were excluded from further analyses, and the maximal Ka/Ks value was set to be 3. Models M7/M8 along with likelihood ratio tests were applied to test for the significance of positive selection, with *P* values generated from chi-square distribution (Nielsen and Yang 1998). Pairwise Ks and Ka/Ks values are presented in a pairwise fashion, except for supplementary table S4, Supplementary Material online, that contains results of Ks and Ka/Ks from both pairwise comparisons and a single multispecies alignment (Ks(all) and Ka/Ks(all)).

McDonald–Kreitman Test

To test for signatures of selection, we also used the McDonald–Kreitman (MK) test that compares the number of synonymous (Ds) and nonsynonymous (Dn) substitutions between species with the number of synonymous (Ps) and nonsynonymous (Pn) polymorphisms within species (McDonald and Kreitman 1991). *Drosophila planitibia* was used as reference for mapping and detection of polymorphisms and substitutions. We used GATK (DePristo et al. 2011) with default parameters for genotyping. Only sites with the minimum depth of 10 and minimum genotyping quality of 30 were used. Sites with at least two reads supporting an alternative allele were considered polymorphic. Sites showing polymorphism in at least one of the three species were counted as polymorphic sites, and those with fixed differences between species were counted as substitutions. *P* values were computed using Fisher exact test. Statistic *DoS* was used to determine the direction of selection (Stoletzki and Eyre-Walker 2011), as given by: $DoS = \frac{D_n}{D_n+D_s} - \frac{P_n}{P_n+P_s}$. Positive and negative *DoS* values suggest positive and purifying selection, respectively.

F_{ST} and d_{XY} Estimates

Sequences were mapped using BWA (Li and Durbin 2010) with default parameters and *Drosophila grimshawi* assembly as reference. Samtools (Li et al. 2009) was used to generate the pileup result. SNPs within 10 bp of an indel were discarded and Poopolation2 (Kofler, Pandey, et al. 2011) was used to estimate the F_{ST} value for each SNP. All pairwise analyses used the maximum number of sites, that is, F_{ST} estimates are based on sites that are polymorphic in at least one of the three species or divergent (if monomorphic) between at least two species. progressiveMauve (Darling et al. 2010) was used for multiple sequence alignments of *D. silvestris*, *D. heteroneura*, *D. planitibia*, and *D. grimshawi*. PoPoolation (Kofler, Orozco-terWengel, et al. 2011) was used to estimate pairwise divergence (d_{XY}) with the window size set to 10 kb.

Phylogeny

A total of 100 orthologs with the highest confidence from BLASTX alignments of *D. heteroneura, D. silvestris, D. planitibia, D. grimshawi*, and *D. melanogaster* were used to construct a phylogenetic tree (fig. 1*B*). Mcmctree based on a Bayesian Markov Chain Monte Carlo algorithm from PAML (Yang 2007) package was used to estimate the divergence time, and the calibration time was set using the divergence time between *D. melanogaster* and *D. grimshawi* from TimeTree (www.timetree.org), that is, 39 (Thomas and Hunt 1991) to 62.9 Ma (Tamura et al. 2004).

Results and Discussion

Pooled genomes from ten noninbred individuals per species were used for construction of paired-end and mate pair libraries that were Illumina-sequenced at a total greater than 80× coverage. First, to determine if *D. silvestris* (SIL) originated from *D. planitibia* (PLA) and *D. heteroneura* (HET) diverged from another species, such as *D. differens*, we predicted that genetic distances should be lower between PLA and SIL than between PLA and HET, even under extensive gene flow between SIL and HET in sympatry. However, our genome-wide analysis of the average number of pairwise differences between sequences d_{XY} showed that SIL and HET were the closest relatives (table 1 and fig. 1*B*). Also, mean fixation index (F_{ST}) values based on 4,558,111 SNPs were lowest between SIL and HET (0.141), thus consistent with shorter divergence

Table 1

Average Pair-Wise Divergence (d_{XY}) Values (Below Diagonal) and F_{ST} Values (Above Diagonal) for HET, SIL, and PLA Pairwise Comparisons

	HET	SIL	PLA
HET	—	0.141	0.277
SIL	0.0077	—	0.306
PLA	0.0130	0.0121	—

Table 2

Genome Assembly Attributes of the Three Hawaiian Picture-Winged *Drosophila*

	Assembly Attributes					
	Total Size	No. of Scaffold	No. of Scaffold ≥ 1k	Contig N50	Scaffold N50	GC Content (%)
D. silvestris	146,901,421	8,486	6,624	16,915	92,229	38.92
D. heteroneura	144,943,455	10,998	7,322	17,229	92,746	39.01
D. planitibia	188,994,020	15.471	11,830	154,334	399,542	40.55

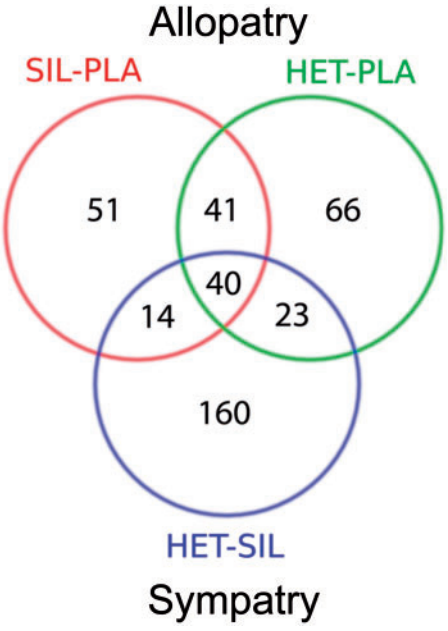

Fig. 2.—A Venn diagram illustrating overlap between *D. silvestris*, *D. heteroneura*, and *D. planitibia* in the number of genes driven by positive selection.

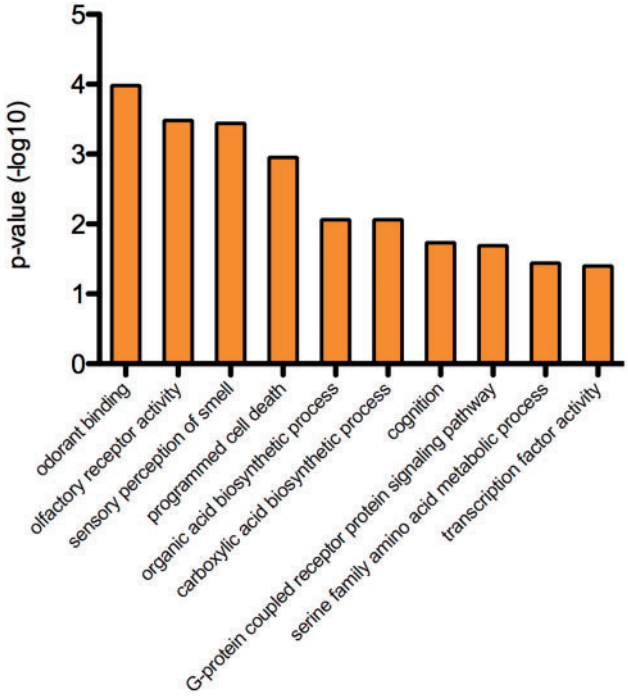

Fig. 3.—Gene Ontology terms and their statistical significance, showing an overrepresentation of genes related to sensory detection and cognition.

time and/or interbreeding between these two, but higher and very similar between PLA and SIL (0.306) and between PLA and HET (0.277). The average number of synonymous substitutions per synonymous site (Ks) showed a similar pattern, with the lowest value (0.041) between SIL and HET, and almost identical values for the other two pairwise comparisons (0.049). These estimates support an evolutionary scenario with the most recent phylogenetic split between SIL and HET.

A remarkable difference among the genomes of these species is that HET and SIL, the two sympatric species on Hawaiì Island and most closely related species of the three examined, have on average approximately 50% more genes with Ka/Ks > 1 than observed in the allopatric species pairs (PLA-SIL and PLA-HET; fig. 2, supplementary table S1, Supplementary Material online), many of them underlying sensory perception (GO term enrichment false discovery rate [FDR] < 0.002, supplementary table S2, Supplementary Material online). For comparison, most genes exhibiting signatures of purifying selection (Ka/Ks < 1) were conserved genes shared by all three species with patterns showing no relationship to sympatry or allopatry. The relative increase of positive selection in sympatry is possibly due to sexual selection and reinforcement, if hybrids are maladapted or when male secondary sex characters evolve through runaway processes (Noor 1995; Higashi et al. 1999). In fact, HET and SIL have very divergent morphologies that have been proposed to be associated with divergent mating or male aggressive behaviors (Boake 2005). The most conspicuous morphological trait of HET is the wide stalk-eyed-shaped head, whereas SIL and PLA have a round head more typical of other *Drosophila* (fig. 1*A*). Among the most abundant groups of overrepresented genes driven by positive selection were many odorant (*Or*) and gustatory (*Gr*) receptor genes (fig. 3 and supplementary table S3, Supplementary Material online).

In addition to *Or* and *Gr* genes, pheromone-binding *antennal protein 10* and neurological system- and courtship-related *spinster* were also driven by positive selection.

We then used the MK test that contrasts levels of polymorphism and divergence at neutral and functional sites (McDonald and Kreitman 1991) to further examine signatures of adaptive evolution at the genomic level. Similar to the Ka/Ks test results, genes related to sensory perception and odorant binding formed one of the largest functionally enriched groups (GO term enrichment FDR < 0.05) of all genes driven by positive selection according to the MK test ($P < 0.05$, supplementary tables S4 and S5, Supplementary Material online). Overall, out of the 62 *Or* and *Gr* genes, the MK test showed 23 (37%) genes to be under significant positive selection (direction of selection DoS > 0, $P < 0.05$, supplementary table S3, Supplementary Material online), which is a threefold enrichment relative to all other genes (372 (11%) out of 3,470 genes, $P = 6.67 \times 10^{-8}$, Fisher exact test). Out of the 11 *Or* and *Gr* genes under positive selection indicated by Ka/Ks tests, eight showed significant signatures of positive selection in the MK tests as well.

Chemosensation in *Drosophila* is critical for detecting food and avoiding toxicants, as well as for courtship and mating. Since all three species share the same primary host plant (lobeliad trees of the genus *Clermontia* [Kaneshiro and Val 1977]), sexual selection operating on traits related to mate discrimination may have taken precedence over food-related adaptations in sensory divergence. HET and SIL males also differ in their concentrations of epicuticular hydrocarbons (Alves et al. 2010), molecules that are important for courtship communication in *Drosophila*. Lastly, courtship displays of SIL and HET also differ in the timing for stage advancement, the degree of female responsiveness, and the speed of body and/or wing movement (Watson 1979; Hoikkala and Kaneshiro 1993). The substantial differences in Ka/Ks values between SIL and HET compared to both SIL-PLA and HET-PLA suggest that strong behavioral reproductive isolation may be driving genome divergences in sympatry. In sum, this study emphasizes the significance of genomic sequences of Hawaiian picture-winged *Drosophila* in the inference of processes involved in speciation and adaptive radiation.

Supplementary Material

Supplementary tables S1–S5 areavailable at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was funded in part by NSF CREST award HRD 0833211 to D.K.P. and VBI s Medical Informatics and Systems funds to P.M. High performance computing was supported by a grant from the NSF (OCI-1124123). The authors also thank Chris Parypa for his help with graphics.

Literature Cited

Ahearn JN, Carson HL, Dobzhansky T, Kaneshiro KY. 1974. Ethological isolation among three species of the planitibia subgroup of Hawaiian *Drosophila.* Proc Natl Acad Sci U S A. 71:901–903.

Alves H, et al. 2010. Evolution of cuticular hydrocarbons of Hawaiian *Drosophilidae.* Behav Genet. 40:694–705.

Aronesty E. 2013. Comparison of sequencing utility programs. Open Bioinform J. 7:1–8.

Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19: 455–477.

Boake CRB. 2005. Sexual selection and speciation in Hawaiian *Drosophila.* Behav Genet. 35:297–303.

Boake CRB, Price DK, Andreadis DK. 1998. Inheritance of behavioural differences between two interfertile, sympatric species, *Drosophila silvestris* and *D. heteroneura.* Heredity 80:642–650.

Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579.

Bonacum J, O'Grady PM, Kambysellis M, Desalle R. 2005. Phylogeny and age of diversification of the *planitibia* species group of the Hawaiian *Drosophila.* Mol Phylogenet Evol. 37:73–82.

Brill E, Kang L, Michalak K, Michalak P, Price DK. Forthcoming 2016. Hybrid sterility and evolution in Hawaiian *Drosophila*: differential gene and allele-specific expression analysis of backcross males. Heredity.

Carson HL. 1982. Evolution of *Drosophila* on the newer Hawaiian volcanoes. Heredity 48:3–25.

Carson HL. 1997. The Wilhelmine E. Key 1996 Invitational Lecture. Sexual selection: a driver of genetic change in Hawaiian *Drosophila.* J Hered. 88:343–352.

Carson HL, Clague DA. 1995. Geology and biogeography of the Hawaiian Islands. In: Wagner W, Funk, V, editors. Hawaiian biogeography: evolution in a hotspot Archipelago. Washington, DC: Smithsonian Institution Press. p. 14–29.

Carson HL, Kaneshiro KY. 1976. Drosophila of Hawaii—systematics and ecological genetics. Ann Rev Ecol Syst. 7:311–345.

Carson HL, Kaneshiro KY, Val FC. 1989. Natural hybridization between the sympatric Hawaiian species *Drosophila silvestris* and *Drosophila heteroneura.* Evolution 43:190–203.

Coyne JA, Orr HA. 1997. Patterns of speciation in *Drosophila* revisited. Evolution 51:295–303.

Coyne JA, Orr HA. 2004. Speciation. Sunderland (MA): Sinauer Associates.

Darling AE, Mau B, Perna NT. 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147.

DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43:491–498.

DeSalle R, Giddings LV. 1986. Discordance of nuclear and mitochondrial DNA phylogenies in Hawaiian *Drosophila.* Proc Natl Acad Sci U S A. 83:6902–6906.

Feder JL, Egan SP, Nosil P. 2012. The genomics of speciation-with-gene-flow. Trends Genet. 28:342–350.

Higashi M, Takimoto G, Yamamura N. 1999. Sympatric speciation by sexual selection. Nature 402:523–526.

Hoikkala A, Kaneshiro K. 1993. Change in the signal-response sequence responsible for asymmetric isolation between *Drosophila planitibia* and *Drosophila silvestris.* Proc Natl Acad Sci U S A. 90: 5813–5817.

Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12:491.

Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4:44–57.
Hunt JA, Bishop JG 3rd, Carson HL. 1984. Chromosomal mapping of a middle-repetitive DNA sequence in a cluster of five species of Hawaiian *Drosophila.* Proc Natl Acad Sci U S A. 81:7146–7150.
Kaneshiro KY. 1976. Ethological isolation and phylogeny in the planitibia subgroup of Hawaiian *Drosophila.* Evolution 30:740–745.
Kaneshiro KY, Val FC. 1977. Natural hybridization between a sympatric pair of Hawaiian *Drosophila.* Am Nat. 111:897–902.
Kofler R, Orozco-terWengel P, et al. 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS One 6:e15925.
Kofler R, Pandey RV, Schlotterer C. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435–3436.
Leggett RM, Clavijo BJ, Clissold L, Clark MD, Caccamo M. 2014. NextClip: an analysis and read preparation tool for Nextera long mate pair libraries. Bioinformatics 30:566–568.
Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595.
Li H, et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.
Luo R, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1:18.
Magnacca KN, Price DK. 2015. Rapid adaptive radiation and host plant conservation in the Hawaiian picture wing *Drosophila* (Diptera: Drosophilidae). Mol Phylogenet Evol. 92:226–242.
McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351:652–654.
Nielsen R, Yang Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929–936.
Noor MA. 1995. Speciation driven by natural selection in *Drosophila.* Nature 375:674–675.
O'Grady PM, et al. 2011. Phylogenetic and ecological relationships of the Hawaiian *Drosophila* inferred by mitochondrial DNA analysis. Mol Phylogenet Evol. 58:244–256.
Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061–1067.
Price DK, Boake CRB. 1995. Behavioral reproductive isolation in *Drosophila silvestris*, *D. heteroneura* and their F1 hybrids (Diptera: Drosophilidae). J Insect Behav. 8:595–616.
Price DK, Souder S, Varys T. 2014. Sexual Selection, epistasis and species boundaries in sympatric Hawaiian picture-winged *Drosophila.* J Insect Behav. 27:27–40.
Price JP, Clague DA. 2002. How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. Proc Biol Sci. 269:2429–2435.
Seehausen O, et al. 2014. Genomics and the origin of species. Nat Rev Genet. 15:176–192.
Sievers F, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7:539.
Simpson JT, et al. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19:1117–1123.
Stoletzki N, Eyre-Walker A. 2011. Estimation of the neutrality index. Mol Biol Evol. 28:63–70.
Tamura K, Subramanian S, Kumar S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. Mol Biol Evol. 21:36–44.
Thomas RH, Hunt JA. 1991. The molecular evolution of the alcohol dehydrogenase locus and the phylogeny of Hawaiian *Drosophila.* Mol Biol Evol. 8:687–702.
Watson GF. 1979. On premating isolation between two closely related species of Hawaiian *Drosophila.* Evolution 33:771–774.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.

Associate editor: Liliana Milani