# Supplementary Information for

## Reproductive capacity evolves in response to ecology through common developmental mechanisms in Hawai'ian Drosophila

Didem P. Sarikaya*, Samuel H. Church, Laura P. Lagomarsino, Karl N. Magnacca, Steven Montgomery, Donald K. Price, Kenneth Y. Kaneshiro and Cassandra G. Extavour*

* Correspondence:

D.P.Sarikaya
didemps@gmail.com, Evolution and Ecology Department, University of California Davis, Davis CA 95616

C.G.Extavour
extavour@oeb.harvard.edu, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138

**This PDF file includes:**

**SUPPLEMENTARY INFORMATION TEXT**

## Extended Methods

*Field collections and Hawai'ian* Drosophila *husbandry*

Field collections of Hawai'ian *Drosophila* were conducted under the Department of Land and Natural Resources of Hawai'i native invertebrate scientific collection permits FHM14-305 and FHM14-353. Collections were made at the Koke'e State Park and Kui'a NAR on Kauai, West Maui Watershed Reserve, Makawao Forest Reserve, and Waikamoi Nature Preserve on Maui, and the Volcanoes National Park and Upper Waiakea Forest Reserve on Hawai'i island. Flies were collected by aspirating flies from traps or sponges containing fermenting fruit and fungi, or by sweeping and sorting leaf litter in forests.

Field-caught females were maintained on yeast-less Wheeler-Clayton medium or *Drosophila* standard laboratory medium at 18°C at 80% humidity. Each vial contained a piece of tissue paper (Kleenex) moistened with water that was steeped with various Hawai'ian *Drosophila* egg-laying substrates to stimulate oviposition.

Larvae of picture wing subgroup and *Antopocerus* species of the AMC subgroup have longer development times than the other species studied herein, and larvae of these species were fed additional food, which was made as follows: 6g of Agar and 225mL distilled water were mixed in a 1L beaker and microwaved for two minutes. 60g cornmeal, 6.6g roasted soybean meal and 7.5g brewer's yeast were mixed, blended and added to the beaker along with an additional 300mL distilled water, and mixed with a spoon. Lastly, three tablespoons of Karo light corn syrup and one tablespoon of unsulfured molasses was added to the mix, and the mixture was microwaved for three minutes. Food was mixed every minute during microwaving until the mixture was close to boiling point and started to rise up within the beaker. The beaker containing hot food was placed at room temperature until the mixture was warm enough to touch. 3mL of propionic acid and 3mL of 99% ethanol were added, and the solidified food was stored at 4 °C. The solidified food was mixed with a small quantity of water to soften the consistency before being used to feed larvae.

Non-picture wing Hawai'ian *Drosophila* species pupariated on the side of the glass vials, and hatched F1 offspring were transferred into new vials. Larvae of the picture wing subgroup species pupate in the soil. To accommodate this behavior, food vials with wandering picture wing larvae were placed in a large jar containing 1-2cm of moist sand at the bottom. A piece of cloth or paper towel held in place using a rubber band was used close the opening of each jar. Larvae migrated from the vials to the sand to pupariate, and thus adults emerged from the sand, and were aspirated out of the jar into a fresh adult food vial.

*Measurement of adult phenotypes*

Adult ovaries were dissected in 1X PBS, and placed in 2% paraformaldehyde in 1X PBS overnight at 4°C. Ovaries were then stained with the nuclear dye Hoechst 33342 (Sigma, 1:500 of 10mg/ml stock solution) in 1X PBS for two hours at room temperature, then washed with 1X PBS for a total of one hour. Ovaries were mounted on glass slides in Vectashield mounting medium (Vector Labs), and ovarioles were spread apart using

tungsten needles for species with high ovariole number. Ovariole number was counted under fluorescent and white light microscopy using a Zeiss AxioImager microscope. Images of eggs were taken from these slides using DIC white light settings. Egg volume was calculated when mature eggs were laid by captive adult flies or when present in dissected ovaries. In these cases, egg volume was estimated by measuring the straight lines across the longest and widest points of the egg, and assuming a prolate spheroid shape following a previously published protocol (1) using ImageJ.

Adult bodies were placed in 99% ethanol after dissection for DNA extraction and adult size analysis. Lateral view images of the thorax were captured using a Zeiss Lumar Stereomicroscope. The highest point of the anterior tip of the thorax and the posterior-most point of the scutellum in the same image plane were used to measure thorax length. A straight line was drawn between these two points in these images using ImageJ's measure function. Thorax volume was calculated as thorax length $(mm)^3$ as a proxy for body size, and proportional egg volume was calculated by dividing egg volume by thorax volume.

*Measurement of larval phenotypes*

Wandering larvae or early pupal stage individuals were dissected in 1X PBS + 0.1% Triton-X and fixed in 4% Paraformaldehyde in 1X PBS for 20 minutes at room temperature. Larval ovaries were stained as previously described (2) using mouse anti-Engrailed/Invected (4D9, Developmental Studies Hybridoma Bank, 1:50), FITC-conjugated Phalloidin (Sigma, 1:120), and Hoechst 33342 (Sigma, 1:500 of 10mg/ml stock solution). Samples were post-fixed in 4% paraformaldehyde in 1X PBS for 15 minutes at room temperature and mounted in Vectashield mounting medium (Vector Labs) for imaging using a Zeiss LSM780 Confocal Microscope at the Harvard Biological Imaging Center. Quantification of TFCs and TFs was conducted as previously described (2).

For some samples, fewer collected specimens could be used for measuring total TF number than for others, as total TF number can only be counted in larval ovaries where TF morphogenesis has completed, which is usually near the end of larval development (2). At the time of larval ovary dissection, some ovaries contained completed TFs, while others were still undergoing morphogenesis and could not be used to gather data on TF number. In the latter cases, TFC number per TF was measured for those TFs that had completed morphogenesis, and total TFC number for that species was assigned based on the average TF number from other specimens from the same species.

*PCR amplification of mitochondrial genes for species identification*

While there are detailed dichotomous keys for species identification of Hawai'ian *Drosophila*, these keys focus on male-specific traits including male genitalia and other sexually dimorphic characters. Therefore we identified female flies using a combination of morphological features (3, 4), collection site information, and DNA barcode-based methods as previously described (5). Following ovary dissection, abdominal at tissue was used for DNA extraction using the Qiagen Blood and Tissue kit. PCRs were conducted using the following primer sets (from 5' to 3') as previously published (6):

**COI F** ATT CAA CCA ATC ATA AAG ATA TTG G

| | |
|---|---|
| **COI R** | TAA ACT TCT GGA TGT CCA AAA AAT CA |
| **COII F** | ATG GCA GAT TAG TGC AAT GG |
| **COII R** | GTT TAA GAG ACC AGT ACT TG |
| **ND2 F** | AGCTATTGGGTTCAGACCCC |
| **ND2 R** | GAAGTTTGGTTTAAACCTCC |

PCR was conducted using Dynazyme DNA polymerase (Thermo Scientific) as follows: 95 °C 5 minutes, (95 °C 30 seconds, 50 °C (COI and COII) or 54 °C (16S) 30 seconds, 72 °C 30 seconds) x 30, 72 °C 5 minutes. PCR products were cleaned using ExoSAP-IT (Affymetrix) and sequenced by Genewiz (Cambridge, MA). Sequences were analyzed by 4Peaks (Nucleobytes).

*Analysis of mitochondrial sequences for species delineation*

All sequences were first analyzed by BLASTn alignment against the NCBI Nr/Nt collection, and the accession numbers were noted for those where there was one clear hit with 98-100% sequence identity to one species (Table S2). Species that returned multiple 99-100% BLASTn hits are summarized in Table S3. We note that in many cases, multiple hits are within a closely related species subgroup (Table S3), as reported in previous phylogenetic studies of Hawai'ian *Drosophila* (6-11). When BLAST results of all three barcodes were consistent, the sample was assigned a species identity. In cases where the BLAST analysis did not provide a clear identity, we tested whether the sample was sister to a single reference species in gene trees constructed using RAxML, as described below in 'Phylogenetic Inference'. Samples with unambiguous BLAST and/or tree-based support were assigned a species identity. Some samples resulted in BLAST and tree-based support for a closely related species group; in these cases, a species group identity was assigned instead of a specific species. Samples that did not have clear support for any species group hit were discarded from the dataset.

*Phylogenetic inference*

Our taxon sampling for phylogenetic inference combines the efforts of four previous studies (6-8, 10) and additional newly identified mitochondrial sequences listed in Table S9. This sampling includes members of all major lineages of Hawai'ian drosophilids. Nucleotide sequences from each of these four studies were downloaded from GenBank, totaling 18 genes. The sequence IDs were parsed using the program phyutility v2.2.6 (12), and the 18 genes were aligned individually using MAFFT v7.130b (13) with the "auto" option selected, and trimmed with Gblocks v0.91b (14) with the "with half" option selection. Trimmed sequences were concatenated using phyutility into two alignments, one including all 18 available genes and one including only the four mitochondrial genes. This second alignment reflects the analysis performed by O'Grady and colleagues (6). PartitionFinder v1.1.1_Mac (15); options 'raxml') was used to find the best fitting model for each partition; GTR + Γ +I was found for nearly all partitions. For species delineations, sequences of the three targeted genes generated in this study were combined with homologous sequences from the four previous studies and aligned

and trimmed using the same procedure as above. Gene trees were generated in RAxML v8.2.3 (16) using a GTR + Γ +I model of sequence evolution).

Phylogenetic relationships and divergence time estimates were inferred simultaneously using both the mitochondrial and mitochondrial+nuclear alignments in a Bayesian framework in BEAST v. 2.3.2 (17, 18). A single calibration at the root (i.e., at the base of Hawai'ian *Drosophila* + *Scaptomyza*) was used to infer divergence times; this was assigned a uniform prior from 23.9–37.1 Ma, following (7). Rate-smoothing was performed using a relaxed lognormal clock model (19). The BEAST analyses followed the same partitioning scheme as RAxML, as determined by PartitionFinder (15) and utilized a birth-death tree prior. Four separate chains were allowed to run for 100 million generations (sampling every 10,000) using the CIPRES supercomputer cluster (https://www.phylo.org). Convergence was assessed using the AWTY web interface (20) and effective sample size (ESS) values of the runs (using values >200 as a cutoff) in Tracer (21). After convergence was reached, the individual runs were combined and the maximum clade credibility tree, including credibility intervals (CI) for ages and posterior probabilities (PP) for node support, was assembled in TreeAnnotator (22). Upstream phylogenetic comparative analyses used either the maximum clade credibility (MCC) tree or a subset of 500 trees from the posterior distribution of trees, as appropriate; analyses were repeated for each of the two BEAST analyses.

*Phylogenetic Generalized Least Squares Analyses*

All phylogenetic comparative analyses and corresponding figures were computed in R version 3.2.0 (23). The evolutionary relationship between ovariole number, egg volume, egg volume over thorax volume, and thorax volume was analyzed in pairs using phylogenetic generalized least squares in **nlme** v.3.1-121(24), using the phylogenetic correlation matrix generated using the *corMartins* function in **ape** v.3.3 (25, 26) with a small starting alpha value. All traits were natural log transformed prior to analysis. Pairwise comparisons were performed over 100 trees randomly drawn from the posterior distribution generated in BEAST. The range and average of both the p-value and the slope from the PGLS models across the 100 trees were calculated, and a cutoff threshold of 0.05 was used to determine significance of the p-values.

*Analysis of Evolutionary Regimes*

We used all reported ecological information about Hawai'ian *Drosophila* as summarized by Magnacca and colleagues (27) to code oviposition site for the species in our dataset. Three different coding schemes were compared: (1) OU8, which considered eight ecological substrates (bark, flower, fruit, fungus, generalist, leaf, sap and spider egg breeders); (2) OU3, which considered three states (bark, flower & spider egg, and 'other substrate' breeders); and (3) a final one with two states (OU2: bark and non-bark breeders). We categorized species as bark-breeders if they utilized the tree stem or trunk, though previous studies distinguished between the two (28, 29). Ancestral states for each of these character codings were calculated over 100 trees randomly drawn from the posterior distribution of trees generated with BEAST. The most likely ecological state was mapped at each node using the *rayDISC* function in the R package **corHMM**, v.1.18 (30), and the resulting tree was pruned to include only tips with ovariole number data.

The fit of three models of trait evolution were assessed on pruned trees using the R package **OUwie** v.1.48 (31). The three models tested were Brownian Motion (BM1), Ornstein-Uhlenbeck with a single optimum for all species (OU1), and Ornstein-Uhlenbeck with optima for each ecological state (OUM; OUM was fit for eight, three and two state models for each of the distinct character codings, respectively OU8, OU3, and OU2 as described above). Corrected Akaike Information Criterion (AICc) values were compared for each of these models for each of the trait coding schemes, where the best-fit model (i.e. the model with the lowest delta AICc score) was moderately supported when other models had delta AICc of 2-10, and strongly supported when delta AICc>10. This analysis was repeated over each of the 100 trees, and the frequency of each best-fitting model was recorded. Optimized theta values for each of the three OUM analyses were untransformed and recorded (Table S5).

All of the commands used to perform phylogenetic comparative analyses, as well as the corresponding commands to generate the figures, are available in the public repository https://github.com/shchurch/hawaiian_drosophila_ovaries_2018, commit e1cfd33.

## Extended Technical Description of Results and Discussion

*Taxon sampling in this and previous studies*

All measurements were taken following methods from (32), and 15 species from that study were also included in our study. The majority of the overlapping data were within two standard deviations of the previous study (Supplemental Table 1; >3 standard deviations highlighted in yellow), suggesting that these traits have remained stable over the last 40 years, such that our measurement methods are comparable to those of the previous study. We therefore included the data from Kambysellis and Heed (32) for subsequent analyses. Our final dataset for analysis contains 35 newly characterized species, 15 species included in both our field-caught dataset and Kambysellis and Heed (32) , and 16 additional species discussed by Kambysellis and Heed (32) but not found by us in the field, yielding a total of 66 Hawai'ian *Drosophila* species across all major species groups that were used in the analysis herein (Figure 1).

*Phylogenetic relationships between Hawai'ian* Drosophila *in this study*

Our study focuses on the Hawai'ian clade Drosophilidae, which comprises an estimated 1000 species in two genera, *Drosophila* and *Scaptomyza*. The phylogeny of this clade and its substituent subclades has been the focus of many recent studies (6-11). The monophyly of Hawai'ian drosophilids is well-supported, as is the monophyly of *Scaptomyza* and Hawai'ian *Drosophila* within them. The relationships between and within the subgroups of Hawai'ian *Drosophila* (*Scaptomyza*, PW, MM, Haleakala, and AMC species groups) remain less resolved. Specifically, the monophyly of each of these individual groups is generally well-supported (but see the moderate support values for the MM group reported in (10) and (6)), but the relationships between these large groups are in conflict across studies.

To generate a phylogenetic tree for evolutionary modeling analysis, we compiled data from published phylogenetic studies and captured the two topologies of Hawai'ian Drosophila species groups that have been published to date. The consensus tree topology that we recovered matched the topology recovered in recently published phylogenetic analyses (6, 10), and the second most common basal topology matched the alternative topology presented in(10). Our phylogenetic analysis also captured monophyletic species subgroups in AMC, PW and *Scaptomyza* (7, 8, 10). Given that species relationships at the species group and subgroup level were recovered in the phylogenies used for phylogenetic comparative methods analysis, we believe our analysis represents an accurate estimate of our current knowledge of Hawai'ian Drosophila phylogeny.

*The effect of larval substrate on ovariole number evolution*

Evolutionary modeling analysis showed that ovariole number in Hawai'ian Drosophila is best explained by evolutionary forces related to egg-laying substrate. Specifically, the three-state model that we tested (OU3), which distinguishes between bark breeding, the specialist substrates of *Scaptomyza* specialists, and other substrates, was the best fit for our ovariole number data across a majority of trees for ovariole number ($\Delta$AICc > 2 as compared to OU2 and OU8 models). The Brownian motion (BM), one-state model (OU1) and two-state (OU2) model lacked support as compared to the OU3 model ($\Delta$AICc > 2). Along with strong support for the OU3 model, we obtained

occasional support for the full eight state model (OU8) (Table S4). It is possible that the limited sample size for some substrate categories in this model contributed to the poor fit of OU8 to our data. Further studies with deeper sampling to obtain increased representation of non-bark oviposition substrates will be needed clarify the extent to which finer distinctions between specific specialist substrates may contribute to adaptive changes in ovariole number.

These results were largely unchanged when comparing models using an alternative topology generated from mtDNA sequences (Table S4), with all analyses supporting a role for ecology in driving trait evolution. Since five out of the 66 species represented in the analysis were categorized into species groups rather than a single species based on the DNA barcoding, we assigned two different species IDs and ran the OU analysis to address whether group assignments had an impact on our analysis. Assigning IDs to closely related species within the group did not alter the results (Table S4). Similarly, results were unchanged when only the data collected for this study were considered, excluding data previously reported by Kambysellis and Heed (32) (Table S4).

Here we note some of the difficulties faced by researchers wishing to rigorously and thoroughly account for oviposition substrate in these analyses. Species keys are not available for females of most non-PW Hawai'ian *Drosophila* species, and DNA barcode data that were used to identify the species were not available for many samples that were collected. Since Haleakala species were difficult to key, most samples from this group were excluded from the analysis. Further, presence of specific Hawai'ian *Drosophila* species in the field can be unpredictable, and difficulty in encountering them during field work is increasingly compounded by the declining numbers of endemic species in Hawai'i. For example, sap flux specialists have been documented to exist in the PW, MM, and AMC groups (27), but we were only able to collect data from two PW sap flux breeders, one from the field (*D. picticornis*) and another from a laboratory line (*D. hawaiiensis*). Lastly, certain egg-laying substrates are observed in very few species. For example, Titanocheta *Scaptomyza* species are spider-egg breeders, and this trait appears to have evolved only once (8, 33). Despite these challenges to taxon sampling, however, our analysis rejected the null model of Brownian evolution, and supported the hypothesis that ovariole number evolves in response to changes in egg-laying substrate across Hawai'ian *Drosophila*.

*The effect of larval substrate on body and egg size evolution*

In addition to ovariole number, we tested whether shifts in larval ecology influenced the evolution of body size and egg index (calculated as the phylogenetic residual of egg volume to thorax volume), as these traits are often correlated with ovariole number and have been predicted to evolve in response to changes in ecology and reproductive strategy. For body size, we found that models that accounted for ecological evolution did not fit the data better than a Brownian Motion model (BM, $\Delta$AICc > 2). For egg index, models that accounted for ecological evolution fit the data better than BM and OU1 ($\Delta$AICc > 2), but we were unable to distinguish within between models ($\Delta$AICc < 2) (Table S10). These results suggest that the evolution of ovariole number, but not overall body size, has been linked to changes in larval ecology within the Hawai'ian radiation of *Drosophila*.

*The relationship between ovariole number and egg size*

One of the life history characteristics commonly observed in animals is the inverse relationship between high reproductive capacity and investment into offspring (34). Egg size is often considered a proxy for maternal investment in insects, and life history theory predicts a trade-off between maternal investment egg size and reproductive output, thereby predicting a negative correlation between ovariole number and egg size (35). Previous empirical studies have found evidence for this predicted inverse correlation between egg size and ovariole number in some insects (36, 37).

We compared the evolutionary relationship between ovariole number and egg size across Hawai'ian *Drosophila*, accounting for the relationship to body size in each variable using phylogenetic residuals (38). We observed a significant negative correlation that explained most of the variation in relative egg size (Figure 3C-D; Table 2). Specifically, we found that when controlling for body size, species with more ovarioles have proportionally smaller eggs (Figure 3C-D; Table 2). This result suggests that the allometric relationship between ovariole number and egg size is complex, and implies that there are constraints preventing the evolution of both large eggs and high ovariole number.

*The developmental basis for evolutionary change in Hawai'ian Drosophila ovariole number*

Overall larval ovary morphology of Hawai'ian *Drosophila* was similar to that of the *melanogaster* subgroup species (Figure 4D-F), with characteristic TFC stacks forming toward the end of larval development (white arrowhead, Figure 4D-F). To determine whether, as in *D. melanogaster* (2, 39, 40), ovariole number is established by the end of larval development and does not change during the pupal phase, we compared total TF number in Hawai'ian *Drosophila* ovaries that had completed TF morphogenesis to adult ovariole number per ovary. We found a close to 1:1 correlation between TF number per ovary and ovariole number per ovary (Supplemental Table 8).

One notable difference was in *S. caliginosa*, which had one ovariole per ovary in our adult samples, and two TFs per ovary in the developmental analysis. Given that Kambysellis and Heed (32) previously reported *S. caliginosa* females with more than two ovarioles, our result may be due to the small sample size of adults of *S. caliginosa* in our study (n = 5 versus n = 24 in the previous study), or to the difficulty of counting ovarioles in this species. However, we note that honeybees can destroy ovarioles that are formed during larval stages through programmed cell death during pupal development (41). Thus, we cannot exclude the possibility that the difference in larval TF and adult ovariole number observed in *S. caliginosa* may be a result of a similar developmental process.

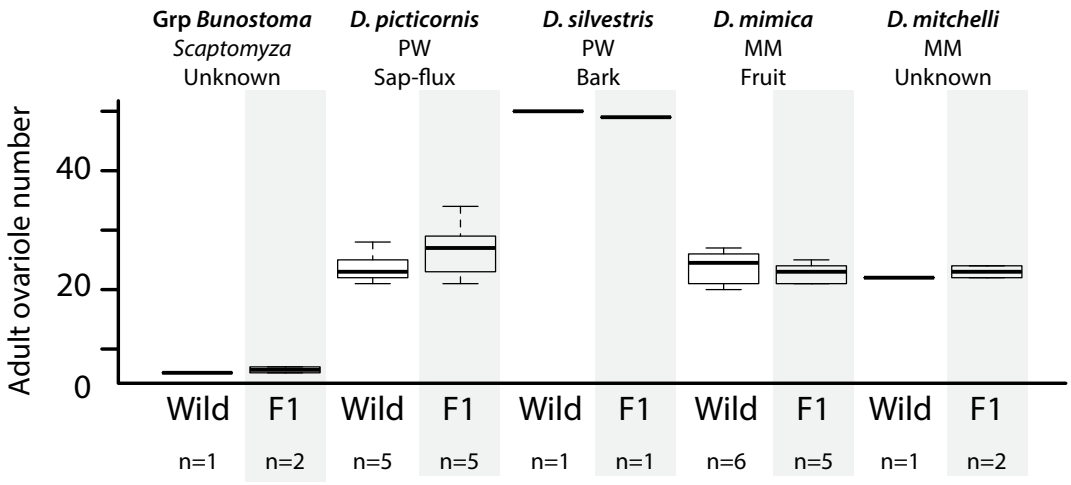


**Fig. S1. Comparison of mean ovariole number of wild caught females and their F1 offspring reared in the laboratory**. Box plot of ovariole number from wild-caught females and their F1 offspring reared in the same laboratory condition for each species. Species name is indicated along with the species group (*Scaptomyza*, PW for picture wing and MM for modified mouthpart) and oviposition substrate. Ovariole number is not significantly different between wild-caught females and F1 females for any species, regardless of oviposition substrate, natural or laboratory diet. Sample size (number of adults) is indicated below the plot.

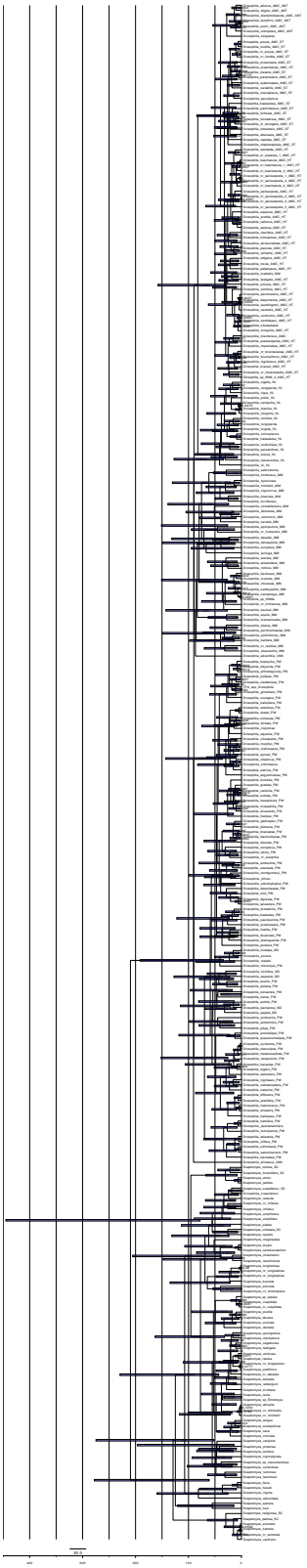

**Fig. S2. Maximum clade credibility chrononogram of Hawaiian *Drosophila* from a BEAST analysis of mitochondrial and nuclear sequences.** Bars at nodes correspond to the 95% highest probability distribution of age estimates. Values at nodes represent relative support for that relationship, and correspond to Bayesian posterior probability values.

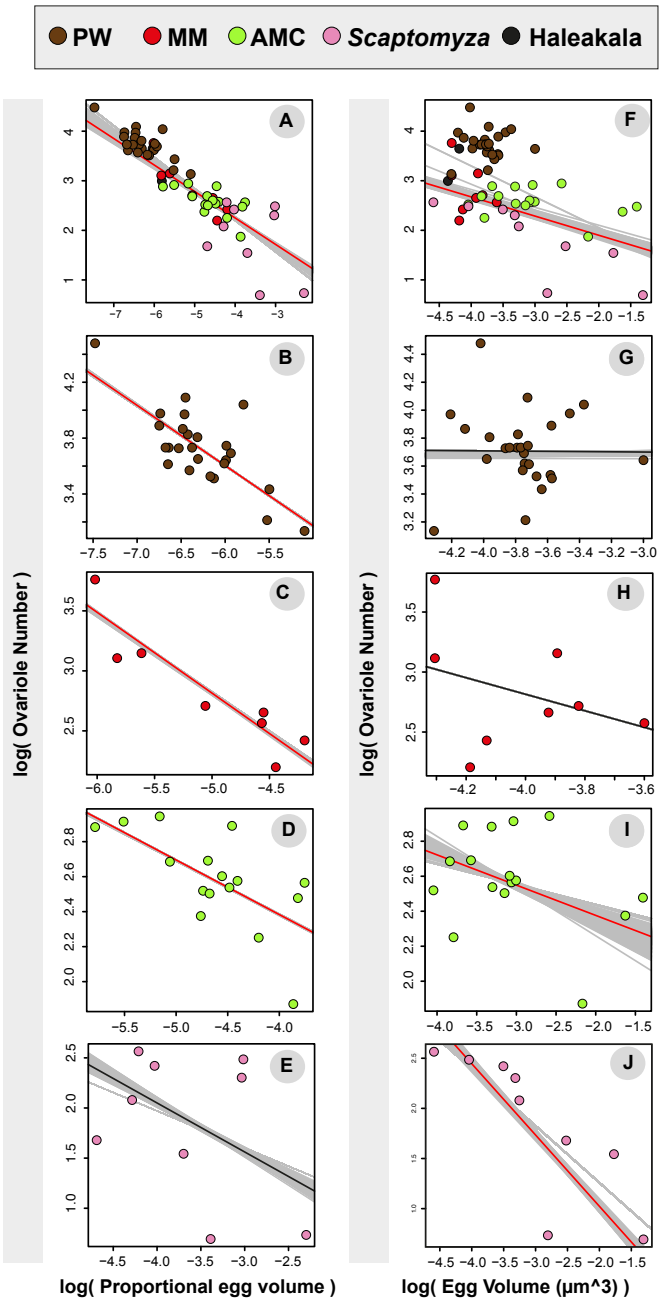


**Figure S3. Relationship between ovariole number and egg size in Hawaiian *Drosophila*.** Scatter plots of log transformed adult measurements with phylogenetically transformed trend lines generated by averaging runs from PGLS analysis across 100 posterior distribution BEAST trees, performed with the R package **nlme** v.3.1-121 (24). Trend line of the consensus tree is denoted in red when there was a significant relationship between the two traits, and black when PGLS analysis did not support a significant relationship (Table 2). (A-E) Ovariole number plotted against proportional egg volume in (A) all specimens, (B) PW, (C) AMC, (D) MM, and (E) *Scaptomyza*. (F-J) Ovariole number plotted against egg volume ($\mu m^3$) in (F) all specimens, (G) PW, (H) AMC, (I) MM, and (J) *Scaptomyza*.

**Table S1. Summary of ovariole number, body size, and egg size from Hawai'ian *Drosophila* species documented in this study, compared to a previous study**. Species name, species subgroup, and egg-laying substrate of each species is listed along with measurements of adult life history traits, along with the standard deviation and number of specimens. Left columns represent data collected in the present study, and right-hand columns represent data from Kambysellis and Heed (32). Species with overlapping data points that were not within two standard deviations are denoted in yellow. For egg volume, the length and width of the eggs were compared between the two studies, as Kambysellis and Heed (32) did not include egg volume calculations and their standard deviations. Thorax volume, calculated as the cube of thorax length, is used as a proxy for body size.

*This table is included in the "Supplemental Tables" spreadsheet.*

**Table S2. Results of best BLAST hit and parsimony analysis using mitochondrial gene sequences from field-caught females**. For each sample, which is denoted by the sample ID, the island of origin is listed. A BLASTn best hit was assigned when there was a hit to a single species that was >98%, and the accession number and species ID are listed. "*m.h.*" denotes cases where there were multiple 98-100% top hits, or there were multiple 92-97% hits. Table S3 contains an expanded list of the m.h. accession. For samples with multiple hits that fell into one specific species group, we noted the name of the species group in the column for BLASTn best hit species. "n.s." indicates samples where sequence data are not available. Parsimony tree results indicate the species group or sister species that the sample was included in by a phylogenetic analysis, for details of the analysis see methods. Final species IDs were assigned to a specific species when BLAST and phylogenetic analysis showed matching results, and were assigned to a species group when BLAST and phylogenetic analysis indicated close similarity to a group of species. For species with morphological keys, the final species ID indicates the species ID based on morphology. Species names with gray background indicate samples where the barcoding and the morphological key did not provide consistent results.

*This table is included in the "Supplemental Tables" spreadsheet.*

**Table S3. Multiple BLAST hit results from mitochondrial gene sequences of field-caught females**. The accession number and species ID for mitochondrial sequences obtained from field-caught females are listed along with percent sequence similarity to the published sequences. Note that there were several instances where the 16S sequences had 99% sequence similarity to more than 20 Hawai'ian *Drosophila* species of various groups; these were indicated as "Over 20 HI Dros spp.".

*This table is included in the "Supplemental Tables" spreadsheet.*

**Table S4. Summary of OU analysis on nuclear and mitochondrial or mitochondrial gene only BEAST trees using data only from the present study, or combined with previously published measurements.** Values are for model fit of Brownian motion (BM) and Ornstein-Uhlenbeck with one optimum (OU1) or with multiple optima (OUM) with different combinations of oviposition substrate categories, estimated with the R package **OUwie** v.1.48 (31). Oviposition substrate was categorized as follows: OU2 categorizes species that lay eggs on bark and non-bark, OU3 categorizes species into bark-breeder, spider egg and flower breeder, and other, and OU8 categorizes each species according to the eight oviposition substrates occupied by the studies considered here (bark, flower, fruit, spider egg, leaf, generalist, fungus, sap flux). Models were tested on phylogeny containing nuclear and mitochondrial gene sequences. Analyses were conducted across 100 trees from a BEAST posterior distribution using nuclear and mitochondrial genes, or mitochondrial genes only. "Combined data" includes data from our study and from that of Kambysellis and Heed (32). "Present study" indicates analyses conducted using data collected from this study alone.

| | Combined data | | | Present study | | |
|---|---|---|---|---|---|---|
| nc+mtDNA | AICc | ΔAICc | w(AIC) | AICc | ΔAICc | w(AIC) |
| BM | 83.702 | 4.399 | 0.057 | 75.796 | 5.565 | 0.046 |
| OU1 | 85.911 | 6.608 | 0.019 | 77.722 | 7.491 | 0.017 |
| OU2 | 83.037 | 3.734 | 0.081 | 76.476 | 6.245 | 0.033 |
| **OU3** | **79.303** | **0** | **0.522** | **70.231** | **0** | **0.737** |
| **OU8** | **80.285** | **0.982** | **0.312** | 73.203 | 2.972 | 0.167 |
| | Combined data | | | Present study | | |
| mtDNA | AICc | ΔAICc | w(AIC) | AICc | ΔAICc | w(AIC) |
| BM | 98.112 | 24.394 | 0 | 71.335 | 5.308 | 0.051 |
| OU1 | 97.245 | 23.527 | 0 | 73.655 | 7.628 | 0.016 |
| OU2 | 84.447 | 10.729 | 0.005 | 70.327 | 4.3 | 0.085 |
| **OU3** | **73.718** | **0** | **0.965** | **66.027** | **0** | **0.729** |
| OU8 | 80.647 | 6.929 | 0.03 | 69.666 | 3.639 | 0.118 |

**Table S5. Estimated OU theta values from OUwie analyses.** Values are the estimated OU theta values for each ecological regime, under different combinations of oviposition substrate categories, calculated using the R package **OUwie** v.1.48 (31). Oviposition substrate was categorized as follows: OU2 categorizes species that lay eggs on bark and non-bark, OU3 categorizes species into bark-breeder, spider egg and flower breeder, and other, and OU8 categorizes each species according to the eight oviposition substrates occupied by the species in our study (bark, flower, fruit, spider egg, leaf, generalist, fungus, sap flux). Table areas shaded in dark grey denote oviposition substrates that are not relevant for the particular model. Values were estimated on a phylogeny based on both nuclear and mitochondrial gene sequences. Analyses were performed across 100 trees from a BEAST posterior distribution using nuclear and mitochondrial genes. The values represent the average of the theta values and standard error over 100 trees.

| | Theta values | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 'Other' | Bark | Flower/Spider Egg | Flower | Fruit | Fungus | Generalist | Leaf | Sap | Spider eggs |
| OU2 | 7.93 ± 1.42 | 29.74 ± 1.56 | | | | | | | | |
| OU3 | 14.61 ± 1.26 | 37.17 ± 1.22 | 3.54 ± 1.42 | | | | | | | |
| OU8 | | 37.03 ± 1.17 | | 5.13 ± 1.44 | 25.57 ± 2.53 | 11.55 ± 1.66 | 13.22 ± 1.38 | 0.14 ± 7.89 | 1.79 ± 258.68 | 0.18 ± 3.62 |

**Table S6. Summary of PGLS analysis on nuclear and mitochondrial or mitochondrial gene only BEAST trees**. Relationships between ovariole number and thorax volume, ovariole number (ON) and egg volume, egg volume and thorax volume, and ovariole number and thorax/egg proportion are listed for regression analyses that were conducted across 100 trees from a BEAST posterior distribution using nuclear and mitochondrial genes or mitochondrial genes only. "Combined data" includes present data and Kambysellis and Heed (32). "Present data" indicates analyses conducted using data collected from this study alone. Minimum, average, and maximum slope and p-value for the analysis is included in the table. P-values below 0.01 are indicated in bold.

| | | nc + mtDNA | | | | | | mtDNA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Combined data | | | Present data | | | Combined data | | | Present data | | |
| | | min | avg | max | min | avg | max | min | avg | max | min | avg | max |
| ON - Thorax volume | Slope | 0.218 | 0.274 | 0.371 | 0.22 | 0.291 | 0.393 | 0.218 | 0.274 | 0.371 | 0.257 | 0.321 | 0.388 |
| | p-value | **<0.000** | **0.004** | **0.017** | **<0.000** | **0.01** | **0.04** | **<0.000** | **0.004** | 0.017 | **<0.000** | **0.004** | 0.017 |
| ON - Egg volume | Slope | -0.418 | -0.383 | -0.344 | -0.528 | -0.48 | -0.427 | -0.418 | -0.383 | -0.344 | -0.506 | -0.425 | -0.349 |
| | p-value | **<0.000** | **<0.000** | **0.001** | **<0.000** | **<0.000** | **0.001** | **<0.000** | **<0.000** | **0.001** | **<0.000** | **0.002** | **0.01** |
| Egg volume - Thorax volume | Slope | 0.285 | 0.371 | 0.426 | 0.153 | 0.201 | 0.255 | 0.285 | 0.371 | 0.426 | 0.124 | 0.199 | 0.276 |
| | p-value | **<0.000** | **<0.000** | **0.003** | 0.021 | 0.054 | 0.102 | **<0.000** | **<0.000** | **0.003** | 0.017 | 0.072 | 0.212 |
| ON - Egg/Thorax volume | Slope | -0.612 | -0.583 | -0.554 | -0.566 | -0.564 | -0.562 | -0.611 | -0.583 | -0.554 | -0.572 | -0.57 | -0.568 |
| | p-value | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** | **<0.000** |

**Table S7. Summary of TF number, TFC number per TF, and total TFC number from Hawai'ian *Drosophila* larval ovaries.** Sample size (n) and standard deviation (sd) is indicated for each species.

| Species ID | Species group | Larval oviposition substrate | TF # | | sd | n | TFC # per TF | | sd | Total TFC # | | sd | n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *D. basimacula* | AMC | Leaf | 7 | ± | 0.53 | 8 | 8.96 | ± | 1.19 | 62.59 | ± | 8.41 | 8 |
| *D. tanythrix* | AMC | Leaf | 5.75 | ± | 0.96 | 4 | 10.55 | ± | 0.86 | 61.03 | ± | 13.65 | 3 |
| *D. mimica* | MM | Fruit | 11 | ± | 1.41 | 2 | 7.5 | ± | 0.14 | 106.9 | ± | 43.7 | 2 |
| *D. mitchelli* | MM | Unknown | 11.5 | ± | 0.71 | 2 | 7.8 | ± | 0.57 | 89.5 | ± | 0.99 | 2 |
| *D. grimshawi* | PW | Bark | 25.25 | ± | 3.96 | 5 | 9.58 | ± | 1.39 | 193.47 | ± | 57.89 | 8 |
| *D. hawaiiensis* | PW | Sap flux | 17.67 | ± | 2.08 | 3 | 9.47 | ± | 1.04 | 167.97 | ± | 33.86 | 3 |
| *D. picticornis* | PW | Sap flux | 13 | ± | 2.83 | 2 | 11.18 | ± | 1.23 | 144.84 | ± | 17.48 | 9 |
| *D. setosimentum* | PW | Bark | 19 | ± | 2.82 | 2 | 9.87 | ± | 0.15 | 186.9 | ± | 20.63 | 3 |
| *D. silvestris* | PW | Bark | 18 | ± | 0 | 1 | 12 | ± | 2.68 | 216 | ± | 48.37 | 2 |
| *D. villocipedis* | PW | Bark | 29 | ± | 1.26 | 4 | 10.02 | ± | 0.43 | 218.01 | ± | 14.92 | 6 |
| *Bunostoma spp* | Scaptomyza | Unknown | 6 | ± | 0 | 2 | 9.4 | ± | 1.56 | 56.5 | ± | 9.19 | 2 |
| *S. caliginosa* | Scaptomyza | Flower | 2 | ± | 0 | 2 | 7.75 | ± | 0.354 | 15.5 | ± | 0.71 | 2 |

**Table S8. Comparison of larval TF number and adult ovariole number.** Number of TFs observed per developing ovary compared to ovariole number per ovary, and the ratio of TF number to ovariole number for various species.

| Species | Species group | TF number per ovary | n | Ovariole number per ovary | n | TF:ON Ratio |
|---|---|---|---|---|---|---|
| *D. basimacula* | AMC | 7 | 8 | 6.2 | 21 | 1.13 |
| *D. tanythrix* | AMC | 5.75 | 4 | 6 | 12 | 0.96 |
| *D. mimica* | MM | 11 | 2 | 11 | 5 | 1 |
| *D. mitchelli* | MM | 11.5 | 2 | 11.2 | 3 | 1.03 |
| *D. grimshawi* | PW | 19.5 | 4 | 20.5 | 12 | 0.95 |
| *D. hawaiiensis* | PW | 17.6 | 3 | 17.58 | 13 | 1 |
| *D. picticornis* | PW | 13 | 2 | 13.3 | 10 | 0.98 |
| *D. setocimentum* | PW | 19 | 2 | 20.6 | 8 | 0.92 |
| *D. villocipedis* | PW | 22 | 2 | 18.5 | 8 | 1.19 |
| *Bunostoma spp* | Scaptomyza | 6 | 2 | 7.4 | 4 | 0.81 |
| *S. caliginosa* | Scaptomyza | 2 | 2 | 1 | 4 | 2 |
| | | | | | Total | 1.11 |

**Table S9. List of mitochondrial COI sequences included in the phylogenetic inference analysis.** Mitochondrial sequences were obtained from field-caught Hawai'ian *Drosophila* that were identified based on morphological keys.

*This table is included in the "Supplemental Tables" spreadsheet.*

**Table S10. Comparison of AICc and weighted AICc values for models testing the relationship between oviposition substrate and egg volume, residual egg volume and thorax volume.** Values are for model fit of Brownian motion (BM) and Ornstein-Uhlenbeck with one optima (OU1) or with multiple optimum (OUM) with different combination of oviposition substrate categories, calculated with the R package **OUwie** v.1.48 (31). Oviposition substrates were categorized as follows: OU2 categorizes species that lay eggs on bark and non-bark; OU3 categorizes species into bark-breeder, spider egg/flower breeder, and other; and OU8 categorizes each species according to the eight oviposition substrates represented (bark, flower, spider egg, fruit, leaf, generalist, fungus, sap flux). Models were tested over 100 posterior distribution BEAST trees using nuclear and mitochondrial gene sequences.

| | Egg volume | | | Residual egg volume | | | Thorax volume | | |
|---|---|---|---|---|---|---|---|---|---|
| nc+mtDNA | AICc | ΔAICc | w(AIC) | AICc | ΔAICc | w(AIC) | AICc | ΔAICc | w(AIC) |
| BM | 82.31 | 8.74 | 0.008 | 76.03 | 5.72 | 0.0322 | **117.92** | **0** | **0.588** |
| OU1 | **73.57** | **0** | **0.593** | 75.48 | 5.17 | 0.042 | 120.14 | 2.21 | 0.194 |
| OU2 | 81.27 | 7.7 | 0.013 | 72.67 | 2.36 | 0.173 | 125.15 | 7.23 | 0.016 |
| OU3 | 77.38 | 3.81 | 0.088 | 72.5 | 2.19 | 0.189 | 122.99 | 5.08 | 0.046 |
| OU7 | **74.94** | **1.37** | **0.299** | **70.31** | **0** | **0.564** | 120.58 | 2.66 | 0.156 |

**References for SI reference citations**


1. Miles CM*, et al.* (2011) Artificial selection on egg size perturbs early pattern formation in *Drosophila melanogaster*. *Evolution* 65(1):33-42.
2. Sarikaya DP*, et al.* (2012) The roles of cell size and cell number in determining ovariole number in *Drosophila*. *Dev. Biol.* 363:279-289
3. Magnacca KN & O'Grady PM (2009) *Revision of the Modified Mouthparts Species Group of Hawaiian Drosophila (Diptera: Drosophilidae). The Ceratostoma, Freycinetiae, Semifuscata, and Setiger Subgroups, and Unplaced Species* (University of California Press, Ltd., London, England) p 104.
4. Magnacca KN & O'Grady PM (2006) A Subgroup Structure for the Modified Mouthparts Species Group of Hawaiian Drosophila. *Proceedings Of The Hawaiian Entomological Society* 38:87-101.
5. Ross HA, Murugan S, & Li WLS (2008) Testing the reliability of genetic methods of species identification via simulation. *Syst. Biol.* 57(2):216-230.
6. O'Grady PM*, et al.* (2011) Phylogenetic and ecological relationships of the Hawaiian Drosophila inferred by mitochondrial DNA analysis. *Mol Phylogenet Evol* 58(2):244-256.
7. Lapoint RT, Magnacca KM, & O'Grady PM (2014) Phylogenetics of the Antopocerus-Modified Tarsus Clade of Hawaiian Drosophila: Diversification across the Hawaiian Islands. *PLoS ONE* 9(11):e113227.
8. Lapoint RT, O'Grady PM, & Whiteman NK (2013) Diversification and dispersal of the Hawaiian Drosophilidae: the evolution of Scaptomyza. *Mol Phylogenet Evol* 69(1):95-108.
9. Lapoint RT, Gidaya A, & O'Grady PM (2011) Phylogenetic relationships in the spoon tarsus subgroup of Hawaiian Drosophila: conflict and concordance between gene trees. *Mol Phylogenet Evol* 58(3):492-501.
10. 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.
11. O'Grady P & Desalle R (2008) Out of Hawaii: the origin and biogeography of the genus Scaptomyza (Diptera: Drosophilidae). *Biol. Lett.* 4(2):195-199.
12. Smith SA & Dunn CW (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* 24(5):715-716.
13. Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30(4):772-780.
14. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17(4):540-552.
15. Lanfear R, Calcott B, Ho SYW, & Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Mol. Biol. Evol.* 29(6):1695-1701.
16. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688-2690.
17. Bouckaert R*, et al.* (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *Plos Computational Biology* 10(4).

18. Drummond AJ & Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7.
19. Drummond AJ, Ho SYW, Phillips MJ, & Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4(5):699-710.
20. Nylander JA, Wilgenbusch JC, Warren DL, & Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24(4):581-583.
21. Rambaut A & Drummond AJ (2003) Tracer: MCMC Trace Analysis Tool. *University of Oxford* Oxford.
22. Rambaut A & Drummond AJ (2013) TreeAnnotator v1. 7.0. *http://beast.community/treeannotator*.
23. Team RC (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
24. Pinheiro J, Bates D, DebRoy S, Sarkar D, & Team RC (2014) Nlme: Linear and Nonlinear Mixed Effects Models. *R Package Version 3.1-117* https://CRAN.R-project.org/package=nlme.
25. Paradis E, Claude J, & Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics (Oxford, England)* 20(2):289-290.
26. Martins EP & Hansen TF (1997) Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* 149(4):646-667.
27. Magnacca K, Foote D, & O'Grady PM (2008) A review of the endemic Hawaiian Drosophilidae and their host plants. *Zootaxa* 1728:1-58.
28. Kambysellis MP*, et al.* (1995) Pattern of ecological shifts in the diversification of Hawaiian *Drosophila* inferred from a molecular phylogeny. *Curr. Biol.* 5(10):1129-1139.
29. Craddock EM & Kambysellis MP (1997) Adaptive Radiation in the Hawaiian Drosophila (Diptera: Drosophilidae): Ecological and Reproductive Character Analyses. *Pacific Science* 51(4):475-489.
30. Beaulieu JM, Oliver JC, & O'Meara B (2012) corHMM: Hidden Markov Models in R, Version 1.0.
31. Beaulieu JM & O'Meara B (2014) OUwie: Analysis of Evolutionary Rates in an OU Framework. *R package version 1*.
32. Kambysellis MP & Heed WB (1971) Studies of Oogenesis in Natural Populations of Drosophilidae. I. Relation of ovarian development and ecological habitats of the hawaiian species. *Am. Nat.* 941(105):31-49.
33. Katoh T, Izumitani HF, Yamashita S, & Watada M (2017) Multiple origins of Hawaiian drosophilids: Phylogeography of Scaptomyza Hardy (Diptera: Drosophilidae). *Entomological Science* 20(1):33-44.
34. Stearns SC (1989) Trade-Offs in Life-History Evolution. *Funct. Ecol.* 3(3):259-268.
35. Fox CW & Czesak ME (2000) Evolutionary ecology of progeny size in arthropods. *Annu. Rev. Entomol.* 45:341-369.
36. Lavista-Llanos S*, et al.* (2014) Dopamine drives Drosophila sechellia adaptation to its toxic host. *eLife* 3:e03785.

37. Berrigan D (1991) The allometry of egg size and number in insects. *OIKOS* 60:313-321.
38. Revell LJ (2010) Phylogenetic signal and linear regression on species data. *Methods Ecol Evol* 1(4):319-329.
39. Hodin J & Riddiford LM (2000) Parallel alterations in the timing of ovarian ecdysone receptor and ultraspiracle expression characterize the independent evolution of larval reproduction in two species of gall midges (Diptera: Cecidomyiidae). *Dev. Genes Evol.* 210(7):358-372.
40. Green II DA & Extavour CG (2012) Convergent Evolution of a Reproductive Trait Through Distinct Developmental Mechanisms in *Drosophila*. *Dev. Biol.* 372(1):120-130.
41. Ronai I, Barton DA, Oldroyd BP, & Vergoz V (2015) Regulation of oogenesis in honey bee workers via programed cell death. *J. Insect Physiol.* 81:36-41.