Complexities of recapitulating polygenic effects in natural populations: replication of genetic effects on wing shape in artificially selected and wild caught populations of *Drosophila melanogaster.*

Katie Pelletier*[1], William R. Pitchers[2]*, Anna Mammel[2,7], Emmalee Northrop-Albrecht[2,5], Eladio J. Márquez[3,6], Rosa A. Moscarella[3,4], David Houle[3], Ian Dworkin[1,2]
*Co-first authors
Corresponding authors: dworkin@mcmaster.ca, dhoule@bio.fsu.edu

Author affiliations: 1- Department of Biology, McMaster University

2 - Department of Integrative Biology, Michigan State University

3- Department of Biological Science, Florida State University

4-Current address: Department of Biology, University of Massachusetts

5- Current address Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester MN

6 – Current address: Dragonfly Therapeutics, Waltham MA

7-Current address: Neurocode USA. 3548 Meridian St, Bellingham, WA 98225

Abstract

Identifying the genetic architecture of complex traits is important to many geneticists, including those interested in human disease, plant and animal breeding, and evolutionary genetics. Advances in sequencing technology and statistical methods for genome-wide association studies (GWAS) have allowed for the identification of more variants with smaller effect sizes, however, many of these identified polymorphisms fail to be replicated in subsequent studies. In addition to sampling variation, this failure to replicate reflects the complexities introduced by factors including environmental variation, genetic background, and differences in allele frequencies among populations. Using *Drosophila melanogaster* wing shape, we ask if we can replicate allelic effects of polymorphisms first identified in a GWAS (Pitchers et al. 2019) in three genes: *dachsous (ds)*, *extra-macrochaete (emc)* and *neuralized (neur),* using artificial selection in the lab and bulk segregant mapping in natural populations. We demonstrate that multivariate wing shape changes associated with these genes are aligned with major axes of phenotypic and genetic variation in natural populations. Following seven generations of artificial selection along the *ds* shape change vector, we observe genetic differentiation of variants in *ds* and genomic regions containing other genes in the hippo signaling pathway. This suggests a shared direction of effects within a developmental network. We also performed artificial selection with the *emc* shape change vector, which is not a part of the hippo signaling network, but showed a largely shared direction of effects. The response to selection along the *emc* vector was similar to that of *ds*, suggesting that the available genetic diversity of a population, summarized by the genetic (co)variance matrix (**G**), influenced alleles captured by selection. Despite the success with artificial selection, bulk segregant analysis using natural populations did not detect these same variants, likely due to the contribution of environmental variation and low minor allele frequencies, coupled with small effect sizes of the contributing variants.

Introduction

Dissecting the genetic architecture underlying complex traits remains challenging because of the joint contributions of many alleles of small effect and genotype-by-environment interactions, among other factors. Progress in sequencing technology in conjunction with development of GWAS statistical methodologies has enabled identification of loci contributing to numerous complex traits and diseases. However, such mapping approaches identify only a subset of loci contributing to trait variation (Visscher et al., 2017). In part, this reflects the low power to detect rare alleles and those with small effects (Tam et al., 2019). For alleles that are relatively common in a population, replication rates between GWAS studies are high, even when effect sizes are small (Marigorta et al., 2018). However, GWAS studies have failed to replicate the effects observed in many candidate gene studies, in part due to the fact that many of the alleles identified in these studies are rare in a population and require very large cohorts to detect (Fritsche et al., 2016; Ioannidis et al., 2011).

In cases where an association is replicated between studies, the magnitude of the effect can vary substantially between different cohorts or populations (CONVERGE consortium, 2015; Marigorta et al., 2018). Differences can arise because of genetic background due to epistatic gene by gene (*GxG*) interactions, or due to gene-by-environment (GxE) interactions. The initial estimates of effect size will be biased upwards if statistical testing in the initial cohort is used to determine which SNPs are chosen for replication studies. It is important to understand which of these causes of differences in effect size are of practical significance when we want to generalize results to different populations or environments.

In this study, we focus on the issue of replication in a multivariate context, where the joint inheritance of multiple features are simultaneously investigated. We will refer to the suite of measured features as a 'multivariate trait' for convenience. In this case, what we want to estimate is the vector of effects of each SNP on all measured features. Each SNP may have a unique combination of effects. Univariate effects vary only in magnitude, as we can only infer effects on a single feature. For a multivariate trait, estimated genetic effects vary in magnitude, the sum of effects on all traits, and also in direction, how the total effect is allocated among different features (Melo et al., 2019). The ability to study the direction along with the

magnitude of genetic effects provides an additional and important way of assessing repeatability. For a univariate trait, there is a 50% chance that the replicate estimate will be in the same direction as the original estimate, even with no true effect. By contrast, the probability of a "replicated" genetic effect sharing a similar direction by chance alone decreases as the number of measured features increases (Marquez and Houle, 2015; Stephens, 2013).

Studying genetic effects in a multivariate context is beneficial in other ways. First, it has been demonstrated both empirically and via simulations, that genetic mapping for multivariate traits generally increases statistical power over trait by trait analyses (Fatumo et al., 2019; Pitchers et al., 2019; Porter and O'Reilly, 2017; Shriner, 2012). Second, some multivariate traits cannot be sensibly reduced to a single measurement. The wing shape we study is a great example of such a multivariate trait. We have good reason to believe that wing shape is important for flight (Ray et al., 2016), but we cannot yet say that any feature, such as wing length or width, is more or less important than any other. Natural selection on wing shape may affect any or all combinations of measurements.

Perhaps most importantly, traits are not inherited in isolation, but are the joint outcome of an integrated developmental process that results in extensive genetic correlations that can have important effects on evolution. The main source of such correlations are the patterns of pleiotropic effects generated by mutational effects. Multivariate studies of inheritance allow pleiotropic effects to be estimated in a rigorous and justifiable manner (Melo et al., 2019). The multivariate breeder's equation, $\Delta \boldsymbol{z} = \boldsymbol{G\beta}$, enables short term prediction of evolutionary responses. Key to understanding how populations respond to selection in the short term requires an understanding of properties of the genetic (co)variance matrix (**G**), and in particular the axis of greatest genetic variation, $\mathbf{g}_{max}$. Studies demonstrate that the direction of $\mathbf{g}_{max}$ influences evolutionary trajectories (Blows and McGuigan, 2015; McGuigan, 2006; Schluter, 1996). The degree to which genetic effects associated with particular variants align to major axes of genetic (co)variance, expressed through **G**, may provide insights into which alleles are most likely to be "captured" by selection (Pitchers et al., 2019).

Due to the polygenic nature of complex traits, including multivariate ones, it is important to consider not only the direction of effect for alleles in a single gene but also

107 correlated effects between genes contributing to the phenotype. Interestingly, initial
108 comparisons of directions of genetic effects among induced mutations in two *Drosophila*
109 *melanogaster* wing development pathways showed only partially correlated effects on wing
110 shape within and between pathways (Dworkin and Gibson, 2006). However, recent work has
111 demonstrated that despite large differences in magnitude, the direction of genetic effects of
112 variants segregating in populations are sometimes similar to those from validation experiments
113 using RNAi knockdown of those same genes (Pitchers et al., 2019). Additionally, Pitchers et al.
114 (2019) demonstrated this shared direction of effect could also be shared between a SNP and
115 RNAi knockdown of other genes in the same signaling pathway, such as those involved with
116 hippo signaling, a key pathway involved with wing growth and morphogenesis (Pan et al.,
117 2018).

118 Pitchers et al. (2019) identified over 500 polymorphisms contributing to wing shape
119 variation in the *Drosophila* genetic resource panel (DGRP). Among these, the hippo pathway
120 was over-represented in SNPs associated with wing shape (Pitchers et al., 2019). The degree to
121 which identified hippo pathway variants reflect allele specific effects, differences in magnitude
122 of genetic effects, and even the large statistical uncertainty associated with genetic effects of
123 small magnitude are unclear. Given common dominance patterns, and the likely non-linear
124 genotype-phenotype relationships of most genetic effects, small to moderate changes in gene
125 function may result in modest phenotypic effects (Green et al., 2017; Melo et al., 2019; Wright,
126 1934). Large effect mutants and many RNAi knockdown studies have moderate to large
127 phenotypic effects that are not reflective of the magnitude of genetic effects of SNPs
128 contributing to phenotypic variance in natural populations.

129 The expression of genetic effects is also dependent on genetic and environmental
130 context, with contributions of gene by gene (GxG) and gene-by-environment (GxE) sources of
131 variation. The context-dependence of genetic effects for a multivariate trait has been
132 demonstrated for *Drosophila* wing shape. Variants in *Egfr* influencing *Drosophila* wing shape are
133 replicable in both lab reared and wild-caught cohorts (Dworkin et al., 2005; Palsson et al., 2005;
134 Palsson and Gibson, 2004). However, in replication studies, effect sizes of alleles were
135 diminished in both outbred populations and wild cohorts. In the latter case the same variant

136 explained 1/10 of the phenotypic variance explained in the initial study (Dworkin et al., 2005).
137 Interestingly, in a series of experimental crosses among strains, the effects of the SNP were
138 replicable for direction and magnitude in multiple experimental assays and crossing schemes.
139 Despite this, the genetic effect on wing shape from this SNP largely disappeared in one natural
140 population (Palsson et al., 2005). A number of reasons have been proposed for the failure to
141 replicate genetic effects including environmental effects, differences between controlled lab
142 and natural environments (Dworkin et al., 2005), and genetic background (Greene et al., 2009),
143 among others. Because both environment and genetic background likely affect the genotype-
144 phenotype map in a non-linear fashion (Wright, 1934), it is important to test observed
145 associations in other experimental contexts.
146 A promising approach to confirm the estimated effects of candidate genetic variants is
147 to test whether they respond to artificial selection in the direction of the inferred effect. This
148 approach is particularly relevant to evolutionary questions, but has rarely been used. In this
149 study, we use both artificial selection and bulk segregant analysis to replicate and validate SNPs
150 associated with three genes: *dachsous (ds), extra-macrochetae (emc)* and *neutralized (neur)*
151 previously identified in a GWAS of *Drosophila* wing shape (Pitchers et al., 2019). Using the
152 vectors of shape change based on RNAi knockdowns of each gene, we demonstrate that the
153 direction of shape change for these genetic effects is aligned with major axes of natural
154 phenotypic and genetic variation. Using artificial selection based on the direction of shape
155 change defined by RNAi knockdown, we were able to replicate the effects observed for *ds,* but
156 not *emc*, likely due to the available genetic diversity in the population. We then asked if these
157 effects could be replicated in a natural population using a bulk segregant approach. We discuss
158 our results in the context of the replicability of genetic effects and the shared direction of
159 genetic effects due to shared developmental processes.
160
161 **Methods**
162 **Source Populations and phenotypic analysis**
163 *Drosophila* strains

164 Phenotype data for the *Drosophila* genetic resource panel (DGRP) was collected for 184
165 strains as part of a GWAS study as described in Pitchers *et al* (2019). Genotype data for these
166 strains was obtained from freeze 2 of the DGRP (Huang et al., 2014). For replication using
167 artificial selection, 30 DGRP strains were used: DGRP-149, 324, 383, 486, 563, 714, 761, 787,
168 796, 801, 819, 821, 822, 832, 843, 849, 850, 853, 859, 861, 879, 887, 897, 900, 907, 911, 913.
169 These strains were selected to increase genetic variation at the *ds* locus (Supplemental Figure 1,
170 Table 1). Reciprocal pairwise crosses between the 30 selected DGRP strains were used to create
171 heterozygotes and these 30 heterozygous genotypes were successively pooled for 4
172 subsequent generations, allowing for recombination. After pooling, the synthetic outbred
173 population was maintained for approximately 47 subsequent generations (allowing for
174 recombination) before the start of artificial selection experiments.
175 For the replication in wild-caught populations using bulk segregant analysis (BSA),
176 individuals were collected via sweep-netting from orchards and vineyards in Michigan and after
177 species identification, stored in 70% ethanol. In 2013 and 2014, cohorts were collected from
178 Fenn Valley Winery (FVW13 and FVW14 respectively, GPS coordinates: 42.578919, -
179 86.144936). Additionally in 2014, cohorts were collected from Country Mill Orchard (CMO, GPS
180 coordinates: 42.635270, -84.796706), and Phillip's Hill Orchard (PHO, GPS coordinates:
181 43.117981, -84.624235). For all collected cohorts, except for the FVW14 collection, only males
182 were used in this study given difficulties differentiating *Drosophila melanogaster* and *D.*
183 *simulans* females. For the genomic analysis of the FVW14 wild caught population (below) we
184 utilized both males and females as the number of individuals was insufficient otherwise. For the
185 collection where females were included in the study, there is no evidence of contamination
186 with *D. simulans* as all dissected wings were classified as *D. melanogaster* using linear
187 discriminant analysis (LDA). The LDA algorithm was trained using male wings from the collected
188 *D. melanogaster* data set as well as males from a *D. simulans* data set acquired from the
189 Hosken lab. There was 100% agreement between the classification of females within each
190 species with our phenotypic classification, indicating that it is very unlikely that *D. simulans*
191 females were included in our samples (Supplemental Figure 2).
192 <u>Morphometric Data</u>

193 Landmark and semi-landmark data were captured from black and white TIFF images
194 using the pipeline described in Houle et al. (2003). First, two landmark locations, the humeral
195 break and alula notch, were digitized using tpsDig2 (version 2.16). Wings (Van der Linde 2004–
196 2014 , v3.72) software was used to fit nine cubic B-splines and correct errors. All shape data
197 was subjected to full Procrustes superimposition to scale images to the same size, translate
198 configurations to the same location and rotate to minimize distance between points (Rohlf and
199 Slice, 1990). Generalized Procrustes superimposition (registration) and extraction of 14
200 landmark and 34 semi landmark positions was done with the program CPR v1.11 (Márquez
201 2012–2014, Figure 1). Sliding of semi-landmarks utilized minimization of Procrustes Distance as
202 the objective function. Superimposition results in the loss of 4 possible dimensions of variation
203 while semi-landmarks are constrained to vary over only one axis, restraining these points to
204 approximately a single dimension of variation. This results in a total of ~58 available dimensions
205 of shape variation, that can be summarized using the first 58 Principal components (PCs).
206 Allometry was adjusted for in the analysis by fitting a model for landmark coordinates onto
207 centroid size, and using the residuals from this model (Klingenberg, 2022). By accounting for
208 the allometric component of shape, shape variation due to size variation is accounted for
209 (Supplemental Figure 3). For most analyses, 'allometry corrected' landmarks were used, with
210 the exception of shape models fit using the Geomorph package in R, where Procrustes
211 landmarks were used and centroid size was included as a predictor in the model.
212 <u>Generation of shape vectors for artificial selection and Bulk segregant analysis</u>
213 A panel of shape change vectors was estimated using the progesterone-inducible
214 Geneswitch GAL4 construct under the regulation of an ubiquitous *tubulin* driver, to drive the
215 expression of RNAi for genes of interest, as previously described in Pitchers *et al.,* 2019. GAL4
216 expression was induced throughout larval development by adding mifepristone, an analog of
217 progesterone, to the larval food. Knockdown was varied by assaying phenotypes at
218 mifepristone concentrations of 0.3, 0.9, and 2.7 μM, plus a control without mifepristone. Wing
219 shape change associated with knockdown of the gene of interest was estimated using
220 multivariate regression of shape on concentration of mifepristone. Shape change vectors for
221 *dachsous* (*ds*)*, extramacrochaete* (*emc*) and *neuralized* (*neur*) were used in this experiment

222 (Figure 1b, Supplemental Figure 4). The magnitude ('length') of the vector measures how much
223 shape change occurs per unit change in mifepristone. In general, vectors of greater magnitudes
224 enable better estimate of direction of effect for shape change. As reported in Pitchers et al.,
225 (2019), the magnitude (l^2-*norm*) of vectors for RNAi knockdown of these genes are 5.5 for *ds*,
226 2.8 for *neur*, and 0.44 for *emc*.
227 Shape data collected as part of a previous study (Pitchers et al., 2019) was used to
228 assess the relationships between shape change vectors from the RNAi titration and $\mathbf{g}_{\mathbf{max}}$, the
229 first eigenvector of the **G** matrix estimated from DGRP line means. The effects of sex, centroid
230 size and their interaction were removed using a linear model and these residuals were used to
231 calculate shape score by projecting the data (see Supplemental Figure 5) onto the shape change
232 vector estimated in each knockdown experiment. To assess major axes of genetic variation
233 among the DGRP strains, principal component analysis was performed on the allometry
234 adjusted model residuals (Supplemental Figure 5B). PCA was done in a similar manner for the
235 individuals from the wild caught cohorts. Correlations between the first three eigenvectors
236 ("genetic PCs" including $\mathbf{g}_{\mathbf{max}}$), the first three PCs from the wild caught cohorts and the shape
237 scores for *ds*, *emc* and *neur* were calculated (Figure 1a, Supplemental Figure 5). From this, *ds,*
238 *emc* and *neur* shape change vectors were selected for further experiments given high
239 correlation with directions of natural genetic variation (Figure 1, Supplemental Figure 5).
240 <u>Artificial selection of synthetic outbred population</u>
241 The synthetic outbred population resulting from pooling DGRP lines was used as the
242 parent population for artificial selection. Both the *ds* and *emc* artificial selection experiment
243 was carried out with three independent replicates of each "up" and "down" selection regimes,
244 along with unselected control lineages. Each generation, wings of live flies were imaged using
245 the 'wingmachine' system and shape data collected (Houle et al., 2003, Van der Linde 2004–
246 2014 ,v3.72). Shape scores were calculated by projecting the data onto the *ds* or *emc* shape
247 change vector as described above, and the 40 individuals each with highest or lowest shape
248 scores, were selected to found the next generation (Supplemental Figure 5A). For the control
249 lineages, 40 individuals were randomly selected for the next generation within each replicate
250 lineage. Following seven generations of selection, 75 individuals from each lineage were

251 selected for pooled sequencing, described below. The response to selection was evaluated both
252 by computing Procrustes distance (PD) between average shape of wings between generations
253 one and seven, and using shape scores (projections) with a linear mixed effect model allowing
254 for the fixed effect factors of treatment and sex, continuous predictors of centroid size and
255 generation, with third order interactions among these effects. The effect of generation was
256 allowed to vary by replicate lineages ($\mathrm{lmer(ds \sim (CS + Sex + line + gen0)\text{^}3 + (1 +}$
257 $\mathrm{gen0|line{:}rep))}$). Realized heritabilities were estimated separately for up and down selection
258 lineages, from the slope of the regression of cumulative selection differentials on cumulative
259 selection response, averaging over sex and with a random effect of replicate lineage.
260 Wild populations
261 For the Bulk segregant Analysis (BSA), wings for wild caught individuals were dissected
262 and mounted in 70% glycerol in PBS. Images of wings were captured using an Olympus DP30B
263 camera mounted on an Olympus BX51 microscope (Olympus software V.3,1,1208) at 20X
264 magnification. When possible, both left and right wings were dissected, imaged and averaged
265 to calculate an individual's mean shape. For some individuals a wing was damaged so only one
266 wing could be used. Shape was captured as described above. The total number of individuals
267 phenotyped from each cohort can be found in Supplemental table 1.
268 To remove allometric effects in the data, shape was regressed onto centroid size and
269 the model residuals were used for all subsequent morphometric analysis. Only data from males
270 was used to compare shape in wild populations, although, including females from the FVW14
271 population and regressing shape onto centroid size and sex gave equivalent results
272 (Supplemental Figure 6). To test for shape differences between collection cohorts, the effect of
273 centroid size and collection cohort on shape were modeled (procD.lm(shape ~ CS + pop_year))
274 using the procD.lm function in Geomorph v 3.1.3. (Adams and Otárola-Castillo, 2013) and
275 distances between populations were calculated using the pairwise function. To select
276 individuals for sequencing, a 'shape score' was calculated using the method described above.
277 Shape data was projected onto the vector of shape change defined by the *ds* or *neur*
278 knockdowns. The *emc* projection vector was not used for bulk segregant analysis due to the
279 high similarity with *ds* shape change (Figure 1), and the similarity of the selection response. Its

280 inclusion would result in selection of largely the same cohorts of individuals for sequencing for
281 both *ds* and *emc*. As an alternative, we utilized the *neur* shape vector as it was largely
282 uncorrelated with that of *emc* and *ds*, but strongly correlated with natural variation of shape.
283 The 75 most extreme individuals within each wild-caught cohort on the shape score distribution
284 were selected for pooled sequencing. Allele frequencies within each population was estimated
285 by sequencing 75 random individuals within each cohort. The difference vector between mean
286 shapes of selected pools (within each population) was used to calculate Procrustes distance
287 (PD) between pools and the correlation of this shape change vector with the selection vector
288 used. An estimate of genetic distances between populations was calculated using allele
289 frequencies (mapping pipeline described below) in the pools of the 75 randomly selected
290 individuals using Bray's distance with the vegdist() function from the vegan package (v2.6-2) in
291 R.

292 Sequencing and Genomic Analysis

293 DNA extractions from pools of selected individuals was performed using a Quiagen
294 DNeasy DNA extraction kit. Library preparation and Illumina sequencing was performed at the
295 research technology support facility at Michigan State University. All library samples were
296 prepared using the Rubicon ThruPLEX DNA Library Preparation kit, without a procedure for
297 automatic size selection of samples. Paired end libraries (150bp) were sequenced using Illumina
298 HiSeq 2500, with each sample (either one pool of 75 individuals in the BSA or one pooled
299 replicate lineage in the artificial selection) being run on two lanes.

300 Reads were trimmed with Trimmomatic (v0.36) to remove adapter contamination and
301 checked for quality using FastQC prior to alignment (Bolger et al., 2014). Trimmed reads were
302 aligned to the *Drosophila melanogaster* genome (v6.23) using BWA-MEM (v0.7.8)(Li and
303 Durbin, 2010). Sequencing replicates of the same biological samples were merged using
304 SAMtools (v1.11). PCR duplicates were removed using Picard with the MarkDuplicates tool (v
305 2.10.3) and reads with a mapping quality score less than 20 were removed using SAMtools (Li et
306 al., 2009). A local realignment around indels was performed using GATK using the
307 IndelRealigner tool (v3.4.46). For artificial selection experiments, reads were merged for all up,
308 down and control selection lines as replicates lineages were independent. For wild cohorts,

309 pools were not merged between populations. mpileup files were created using SAMtools and
310 used for subsequent genomic analysis. Highly repetitive regions of the *Drosophila* genome were
311 identified and subsequently masked in mpileup files using RepeatMasker (v4.1.1) with default
312 settings. INDELs and regions within 5bp of an indel were identified and masked using
313 popoolation2 scripts. Population genetic statistics were calculated using PoPoolation (v1.2.2)
314 and PoPoolation2 (v1.201) (Kofler et al., 2011b, 2011a).

315 For the bulk segregant analysis (BSA) in the wild-caught cohorts, a modified Cochran-
316 Mantel-Haenszel (CMH) test was used to measure significantly differentiated sites between
317 pools of individuals. Sampling effects were accounted for using the ACER package (v.1.0) in R,
318 assuming $N_e = 10^6$ with 0 generations of differentiation between selected pools (Spitzer et al.,
319 2020). To adjust for multiple testing, the p-value was corrected using a Benjamini and Hochberg
320 correction (Benjamini and Hochberg, 1995) with an adjusted alpha of 0.05. For each significant
321 site from the CMH test, using an adjusted p-value cut-off of 0.05, we identified the nearest
322 gene using BEDtools (v2.19.1) (Quinlan and Hall, 2010). In addition, to account for sampling
323 variation, we sampled genomic coverage to 75x for all samples, dropping sites that did not
324 meet this threshold and repeating the CMH test. We confirmed that there was no association
325 between genetic and shape differentiation between populations, and that the populations do
326 not show strong phenotypic differentiation based on either overall shape variation, or shape
327 scores used to identify selected individuals for BSA (Supplemental Figures 3 and 7). There was
328 some variation among populations in overall wing size (Supplemental Figure 8), however we
329 (assuming common allometry) adjusted for allometric effects on shape.

330 For artificial selection experiments, F_{ST} was calculated in 5000bp windows. We chose
331 this window size as it is expected that blocks of LD in the synthetic outbred population will be
332 much larger in comparison to that of the wild caught samples (King et al., 2012a; King et al.,
333 2012b; Marriage et al., 2014). This statistic was used to compare the "up" selected pools to the
334 "down" selected pools to help identify regions of differentiation between selected populations.

335 For the artificial selection comparisons, genes in regions of high F_{ST} were identified by
336 finding overlaps between outlier windows and annotated *Drosophila* genes using
337 GenomicRanges (v1.46.1) in Bioconductor. High F_{ST} was defined as F_{ST} values greater than three

338 standard deviations above the mean. GO terms associated with identified genes were
339 annotated using TopGO package (v2.34.0) (Alexa et al., 2006) in Bioconductor. GO enrichment
340 was then performed to identify those terms overrepresented in the identified list using TopGO
341 and a Fisher's exact test. Over representation of 2 GO terms in outlier windows (hippo
342 signaling, GO:0035329; negative regulation of hippo signaling GO:0035331) were tested using a
343 permutation test that randomly sampled genomic windows from the total windows for which
344 F_{ST} was calculated and the permutation was run 1000 times. The distribution of the ratio of
345 observed to expected genes annotated with the term of interest within randomly sampled
346 regions was compared to the number observed in the data.

347 Verification of *ds* indel in DGRP

348 Sanger sequencing was performed on individuals from a cross between DGRP lines predicted to
349 have the polymorphism (DGRP 195, 28, 96, 48, 59, 801) and those without (DGRP 129, 301, 69,
350 385, 75, 83, 491, 34, 774) crossed to a line carrying a deletion in the region of interest (BDSC
351 24960) to account for potential residual heterozygosity in otherwise inbred strains. DNA was
352 prepared by incubating flies in DNA extraction buffer (1mM EDTA, 25mM NaCl, 10 mM TrisHCl
353 pH 7.5) for 10 minutes, followed by storage at -20 C. PCR application of the region of interest
354 (Forward primer ggagtacaagttgctcgaac: Reverse Primer: cagatcgtgttccctttagc) using Taq DNA
355 polymerase (Gene DirectX) (PCR mix: 1uL DNA, 1uL forward primer, 1uL reverse primer, 0.5uL
356 10mM dNTPs, 2uL 10x PCR Buffer, 0.1 uL taq, H20 to 20uL). PCR conditions were as follows: 5
357 minutes 95°, (30 seconds 95°, 30 seconds 55°, 30 seconds 72°)x30. Reactions were checked on a
358 gel and cleaned with the GenepHlow™ Gel/PCR Kit (Geneaid). Sanger sequencing reactions
359 were performed by the Mobix Lab at McMaster University. All alignments were created using
360 ClustalOmega (Madeira et al., 2022).

361

362 **Results**

363 *dachsous (ds)* shape change is aligned with major axes of genetic and phenotypic variation in
364 natural populations

365 To assess the relationship between shape change vectors and axes of variation in the
366 DGRP, mean shape vectors were calculated for each DGRP line then used in a PCA to summarize

367 axes of variation among strains. The mean shape vectors for each DGRP line were projected
368 onto vectors for *ds* and *emc* defined from the RNAi knockdown and the correlations between
369 them (Figure 1, Supplemental Figure 5) and with PC1 generated from DGRP line means (*ds*: r = -
370 0.56; *emc*: r = -0.45; Figure 1). A similar response was observed with the data projected onto
371 the *ds* and *emc* shape change vectors (Figure 1, r= 0.69) due to the high correlation between
372 shape change vectors themselves (r = 0.65). Projections of the data onto the vector defining the
373 *neur* shape change is aligned with PC1 (r = -0.69) and PC3 (r = -0.64), indicating that this is also
374 an important axis of shape variation in this population (Figure 1) that differs from projections
375 onto both *ds* (r = 0.034) and *emc* (r = 0.30) shape change vectors. Because of these observed
376 correlations, and previous associations observed (Pitchers et al., 2019) *ds, emc* and *neur* were
377 selected as focal genes for subsequent studies.
378 We also examined the relationship between direction of phenotypic effects with the
379 wild caught cohorts, where phenotypic variance for shape (and thus principal components) is
380 due to the joint contribution of genetic and environmental effects. To illustrate the difference
381 in shape variance in wild populations and the DGRP, we calculated correlations between the
382 first three eigenvectors for shape in the DGRP, the combined wild cohorts as well as the CMO
383 cohort alone, and observed low correlation between the DGRP eigenvectors and those
384 estimated from wild populations (Supplemental Table 2). As observed with the DGRP, there is a
385 substantial correlation between projections of individual shapes onto the *ds* shape change
386 vector and PC1 (defined by phenotypic variation among wild caught files) in most of the
387 sampled cohorts (PHO: r = 0.78; CMO: r = 0.87; FVW13: r = -0.22; FVW14: r = 0.95, Figure 2,
388 Supplemental Figure 9). In the cohorts where the *ds* shape change vector was not correlated
389 with PC1, specifically the FVW13 collection, this vector is correlated with PC2 (PHO: r = 0.12;
390 CMO r = -0.44; FVW13: r = -0.63; FVW14: r = 0.19; Figure 2; Supplemental Figure 5) from the
391 wild caught individuals. We also observe a correlation between *neur* shape change and PC1 in
392 most cohorts (PHO: r = 0.51; CMO: r = -0.051; FVW12: r = -0.95 FVW13; FVW14: r = -0.084;
393 Figure 2; Supplemental Figure 7). As with the *ds* shape change vector, in cases such as the CMO
394 cohort and PC1 there is a correlation between the *neur* shape change vector and PC2 (PHO: r =
395 0.22; CMO: r = -0.57; FVW13: r = -0.0.059; FVW14: r = 0.85; Figure 2; Supplemental Figure 7).

396 Interestingly, in the CMO natural cohort of flies, the correlations between the projection of
397 shape data onto the *ds* and *neur* shape change vectors is low (r = 0.11, Figure 2).
398 Multiple loci linked to hippo signaling - including *ds*- respond to artificial selection for *ds* and
399 *emc* shape changes.
400 To examine if variants in *ds* are contributing to shape variation, replicating the findings
401 of the earlier GWAS (Pitchers et al. 2019), we performed artificial selection experiment for wing
402 shape along the *ds* shape change vector and then examined the genomic response to selection.
403 By the final generation of selection, we observed a substantial shape change in both the "up"
404 (females: Procrustes Distance (PD) = 0.039, males: 0.044) and "down" directions (females: PD
405 = 0.022, males: PD = 0.022) compared to the base population at the start of the selection
406 experiment. In comparison, the shape change among unselected control lineages was much
407 smaller (females: PD = 0.005, males: PD = 0.005) (Figure 3, Supplemental Figure 10). The
408 direction of phenotypic shape change after seven generations of selection was in a similar
409 direction to the *ds* shape change vector (defined by the RNAi knockdown) for both the up
410 (females: r = 0.90, males: r = 0.90) and down (females: r = -0.82, males: r = -0.77) selection
411 lineages. Realized heritabilities, averaged over sex and replicate (Supplemental Figure 11, up =
412 0.38, 95% CI: 0.25 – 0.50; down = 0.28, 95% CI: 0.24 – 0.50) were moderate. Hippo signaling,
413 including the effects of *ds*, is often associated with changes in size (Pan, 2007). However, we do
414 not observe a significant change of wing size in our selection lineages in either sex
415 (Supplemental Figure 12). It is possible that with more generations of selection we would have
416 observed a clear change in size, as there is a trend indicating such divergence (Supplemental
417 Figure 12).
418 Genome-wide patterns of F_{ST} were examined between up and down *ds* selection
419 lineages. We observed strong genetic differentiation linked with the *ds* locus (Figure 3,
420 Supplemental Figure 13). We also observed striking genetic differentiation across several
421 regions in the genome. In addition, one of the SNPs (2L:702560) in the intron of *ds* identified in
422 Pitchers et al. (2019) through GWAS, showed the expected pattern of response to selection,
423 with opposing sign in up and down selection lineages, with the SNP going to high frequency in
424 all three up selection lineages (Table 1). It should be noted that this SNP represents a complex

425 polymorphism including an insertion of 18bp that may result in inaccurate genotyping at this
426 locus which we further evaluated with Sanger sequencing of this region (Supplemental Figure
427 14). Gene ontology analysis for genes in regions of the genome with an F_{ST} greater than 0. 345
428 (three standard deviations from mean F_{ST}), show enrichment for hippo signaling loci
429 (Supplemental Table 3). The top 20 enriched terms are all related to cell signaling and
430 development. Of particular note is the inclusion of the terms for 'negative regulation of hippo
431 signaling' (GO:0035331), and 'hippo signaling' (GO:0035329) in this list (Supplemental Table 3,
432 Supplemental Figure 13). Using a permutation test we confirmed this result by selecting
433 random sets of genomic intervals equal in size to the number of observed outlier windows, and
434 measuring the ratio of genes annotated to the expected number of genes in these regions. The
435 observed value for the terms for hippo signaling (ratio = 4.76) and negative regulation of hippo
436 signaling (ratio = 9.23) were compared to the distribution of values for 1000 random sets of loci
437 (Supplemental Figure 15). For the hippo signaling term the observed to expected ratio for
438 random loci was larger than the observed value 5 times out of 1000, while the value from
439 random loci was never larger than the observed.
440 For the artificial selection experiment based on the *emc* shape change vector we
441 observed phenotypic differentiation under artificial selection in both up (females: PD = 0.043,
442 males: PD = 0.040) and down directions (females: PD = 0.021, males: PD = 0.020) with little
443 change in control lineages (females: PD = 0.009, males: PD = 0.008) (Figure 4). The direction of
444 phenotypic change is correlated with the *emc* (RNAi knockdown) shape change vector in both
445 up (females: r = 0.75, males: r = 0.69) and down (females: r = -0.69, males: r = -0.75) directions.
446 Realized heritabilities, averaged over sex and replicate were calculated for both up and down
447 lineages (Supplemental Figure 16, up = 0.38, 95% CI: 0.29 – 0.47; down = 0.28, 95% CI: 0.21 –
448 0.35). Genetic differentiation linked to the *emc locus* was modest following selection, but we
449 again observed striking genetic differentiation linked to *ds* (Figure 4, Supplemental Figure 13).
450 Notably, as seen in Supplemental Figure 1, the site frequency spectrum (SFS) suggests modest
451 allelic variation at the *emc* locus in the synthetic outbred population. Using a three standard
452 deviation cut-off for F_{ST}, we did observe enrichment for various developmental GO terms, but
453 not of hippo signaling terms (Supplemental Table 4, Supplemental Figure 13).

454

455 Bulk segregant analysis in wild caught cohorts does not recapitulate effects of the GWAS or
456 artificial selection

457 Having demonstrated that variants in *ds* (or linked to *ds*) respond to artificial selection
458 for wing shape along the *ds* shape change vector, we next wanted to determine whether we
459 could recapitulate these findings with wild caught individuals. In addition to determining
460 whether we can replicate effects in wild cohorts, it provides the opportunity to identify causal
461 SNPs because of low LD generally observed in wild caught *Drosophila*. Wild caught populations
462 introduce considerably more environmental variation for shape along with a different site
463 frequency spectrum for variants contributing to shape variation (and *ds* like shape changes
464 specifically). In particular, it is known that several of the variants that the original GWAS
465 detected in *ds* have low minor allele frequency (MAF) (Pitchers et al., 2019) (Table 2). The SNP
466 at 2L:702560 does appear to be at intermediate frequency but it occurs both directly before
467 and after an indel, making the alignment and variant calling in this region challenging
468 (Supplementary Figure 14). We have included the frequencies (Table 2), but these results
469 should be interpreted with caution due to the technical complexities of mapping and variant
470 calling close to indels.

471 As we sampled multiple cohorts of wild-caught flies in different locations and years in
472 Michigan (USA), we wanted to confirm that any phenotypic differentiation among these
473 samples was modest and would not impact genomic analysis for the BSA. We observe modest,
474 statistically significant wing shape differences among cohorts from a Procrustes ANOVA,
475 utilizing permutations of the residuals for the relevant "null" model (Supplemental Table 5; R^2=
476 0.16, F = 351, Z_{RRPP} = 18.3, p = 0.001) (Collyer and Adams, 2018). This appears to be due to
477 differences in wing shape between the PHO population and other populations based on
478 pairwise Procrustes Distances (Supplemental Table 5, Supplemental Figure 3). In a joint PCA
479 including all populations, there is very modest separation between populations using allometry
480 adjusted shape (Supplemental Figure 3). Most relevant to the bulk segregant analysis approach
481 we used, when we project all wild caught individuals onto the *ds* and *neur* vectors, we see no
482 clear separation among sampling locales (Supplemental Figure 7). There is some variation in

483 wing size between populations (Supplemental Figure 8), but this is unlikely to influence
484 downstream analysis as we use size adjusted estimates. There is little evidence of genetic
485 differentiation between populations with the two collections from Fenn Valley Winery
486 separating more on a Principal Co-ordinate Analysis (PCoA)(Supplemental Figure 17) than other
487 sampling locales. There is also no relationship between genetic and phenotypic distances
488 between samples (Supplemental Figure 18). These results suggests that the multiple sampling
489 locales should not influence downstream genomic analysis as individuals used for generating
490 pools were compared within each population, and we observe little evidence for substantial
491 differences among populations.

492 Because there is a single bout of phenotypic selection distinguishing pools in the bulk
493 segregant analysis (BSA), shape and allele frequency changes between pools are expected to be
494 modest. We observe shape difference between the two pools within each population (PD =
495 CMO: 0.033; PHO: 0.036; FVW13: 0.040; FVW14: 0.041; Supplemental Figure 19). Correlations
496 of the shape difference vectors of the pools (i.e. difference between the two pools created
497 from the extremes along the *ds* shape change axis) and the direction of the *ds* shape change
498 vector used for selection is high (CMO: 0.94, PHO: 0.79, FVW13: 0.92, FVW14: 0.90).

499 The genome scans from the bulk segregant analysis show little evidence of genetic
500 differentiation linked to the *ds gene* (Figure 5). Across the genome, 15 sites were detected as
501 significantly differentiated between "up" and "down" selected pools based on a CMH test with
502 a FDR cut-off of 5% (Figure 5, Table 3). The genes nearest to these sites are not associated with
503 hippo signaling pathways or implicated in the development of the *Drosophila* wing (Table 3).
504 Because PHO had somewhat distinct shape variation from the other populations and had a
505 lower correlation of the difference vector between selected pools and *ds* shape change vector,
506 we repeated the CMH test with this population left out. We observe significant differentiation
507 at 174 sites between "up" and "down" pools (Supplemental Table 6, Supplemental Figure 20).
508 We identified the nearest genes to these sites and GO analysis indicated enrichment for wing
509 development terms, in particular related to Wnt signaling, but not hippo signaling terms
510 (Supplemental Table 6). Importantly, we do not observe differentiation linked to *ds* or any
511 other hippo loci. To ensure that the results we obtained were not due to uneven coverage

512 between samples, we down-sampled genomic coverage to 75x for each sample, dropping sites
513 that did not meet this threshold. Significant differences were detected at 19 sites
514 (Supplemental Figure 21, Supplemental Table 7), but none of these overlapped with those
515 identified using all the genomic data. Two of the significant sites are located in the *dumpy* gene,
516 a gene known to have a role in wing morphogenesis during pupation (Etournay et al., 2015). F_{ST}
517 between selected and random pools within each cohort are generally low (Supplemental Figure
518 22).

519 In addition to the BSA selection based upon the *ds* shape change, we also selected pools
520 of individuals based on the *neur* shape change vector. We did not use *emc* shape change in this
521 experiment due to the high similarity between the *ds* and *emc* shape change vectors (r = 0.65).
522 We selected the *neur* shape change vector as it is not aligned with *ds* but does align with
523 directions of variation in wild populations (Figure 1, 2, Supplemental Figure 9). Additionally,
524 there is little relationship between the *ds* and *neur* shape change axis (r = 0.12, Supplemental
525 Figures 7 and 8). We observe shape changes between pools of individuals (PD = CMO: 0.027;
526 PHO: 0.028; FVW14: 0.041; FVW13: 0.038, Supplemental Figure 23). There is little evidence of
527 genetic differentiation between *neur* selected pools (Supplemental Figure 24). Only 4 sites were
528 identified as being significantly differentiated between pools and none of these sites are
529 associated with wing development (Supplemental Table 9). When population differentiation
530 between pools within populations is measured using F_{ST}, genetic differentiation remains low
531 across the genome (Supplemental Figure 25)
532

533 **Discussion**

534 The primary goal of this study was to determine whether we could recapitulate genetic
535 effects initially observed through a traditional GWAS using an "inverted" approach: artificially
536 selecting on phenotypes and observing changes in allele frequencies. We observed that shape
537 changes associated with the *ds*, *emc* and *neur* genes were associated with major axes of genetic
538 variation among a panel of wild type strains (DGRP) reared in the lab, and axes of phenotypic
539 variation among wild caught individuals (Figure 1, 2). After observing a strong response to
540 artificial selection along two shape change vectors (*ds* and *emc*), we examined patterns of

541 genomic differentiation and observed substantial changes in allele frequency for markers linked
542 with *ds* itself, and also markers linked to numerous other genes associated with hippo signaling
543 (Figure 3, 4).
544 In contrast, our bulk segregant experiments using pools of wild caught individuals chosen to
545 be phenotypically divergent on the same shape vectors, did not detect differences in the loci
546 identified in the artificial selection experiments (Figure 5, Supplemental Figure 24). As we
547 discuss in detail below, these seemingly contradictory results are in fact not that surprising.
548 Following artificial selection based on *ds* shape change we observe allele frequency changes
549 not only at *ds* but also linked to a number of other hippo signalling loci (Figure 3, Supplemental
550 Table 3). The previous GWAS study identified a number of loci associated with wing shape
551 variation in the DGRP, however, this approach cannot predict which alleles are causative
552 (Pitchers et al., 2019). In our synthetic outbred population, we maximized variation among
553 haplotype blocks containing many of the candidate SNPs in *ds*, increasing our ability to detect
554 frequency changes at and near the implicated variants. Although LD blocks in the outcrossing
555 population from this study remain large, *ds* variants exist on multiple distinct haplotypes,
556 allowing for an examination of allele frequency changes for variants. Of particular interest is
557 SNP 2L:702560, which was previously identified though GWAS (Pitchers et al., 2019) as
558 influencing wing shape variation. It was driven to near fixation in each of the artificial selection
559 lineages (Table 1). Although this polymorphism is annotated as a SNP, this region may contain a
560 complex polymorphism (Supplemental Figure 14), making it difficult to accurately assess
561 genotypic calls. Because of this, the predicted allele frequency in the founding population and
562 allele frequencies in this region may be inaccurate. Previous studies demonstrate the
563 importance of alleles at intermediate frequency in founding populations to those contributing
564 to responses to selection over short timescales (Kelly and Hughes, 2019). If this polymorphism
565 is at a more intermediate frequency in the founding population, it would be more likely to be
566 captured by selection during these experiments. Additionally, we recognize that the haplotype
567 blocks are large, and contain many potential functional variants. However, based on the results
568 of both the current and previous studies, these *ds* variants associated with 2L:702560 are good
569 candidates for functional validation in future work.

570 When selecting on the *emc* shape change vector, which is similar but not identical to *ds*
571 shape change, we observe only a modest allele frequency change at *emc*, and a more robust
572 response at *ds* (Figure 4). In hindsight, this is not particularly surprising and there are multiple
573 contributing factors. Given the increased genetic diversity at *ds* compared to *emc* in the
574 founding population, alleles in *ds* may have provided a more robust genetic target influencing
575 phenotypic variation for wing shape along this axis of variation. Additionally, if our estimated
576 direction of effects and selection for *emc* (based on RNAi knockdown) was not well aligned with
577 the actual direction of *emc* SNP effects, this could result in weaker selection on variants at the
578 *emc* locus. It is worthwhile pointing out the small magnitude of the *emc* shape change vector
579 (0.44) relative to *ds* (5.5). However, previous work has indicated that there is a relationship
580 between this estimated *emc* shape change vector (from RNAi) and the effect of SNPs in *emc* on
581 shape change (Pitchers et al., 2019).
582 In addition to a response on allele frequency associated with *ds*, our results suggest a
583 response on segregating variation at other hippo signaling loci. Earlier work has suggested that
584 the direction of effects within signaling pathways are inconsistent for alleles of small effect
585 (Dworkin and Gibson, 2006). However, allelic effect sizes in the 2006 study were heterogeneous
586 and may result in direction and magnitude being confounded. In contrast, in both the current
587 and the Pitchers et al. (2019) studies, we estimated the direction of genetic effects by titrating
588 gene knockdown. The strength of this approach is highlighted in the result that segregating
589 variation at multiple hippo loci was selected on (Figure 3). Our finding is consistent with models
590 for the architecture of complex traits that predict that many alleles of small effect will
591 contribute to trait variation with many genes within developmental pathways (Boyle et al.,
592 2017; Wray et al., 2018). This pathway response has also been demonstrated in human
593 adaptation to pathogen resistance (Daub et al., 2013) and adaptation to high altitude (Gouy et
594 al., 2017). These results are consistent with the expectation that polymorphisms in the same
595 developmental pathway would show correlated phenotypic effects and therefore correlated
596 genomic responses to selection. However, this may not be reflective of all wild caught
597 populations. In this study, we generated a population that had high diversity at *ds*, while these
598 variants are at much lower frequency in natural populations (Table 2). The amount of

599 selectable variation a variant causes is dependent on both effect size, *a*, and variant
600 frequencies, *p* as $V_A = 2p(1-p)a^2$. When allele frequencies are near 0 or 1, even variants with
601 large effects will have only a small contribution to short term selection response. Therefore,
602 the outcrossed population we created here is an ideal situation to validate the existence of the
603 measured effects. It is unlikely to be typical of natural populations where functional variants
604 may be rare.

605 Given the clear and robust response observed in the artificial selection experiment, it may
606 seem surprising that we do not observe allele frequency changes in the bulk segregant analysis
607 using the wild cohorts. Indeed, previous work has demonstrated that variants in *Egfr*, could be
608 replicated in wild caught samples (Dworkin et al., 2005; Palsson et al., 2005) and were also
609 found in genome wide associations (Pitchers et al., 2019). However, there are many
610 explanations for why we may not have been able to detect these allele frequency changes in
611 our experiment. First, the addition of environmental variation to the system introduces
612 additional complications. In the aforementioned example of variants in *Egfr*, the genetic effect
613 of the SNP in wild caught cohorts was only 10% of the magnitude estimated in lab reared flies.
614 As discussed previously, the *ds* variants implicated in the previous GWAS study are at low
615 frequency in the natural cohorts (Table 2). Given that natural populations of *Drosophila* are
616 generally large and wing shape is likely under weak selection (Gilchrist and Partridge, 2001),
617 mutation-drift-selection balance may maintain most variation, resulting in low minor allele
618 frequencies at these sites. Because allelic contribution to wing shape are expected to be both
619 rare in wild populations and of small phenotypic effect, we do not expect large allele frequency
620 changes given only one "generation" of selection. Using the approach of ACER (Spitzer et al.,
621 2020) to account for sampling effects, we observe few differentiated sites, and none in the *ds*
622 gene, indicating that bulk segregant analysis (BSA) may not be a well-suited approach to
623 identify modest allele frequency changes and as such may not be particularly effective for
624 highly polygenic traits. Although our approach was tailored to look for variants that had
625 consistent direction of frequency changes across the four collection cohorts, it is possible that
626 different loci were contributing variation within each cohort. We attempted to address this
627 question by examining allele frequency changes between selected pools within each cohort

628 (Supplemental Figures 22 and 24) but could not identify specific loci contributing to differences
629 within any one population. Previous successful BSA studies identified a smaller number of
630 contributing loci with few polymorphisms contributing to the trait of interest. For example, in
631 *Drosophila,* a number of melanin synthesis genes contributing to variance in pigmentation
632 between populations were identified using a BSA (Bastide et al., 2013). Pigmentation may
633 represent a relatively 'simpler' genetic architecture (fewer variants of individually larger genetic
634 effect, smaller impact of environmental variation, smaller mutational target size) and if so, this
635 may have enabled the success of the BSA approach with such systems. In the case of wing
636 shape, we know that many alleles of small effect contribute to variation in the trait (Pitchers et
637 al., 2019).
638 Our approach for the BSA was to perform the same phenotypic selection within each of
639 four distinct "populations". It is important to recognize that there was heterogeneity among
640 our populations, not only in allele frequencies, but in environmental variance and potentially
641 GxE, even though all were caught in locales in lower Michigan. We detected small degrees of
642 phenotypic and genetic differences between the four cohorts, however these effects are
643 neither correlated with one another, nor related to the *ds* and *neur* shape scores we used for
644 selecting individuals (Supplemental Figures 3, 7, 18).The population from the Phillips Orchard
645 (PHO) was phenotypically distinct from the other populations. When we performed the BSA
646 without this population, we observed a larger set of variants associated with shape
647 (Supplemental Figure 20), albeit still not showing any effects at the *ds* or *neur* genes
648 themselves. One possibility is that the increased number of sites when the PHO sample is
649 removed from analysis represents an unknown statistical artefact we have not identified.
650 However, a more likely explanation is that there are some large unknown environmental
651 influences (**E**), or that the genetic effects show a degree of GxE (with a specific environment in
652 PHO) that contributed to shape variation along the *ds* direction in this population. Such
653 obfuscating effects have been observed before with the previously discussed *Egfr* example,
654 where the same SNP effect that was initially identified and validated in multiple contexts
655 (Dworkin et al., 2005; Palsson et al., 2005; Palsson and Gibson, 2004) could not be detected in
656 one natural population, despite being at intermediate frequencies in each sample (Palsson et

al., 2005). Importantly, we did detect differentiation at sites associated with developmental processes in the wild cohorts, suggesting that the failure to detect variation linked to *ds* or other hippo signaling loci (Table 3, Supplemental Table 6,7) is not due simply a lack of power.

The response to selection at *ds* and other hippo signaling loci in the artificial selection experiment based on *ds* shape change indicates that this is an important axis of variation for wing shape. Coupled with the alignment of phenotypic effects of perturbations in genes in this pathway with directions of **G** and **P**, this finding may seem to suggest a developmental bias in available variation. However, we caution against such interpretations based solely on the findings in this study. The structure of the **G** matrix strongly influenced our findings as we artificially created a population to maximize genetic diversity at *ds*. When another effect is aligned with *ds* shape change, as in the case of *emc* shape change, we observed the same response at the hippo signaling loci and not at *emc.* Only the genetic diversity in the starting population was available to be selected on so this influenced selection towards the "spiked in" *ds* variants, even if the inferred phenotypic effects of *emc* variants are very similar. Alternatively, the inferred *emc* direction of effects (via RNAi knockdown) may be sufficiently "distant" from true effects of *emc* variants. If this was the case, we were ineffectively selecting for *emc* shape changes. In other cases where single genes are implicated in divergence between multiple populations, such as *mc1r* in mice (Steiner et al., 2007) or *pitx1* in stickleback (Chan et al., 2010), other factors such as low pleiotropy, developmental and mutational constraints and history of selection in the population are used to explain why these genes are so often implicated in evolutionary change (Gompel and Prud'homme, 2009; Martin and Orgogozo, 2013; Stern and Orgogozo, 2008). In our case, it is not *ds* itself that is special but rather the orientation of the **G** matrix to align $\mathbf{g}_{\mathbf{max}}$ with the direction of effect for *ds* that shapes our results. Selection acts on variants aligned with the vector of selection (Reddiex and Chenoweth, 2021). By varying the orientation of $\mathbf{g}_{\mathbf{max}}$ in the parental population, we would be able to address questions about the repeatability of hippo overrepresentation and if this can be explained by more than just the orientation of **G.**

Despite the need for skepticism about the potential for developmental bias influencing directions of variation, the correlated response of sites linked to multiple other hippo signaling

genes is intriguing. Coupling of more traditional mapping approaches like GWAS with short term artificial selection provides an additional route to validation and replication of genetic effects. It also suggests that using multivariate data to address the distribution of genetic effects will pay long-term dividends to our understanding of both inheritance and the evolution of multivariate traits.

Data Availability:

All code is available on GitHub at: https://github.com/KatiePelletier/WingShapeBSA

All images and phenotypic data are available on Dryad at:

All sequence data has been submitted to NCBI

Acknowledgements:

We would like to thank the Hosken lab for *D. simulans* wing shape data. Yun Bo Xi aided with the sequencing of *ds* haplotype.

References:

Adams, D.C., Otárola-Castillo, E., 2013. geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods Ecol. Evol. 4, 393–399. https://doi.org/10.1111/2041-210X.12035

Alexa, A., Rahnenfuhrer, J., Lengauer, T., 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22, 1600–1607. https://doi.org/10.1093/bioinformatics/btl140

Bastide, H., Betancourt, A., Nolte, V., Tobler, R., Stöbe, P., Futschik, A., Schlötterer, C., 2013. A Genome-Wide, Fine-Scale Map of Natural Pigmentation Variation in Drosophila melanogaster. PLoS Genet. 9, e1003534. https://doi.org/10.1371/journal.pgen.1003534

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Blows, M.W., McGuigan, K., 2015. The distribution of genetic variance across phenotypic space and the response to selection. Mol. Ecol. 24, 2056–2072. https://doi.org/10.1111/mec.13023

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

Boyle, E.A., Li, Y.I., Pritchard, J.K., 2017. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell 169, 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

722 Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M.,
723 Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., Petrov, D., Jónsson, B., Schluter,
724 D., Bell, M.A., Kingsley, D.M., 2010. Adaptive Evolution of Pelvic Reduction in
725 Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. Science 327, 302–305.
726 https://doi.org/10.1126/science.1182213
727 Collyer, M.L., Adams, D.C., 2018. RRPP: An R package for fitting linear models to high-
728 dimensional data using residual randomization. Methods Ecol. Evol. 9, 1772–1779.
729 https://doi.org/10.1111/2041-210X.13029
730 CONVERGE consortium, 2015. Sparse whole-genome sequencing identifies two loci for major
731 depressive disorder. Nature 523, 588–591. https://doi.org/10.1038/nature14659
732 Daub, J.T., Hofer, T., Cutivet, E., Dupanloup, I., Quintana-Murci, L., Robinson-Rechavi, M.,
733 Excoffier, L., 2013. Evidence for Polygenic Adaptation to Pathogens in the Human
734 Genome. Mol. Biol. Evol. 30, 1544–1558. https://doi.org/10.1093/molbev/mst080
735 Dworkin, I., Gibson, G., 2006. Epidermal Growth Factor Receptor and Transforming Growth
736 Factor-β Signaling Contributes to Variation for Wing Shape in *Drosophila melanogaster*.
737 Genetics 173, 1417–1431. https://doi.org/10.1534/genetics.105.053868
738 Dworkin, I., Palsson, A., Gibson, G., 2005. Replication of an *Egfr* -Wing Shape Association in a
739 Wild-Caught Cohort of *Drosophila melanogaster*. Genetics 169, 2115–2125.
740 https://doi.org/10.1534/genetics.104.035766
741 Etournay, R., Popović, M., Merkel, M., Nandi, A., Blasse, C., Aigouy, B., Brandl, H., Myers, G.,
742 Salbreux, G., Jülicher, F., Eaton, S., 2015. Interplay of cell dynamics and epithelial tension
743 during morphogenesis of the Drosophila pupal wing. eLife 4, e07090.
744 https://doi.org/10.7554/eLife.07090
745 Fatumo, S., Carstensen, T., Nashiru, O., Gurdasani, D., Sandhu, M., Kaleebu, P., 2019.
746 Complimentary Methods for Multivariate Genome-Wide Association Study Identify New
747 Susceptibility Genes for Blood Cell Traits. Front. Genet. 10, 334.
748 https://doi.org/10.3389/fgene.2019.00334
749 Fritsche, L.G., Igl, W., Bailey, J.N.C., Grassmann, F., Sengupta, S., Bragg-Gresham, J.L., Burdon,
750 K.P., Hebbring, S.J., Wen, C., Gorski, M., Kim, I.K., Cho, D., Zack, D., Souied, E., Scholl,
751 H.P.N., Bala, E., Lee, K.E., Hunter, D.J., Sardell, R.J., Mitchell, P., Merriam, J.E., Cipriani,
752 V., Hoffman, J.D., Schick, T., Lechanteur, Y.T.E., Guymer, R.H., Johnson, M.P., Jiang, Y.,
753 Stanton, C.M., Buitendijk, G.H.S., Zhan, X., Kwong, A.M., Boleda, A., Brooks, M., Gieser,
754 L., Ratnapriya, R., Branham, K.E., Foerster, J.R., Heckenlively, J.R., Othman, M.I., Vote,
755 B.J., Liang, H.H., Souzeau, E., McAllister, I.L., Isaacs, T., Hall, J., Lake, S., Mackey, D.A.,
756 Constable, I.J., Craig, J.E., Kitchner, T.E., Yang, Z., Su, Z., Luo, H., Chen, D., Ouyang, H.,
757 Flagg, K., Lin, D., Mao, G., Ferreyra, H., Stark, K., von Strachwitz, C.N., Wolf, A., Brandl,
758 C., Rudolph, G., Olden, M., Morrison, M.A., Morgan, D.J., Schu, M., Ahn, J., Silvestri, G.,
759 Tsironi, E.E., Park, K.H., Farrer, L.A., Orlin, A., Brucker, A., Li, M., Curcio, C.A., Mohand-
760 Saïd, S., Sahel, J.-A., Audo, I., Benchaboune, M., Cree, A.J., Rennie, C.A., Goverdhan, S.V.,
761 Grunin, M., Hagbi-Levi, S., Campochiaro, P., Katsanis, N., Holz, F.G., Blond, F., Blanché,
762 H., Deleuze, J.-F., Igo, R.P., Truitt, B., Peachey, N.S., Meuer, S.M., Myers, C.E., Moore,
763 E.L., Klein, R., Hauser, M.A., Postel, E.A., Courtenay, M.D., Schwartz, S.G., Kovach, J.L.,
764 Scott, W.K., Liew, G., Tan, A.G., Gopinath, B., Merriam, J.C., Smith, R.T., Khan, J.C.,
765 Shahid, H., Moore, A.T., McGrath, J.A., Laux, R., Brantley, M.A., Agarwal, A., Ersoy, L.,

766 Caramoy, A., Langmann, T., Saksens, N.T.M., de Jong, E.K., Hoyng, C.B., Cain, M.S.,
767 Richardson, A.J., Martin, T.M., Blangero, J., Weeks, D.E., Dhillon, B., van Duijn, C.M.,
768 Doheny, K.F., Romm, J., Klaver, C.C.W., Hayward, C., Gorin, M.B., Klein, M.L., Baird, P.N.,
769 den Hollander, A.I., Fauser, S., Yates, J.R.W., Allikmets, R., Wang, J.J., Schaumberg, D.A.,
770 Klein, B.E.K., Hagstrom, S.A., Chowers, I., Lotery, A.J., Léveillard, T., Zhang, K., Brilliant,
771 M.H., Hewitt, A.W., Swaroop, A., Chew, E.Y., Pericak-Vance, M.A., DeAngelis, M.,
772 Stambolian, D., Haines, J.L., Iyengar, S.K., Weber, B.H.F., Abecasis, G.R., Heid, I.M., 2016.
773 A large genome-wide association study of age-related macular degeneration highlights
774 contributions of rare and common variants. Nat. Genet. 48, 134–143.
775 https://doi.org/10.1038/ng.3448
776 Gilchrist, A.S., Partridge, L., 2001. The contrasting genetic architecture of wing size and shape in
777 Drosophila melanogaster. Heredity 86, 144–152. https://doi.org/10.1046/j.1365-
778 2540.2001.00779.x
779 Gompel, N., Prud'homme, B., 2009. The causes of repeated genetic evolution. Dev. Biol. 332,
780 36–47. https://doi.org/10.1016/j.ydbio.2009.04.040
781 Gouy, A., Daub, J.T., Excoffier, L., 2017. Detecting gene subnetworks under selection in
782 biological pathways. Nucleic Acids Res. 45, e149–e149.
783 https://doi.org/10.1093/nar/gkx626
784 Green, R.M., Fish, J.L., Young, N.M., Smith, F.J., Roberts, B., Dolan, K., Choi, I., Leach, C.L.,
785 Gordon, P., Cheverud, J.M., Roseman, C.C., Williams, T.J., Marcucio, R.S., Hallgrímsson,
786 B., 2017. Developmental nonlinearity drives phenotypic robustness. Nat. Commun. 8,
787 1970. https://doi.org/10.1038/s41467-017-02037-7
788 Greene, C.S., Penrod, N.M., Williams, S.M., Moore, J.H., 2009. Failure to Replicate a Genetic
789 Association May Provide Important Clues About Genetic Architecture. PLoS ONE 4,
790 e5639. https://doi.org/10.1371/journal.pone.0005639
791 Houle, D., Mezey, J., Galpern, P., Carter, A., 2003. Automated measurement of Drosophila
792 wings. BMC Evol. Biol. 3, 25. https://doi.org/10.1186/1471-2148-3-25
793 Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ramia, M., Tarone, A.M., Turlapati, L., Zichner,
794 T., Zhu, D., Lyman, R.F., Magwire, M.M., Blankenburg, K., Carbone, M.A., Chang, K., Ellis,
795 L.L., Fernandez, S., Han, Y., Highnam, G., Hjelmen, C.E., Jack, J.R., Javaid, M., Jayaseelan,
796 J., Kalra, D., Lee, S., Lewis, L., Munidasa, M., Ongeri, F., Patel, S., Perales, L., Perez, A., Pu,
797 L., Rollmann, S.M., Ruth, R., Saada, N., Warner, C., Williams, A., Wu, Y.-Q., Yamamoto,
798 A., Zhang, Y., Zhu, Y., Anholt, R.R.H., Korbel, J.O., Mittelman, D., Muzny, D.M., Gibbs,
799 R.A., Barbadilla, A., Johnston, J.S., Stone, E.A., Richards, S., Deplancke, B., Mackay, T.F.C.,
800 2014. Natural variation in genome architecture among 205 Drosophila melanogaster
801 Genetic Reference Panel lines. Genome Res. 24, 1193–1208.
802 https://doi.org/10.1101/gr.171546.113
803 Ioannidis, J.P.A., Tarone, R., McLaughlin, J.K., 2011. The False-positive to False-negative Ratio in
804 Epidemiologic Studies: Epidemiology 22, 450–456.
805 https://doi.org/10.1097/EDE.0b013e31821b506e
806 Kelly, J.K., Hughes, K.A., 2019. Pervasive Linked Selection and Intermediate-Frequency Alleles
807 Are Implicated in an Evolve-and-Resequencing Experiment of *Drosophila simulans*.
808 Genetics 211, 943–961. https://doi.org/10.1534/genetics.118.301824

King, Elizabeth G., Macdonald, S.J., Long, A.D., 2012. Properties and Power of the Drosophila Synthetic Population Resource for the Routine Dissection of Complex Traits. Genetics 191, 935–949. https://doi.org/10.1534/genetics.112.138537

King, E. G., Merkes, C.M., McNeil, C.L., Hoofer, S.R., Sen, S., Broman, K.W., Long, A.D., Macdonald, S.J., 2012. Genetic dissection of a model complex trait using the Drosophila Synthetic Population Resource. Genome Res. 22, 1558–1566. https://doi.org/10.1101/gr.134031.111

Klingenberg, C.P., 2022. Methods for studying allometry in geometric morphometrics: a comparison of performance. Evol. Ecol. 36, 439–470. https://doi.org/10.1007/s10682-022-10170-z

Kofler, Orozco-terWengel, P., De Maio, N., Pandey, R.V., Nolte, V., Futschik, A., Kosiol, C., Schlötterer, C., 2011a. PoPoolation: A Toolbox for Population Genetic Analysis of Next Generation Sequencing Data from Pooled Individuals. PLoS ONE 6, e15925. https://doi.org/10.1371/journal.pone.0015925

Kofler, Pandey, R.V., Schlotterer, C., 2011b. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27, 3435–3436. https://doi.org/10.1093/bioinformatics/btr589

Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26, 589–595. https://doi.org/10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., Lopez, R., 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. 50, W276–W279. https://doi.org/10.1093/nar/gkac240

Marigorta, U.M., Rodríguez, J.A., Gibson, G., Navarro, A., 2018. Replicability and Prediction: Lessons and Challenges from GWAS. Trends Genet. 34, 504–517. https://doi.org/10.1016/j.tig.2018.03.005

Marquez, E.J., Houle, D., 2015. Dimensionality and the statistical power of multivariate genome-wide association studies (preprint). Genomics. https://doi.org/10.1101/016592

Marriage, T.N., King, E.G., Long, A.D., Macdonald, S.J., 2014. Fine-Mapping Nicotine Resistance Loci in *Drosophila* Using a Multiparent Advanced Generation Inter-Cross Population. Genetics 198, 45–57. https://doi.org/10.1534/genetics.114.162107

Martin, A., Orgogozo, V., 2013. The Loci of Repeated Evolution: A Catalog of Genetic Hotspots of Phenotypic Variation. Evolution 67, 1235–1250. https://doi.org/10.1111/evo.12081

McGuigan, K., 2006. Studying phenotypic evolution using multivariate quantitative genetics: EVOLUTIONARY QUANTITATIVE GENETICS. Mol. Ecol. 15, 883–896. https://doi.org/10.1111/j.1365-294X.2006.02809.x

Melo, D., Marroig, G., Wolf, J.B., 2019. Genomic Perspective on Multivariate Variation, Pleiotropy, and Evolution. J. Hered. 110, 479–493. https://doi.org/10.1093/jhered/esz011

Palsson, A., Dodgson, J., Dworkin, I., Gibson, G., 2005. Tests for the replication of an association between Egfr and natural variation in Drosophila melanogaster wing morphology. BMC Genet. 6, 44. https://doi.org/10.1186/1471-2156-6-44

Palsson, A., Gibson, G., 2004. Association Between Nucleotide Variation in Egfr and Wing Shape in Drosophila melanogaster. Genetics 167, 1187–1198. https://doi.org/10.1534/genetics.103.021766

Pan, D., 2007. Hippo signaling in organ size control. Genes Dev. 21, 886–897. https://doi.org/10.1101/gad.1536007

Pan, Y., Alégot, H., Rauskolb, C., Irvine, K.D., 2018. The dynamics of Hippo signaling during *Drosophila* wing development. Development 145, dev165712. https://doi.org/10.1242/dev.165712

Pitchers, W., Nye, J., Márquez, E.J., Kowalski, A., Dworkin, I., Houle, D., 2019. A Multivariate Genome-Wide Association Study of Wing Shape in *Drosophila melanogaster*. Genetics 211, 1429–1447. https://doi.org/10.1534/genetics.118.301342

Porter, H.F., O'Reilly, P.F., 2017. Multivariate simulation framework reveals performance of multi-trait GWAS methods. Sci. Rep. 7, 38837. https://doi.org/10.1038/srep38837

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

Ray, R.P., Nakata, T., Henningsson, P., Bomphrey, R.J., 2016. Enhanced flight performance by genetic manipulation of wing shape in Drosophila. Nat. Commun. 7, 10851. https://doi.org/10.1038/ncomms10851

Reddiex, A.J., Chenoweth, S.F., 2021. Integrating genomics and multivariate evolutionary quantitative genetics: a case study of constraints on sexual selection in *Drosophila serrata*. Proc. R. Soc. B Biol. Sci. 288, 20211785. https://doi.org/10.1098/rspb.2021.1785

Rohlf, F.J., Slice, D., 1990. Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. Syst. Zool. 39, 40. https://doi.org/10.2307/2992207

Schluter, D., 1996. Adaptive Radiation Along Genetic Lines of Least Resistance. Evolution 50, 1766. https://doi.org/10.2307/2410734

Shriner, D., 2012. Moving toward System Genetics through Multiple Trait Analysis in Genome-Wide Association Studies. Front. Genet. 3. https://doi.org/10.3389/fgene.2012.00001

Spitzer, K., Pelizzola, M., Futschik, A., 2020. Modifying the Chi-square and the CMH test for population genetic inference: Adapting to overdispersion. Ann Appl Stat 14, 202–220. https://doi.org/10.1214/19-AOAS1301

Steiner, C.C., Weber, J.N., Hoekstra, H.E., 2007. Adaptive Variation in Beach Mice Produced by Two Interacting Pigmentation Genes. PLoS Biol. 5, e219. https://doi.org/10.1371/journal.pbio.0050219

Stephens, M., 2013. A Unified Framework for Association Analysis with Multiple Related Phenotypes. PLoS ONE 8, e65245. https://doi.org/10.1371/journal.pone.0065245

Stern, D.L., Orgogozo, V., 2008. THE LOCI OF EVOLUTION: HOW PREDICTABLE IS GENETIC EVOLUTION? Evolution 62, 2155–2177. https://doi.org/10.1111/j.1558-5646.2008.00450.x

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., Meyre, D., 2019. Benefits and limitations of genome-wide association studies. Nat. Rev. Genet. 1. https://doi.org/10.1038/s41576-019-0127-1

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., Yang, J., 2017. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am. J. Hum. Genet. 101, 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Wray, N.R., Wijmenga, C., Sullivan, P.F., Yang, J., Visscher, P.M., 2018. Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. Cell 173, 1573–1580. https://doi.org/10.1016/j.cell.2018.05.051

Wright, S., 1934. Physiological and Evolutionary Theories of Dominance. Am. Nat. 68, 24–53. https://doi.org/10.1086/280521

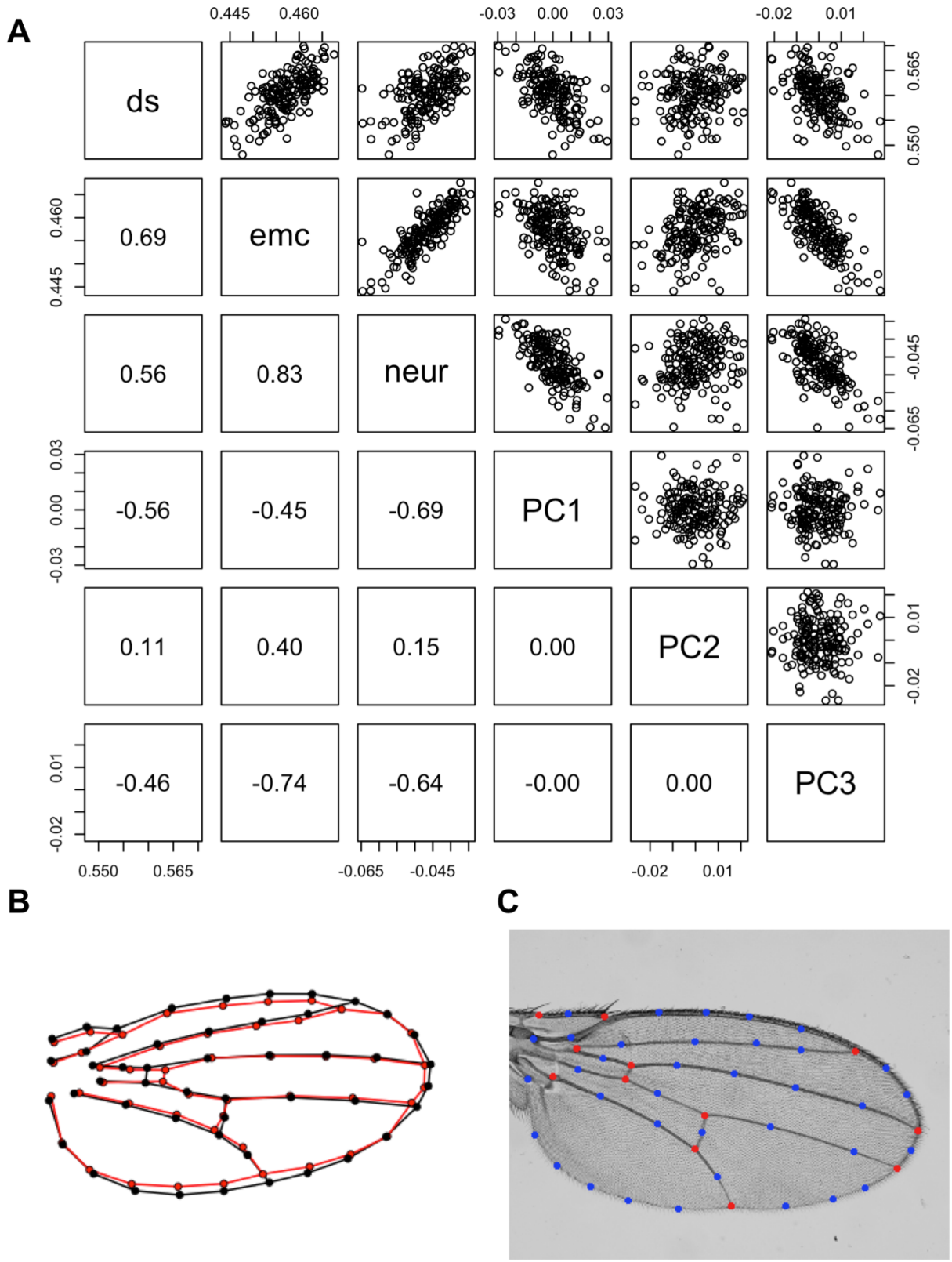

A
ds
emc
neur
PC1
PC2
PC3
0.69
0.56
0.83
-0.56
-0.45
-0.69
0.11
0.40
0.15
0.00
-0.46
-0.74
-0.64
-0.00
0.00
0.445
0.460
-0.03
0.00
0.03
-0.02
0.01
0.550
0.565
-0.065
-0.045
-0.02
0.01
B
C

Figure 1. Projections of data onto RNAi shape change vectors are correlated with major axes of shape variation among DGRP strains. (A) Shape change vectors from RNAi titration experiments for *ds, emc* and *neur* were used, and DGRP line means were projected onto these vectors to calculate shape scores. Eigenvectors for the PCA were estimated based on the same DGRP line means. Vector correlations between shape change vectors from RNAi knockdown: *ds* – *emc*: 0.65, *ds* – *neur*: 0.03, *emc* – *neur*: 0.30. (B) Effect of *ds* shape change estimated from RNAi knockdown, effects not magnified. (C) Landmarks (red) and semi-landmarks (blue) used in geomorphic morphometric analysis on a *Drosophila* wing. PCs 1–3 account for 22%, 20% and 9% of the overall, among DGRP shape variance.

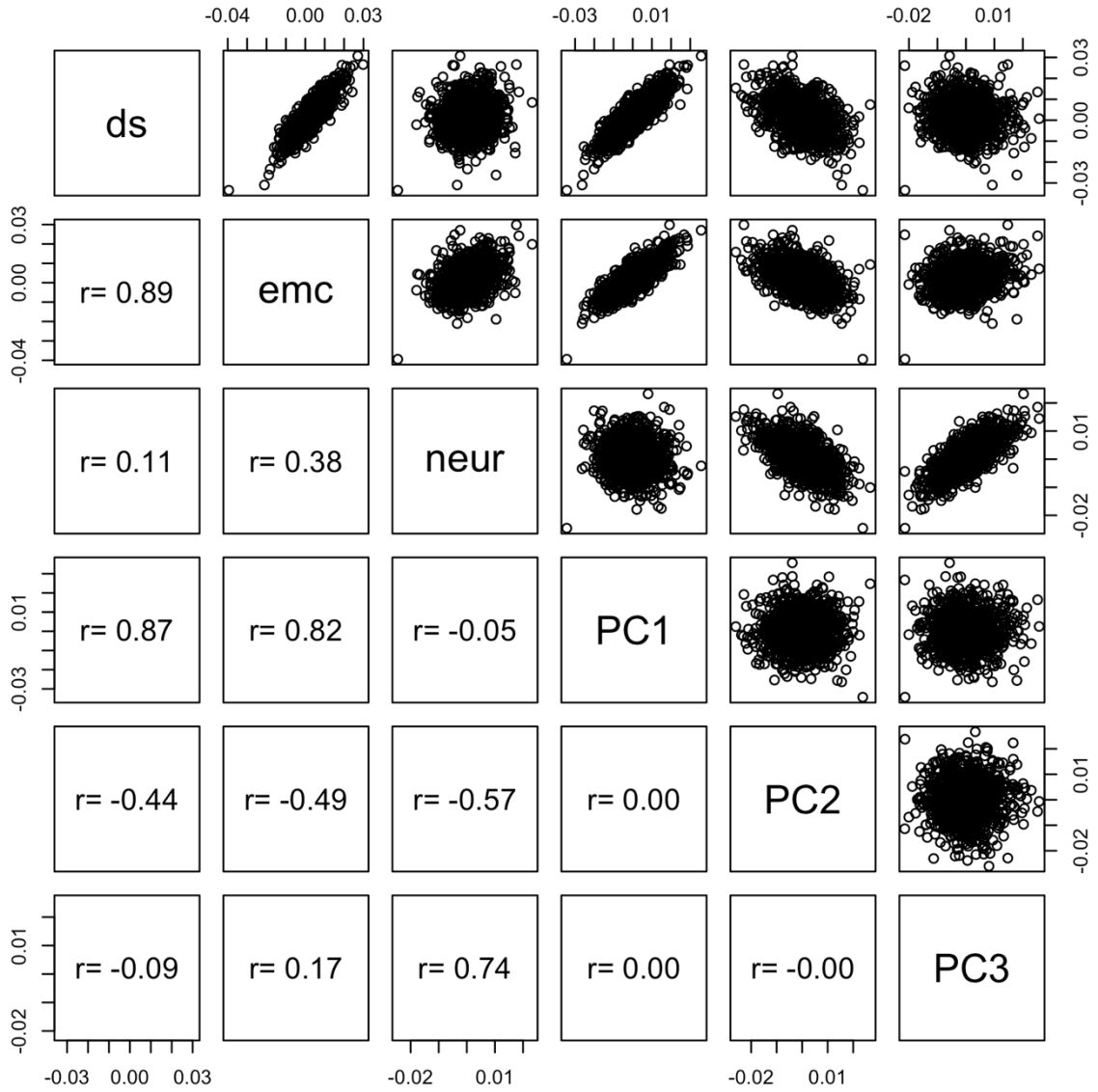


Figure 2. Projections of data onto RNAi shape change vectors are correlated with major axes of shape variation in wild-caught *Drosophila*. Correlations between projection of shape data from CMO population onto *ds, emc* and *neur* RNAi shape change vectors, and the first three eigenvectors from the PCA, calculated from shape data from all samples in the CMO population. PC1–3 account for 24%, 18% and 9% of overall shape variance in the CMO population.

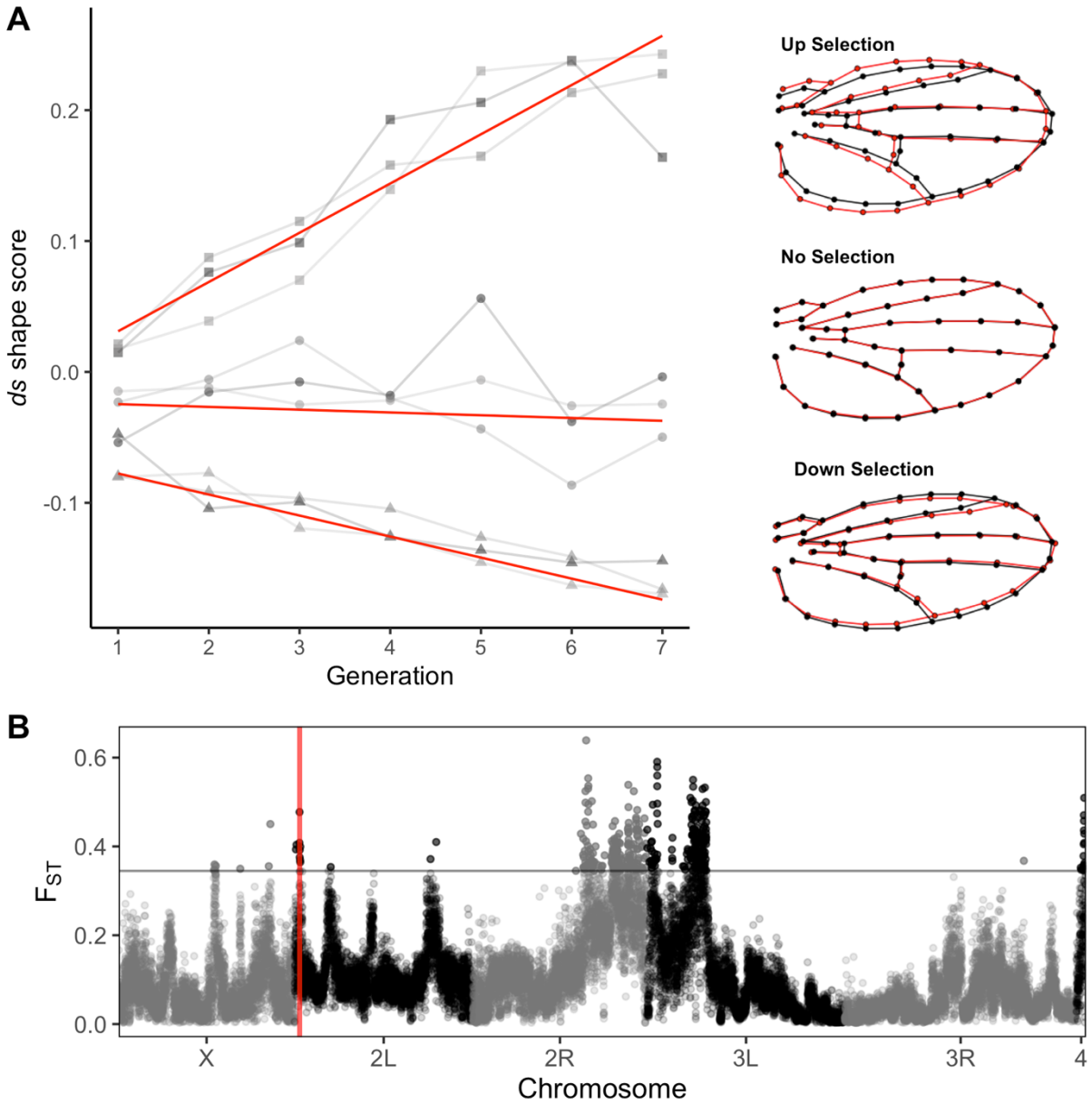


Figure 3. Artificial selection along *ds* shape change vector influences allele frequencies of variants at *ds*, and other hippo signalling genes. (A) Phenotypic response to selection based on *ds* shape change vector. Only data from females is plotted for ease of visualization. Each replicate of up (squares), control (dots) and down (triangles) selection lineages are plotted (greys). Estimated response to selection shown along red lines. Wing plots represent the effect of selection on shape change between generation one and seven (red, effects not magnified). (B) Genomic differentiation (F_{ST}) between up and down selection treatments measured in 5000bp windows. Red line represents the location of the *ds* locus. Grey line represents 3sd from genome wide mean F_{ST}.

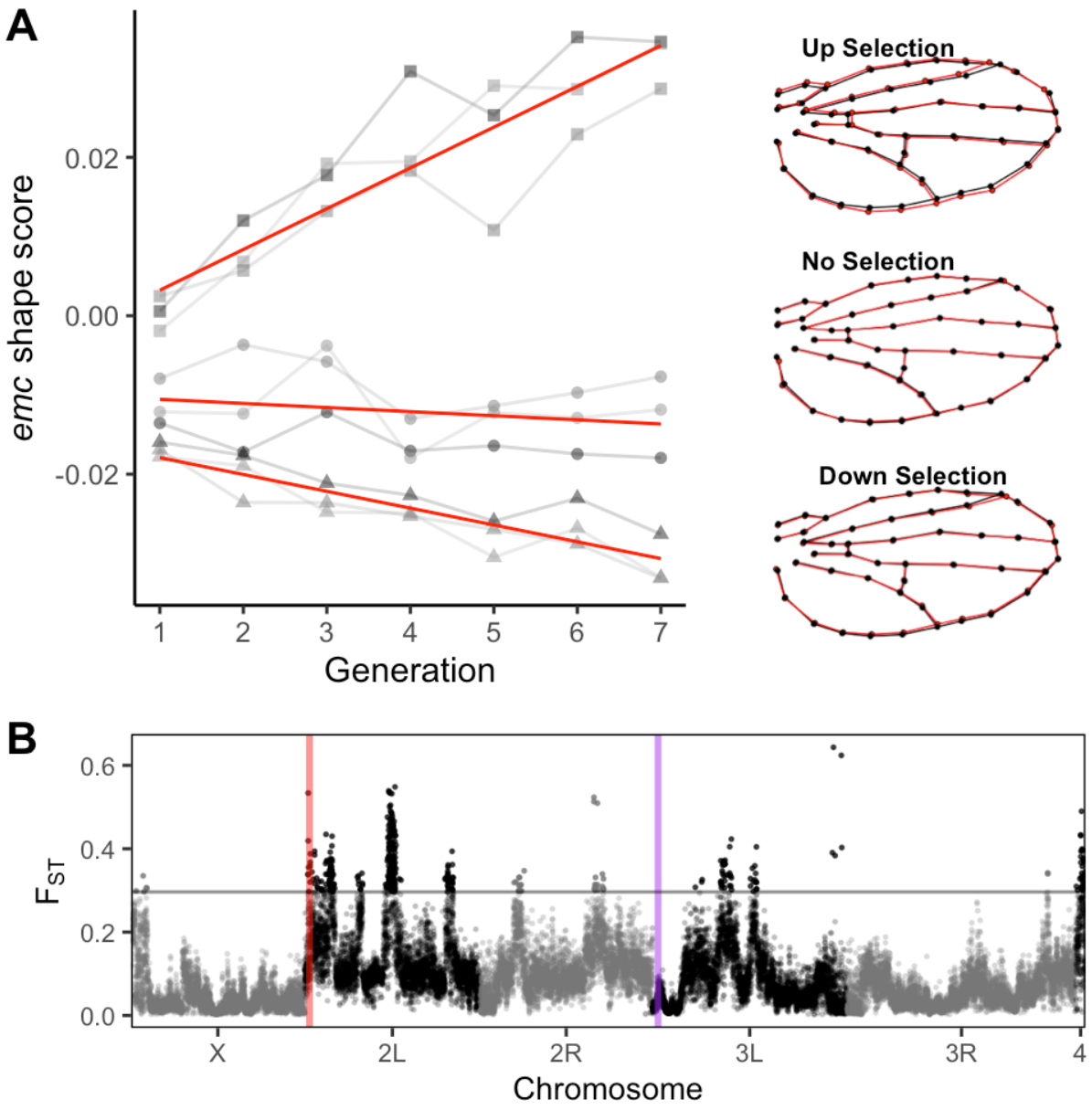


Figure 4. Artificial selection along *emc* shape change vector has modest influence on allele frequencies at *emc*, but a greater impact at the *ds locus*. (A) Phenotypic response to selection based on the *emc* shape change vector. Only data from females is plotted for ease of visualization. Each replicate of up (squares), control (dots) and down (triangles) selection lineages are plotted in greys. Estimated response to selection shown along red lines. Shape change between generation 1 and 7 is indicated on the right. Shape effects have been magnified 5x. (B) Genomic differentiation between up and down selection lineages (F_{ST}) measured in 5000bp sliding windows. Red and purple vertical lines represents genomic locations of *ds* and *emc* respectively. Grey line represents 3sd from genome wide mean F_{ST}.

.

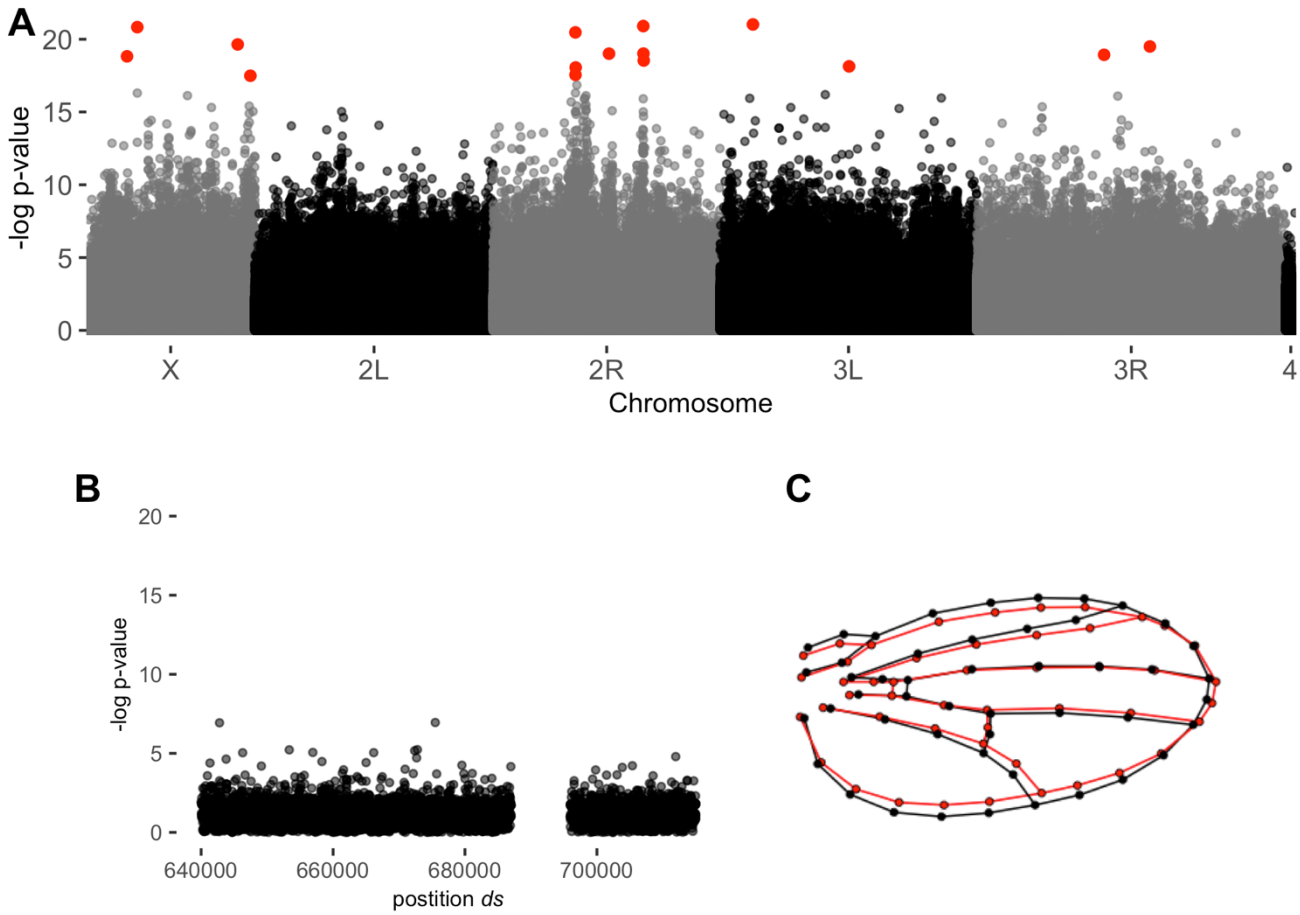


Figure 5. Genetic Differentiation between pools selected based on *ds* shape change among the wild-caught cohorts. (A) Genome-wide scan for differentiated loci between pools selected based on *ds* shape change vector using the CMH test implemented in ACER. Points in red indicate sites with significant differentiation. (B) Genomic differentiation at *ds* between pools selected based on *ds* shape change vector. No sites are significantly differentiated in *ds*. The large gap in sites is due to a masked region in the genome due to repetitive sequence and poor (syntenic) mapping scores. (C) Shape difference between selected pools of individuals from one representative (CMO) population, with the mean shape of pools represented in black and red.

Table 1. Variants from Pitchers et al. (2019) in *ds* artificial selection experiment. Estimated effect sizes for SNPs are estimated from a GWAS in the DGRP using LASSO regularized coefficients. Average frequency is given with replicate lineage frequencies in brackets. Estimated effect is the l^2-*norm* of shape differences associated with the variant. MAF = minor allele frequency.

Variant	Estimated Effect	DGRP MAF	Estimated MAF in synthetic outcross	Average allele frequency "up" selection	Average allele frequency "down" selection	Average allele frequency "control" selection
2L:655894	0.072	0.44	0.067	0 (0, 0, 0)	0 (0, 0, 0)	0.003 (0, 0, 0.0105)
2L:702560*	**0.159**	**0.056**	**0.06**	**0.995 (1, 0.98, 1)**	**0.446 (0.32, 0.35, 0.67)**	**0.705 (0.69, 0.56, 0.87)**
2L:702798	0.101	0.089	0.1	0.007 (0, 0.0217, 0)	0 (0, 0, 0)	0.005 (0, 0, 0.139)
2L:718623	0.225	0.033	0	0 (0, 0, 0)	0 (0, 0, 0)	0 (0, 0, 0)
2L:718627	0.11	0.033	0	0 (0, 0, 0)	0 (0, 0, 0)	0 (0, 0, 0)

* This is a complex polymorphism with linked SNPs and INDELs, in Pitchers *et al* (2019) a SNP in this region was found to be linked. However, the variant calling pipeline used in this work recognized an INDEL in this region which was used for counting.

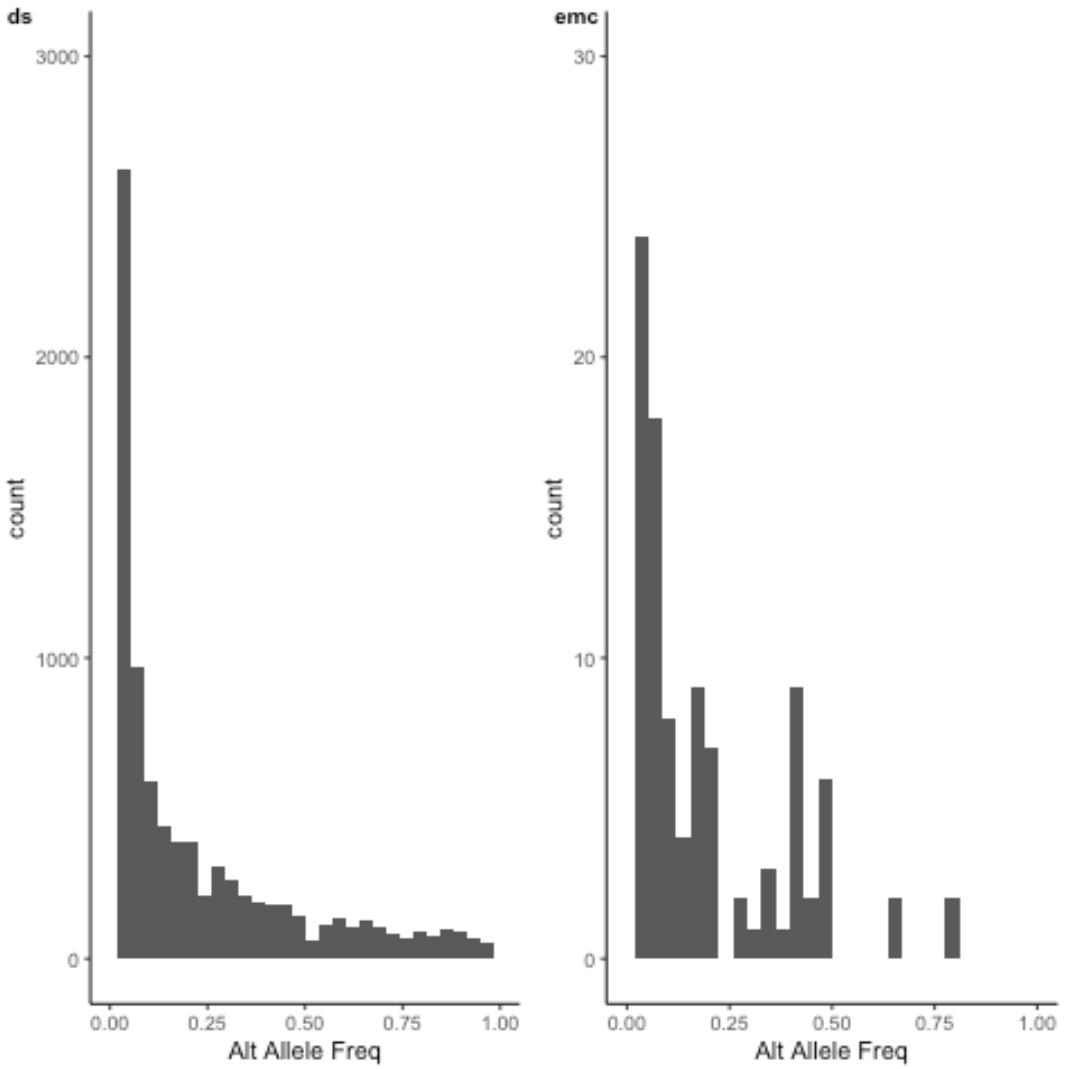
Table 2. *ds* Variants from Pitchers et al. (2019) in wild-caught cohorts used in the present study. Estimated effect sizes for SNPs are estimated from the DGRP GWAS with LASSO regularized coefficients. MAF in wild cohorts was estimated from sequenced pools of 75 random individuals.

Variant	Estimated Effect (Pitchers 2019)	DGRP MAF	Estimated MAF CMO	Estimated MAF FVW13	Estimated MAF FVW14	Estimated MAF PHO
2L:655894	0.072	0.445	0	0	0	0
2L:702560*	0.159	0.056	0.375	0.473	0.485	0.336
2L:702798	0.101	0.089	0.077	0.101	0.044	0.034
2L:718623	0.225	0.033	0.051	0.021	0.044	0.100
2L:718627	0.11	0.033	0.055	0.020	0.046	0.099

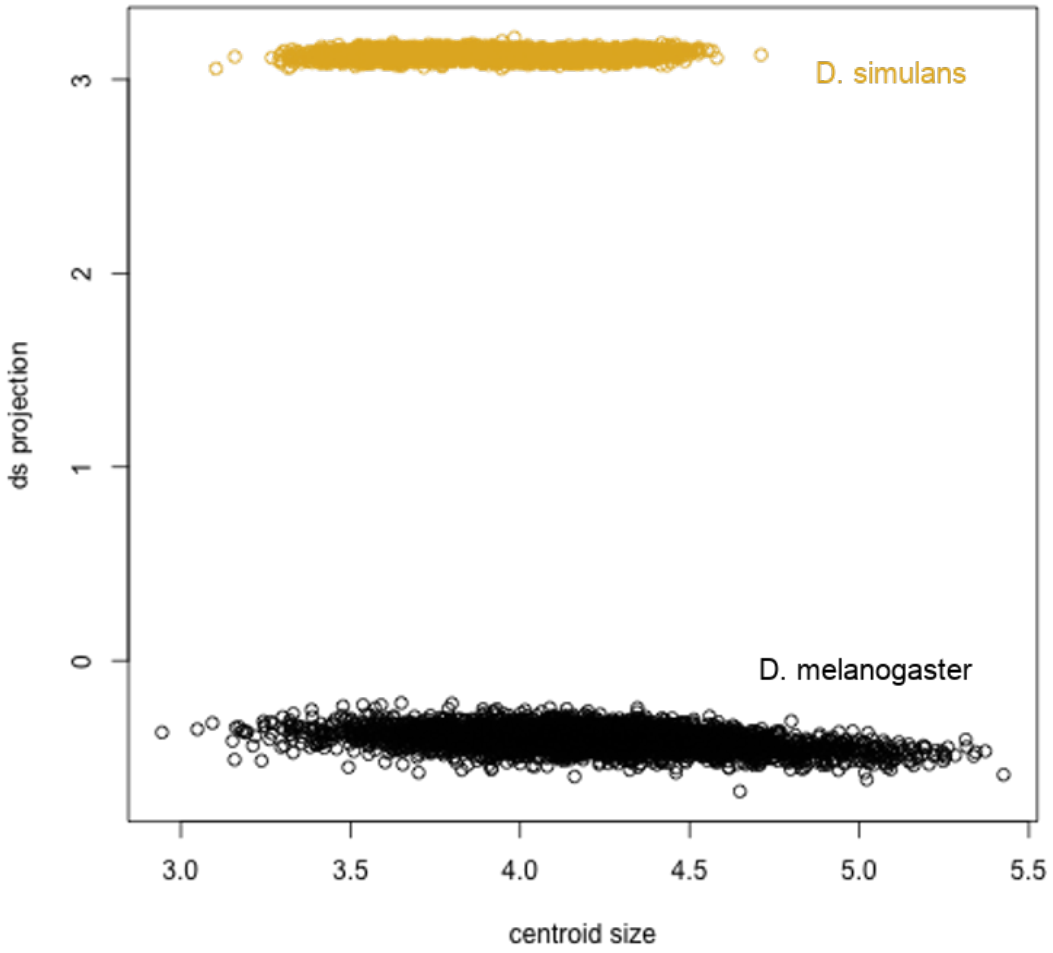
* This is a complex polymorphism with linked SNPs and INDELs, in Pitchers *et al* (2019) a SNP in this region was found to be linked. However, the variant calling pipeline used in this work recognized an INDEL in this region which was used for counting.

Table 3. Significantly differentiated variants for *ds* shape change from the wild-caught cohorts (BSA).

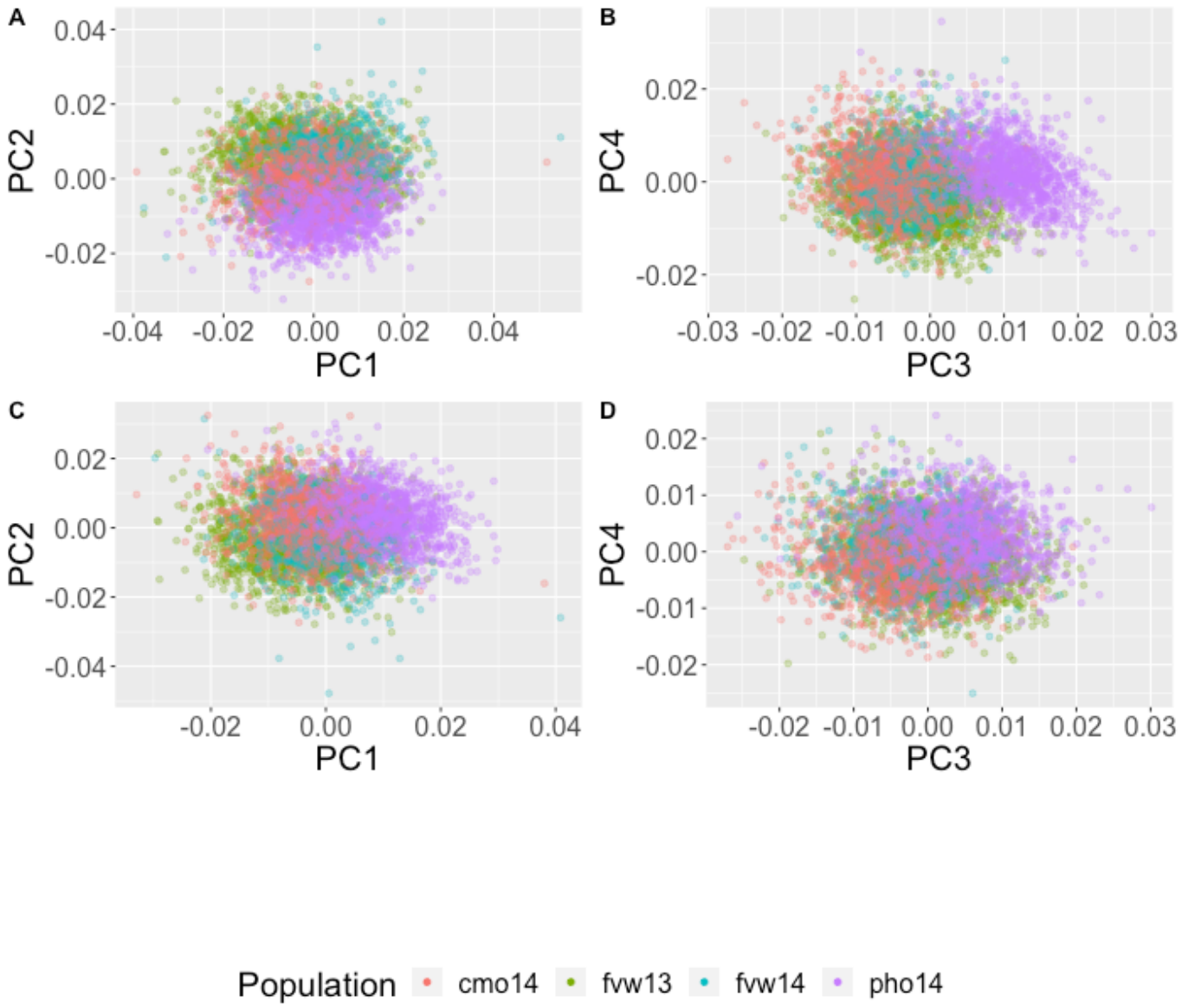
Location	CMH p-value (FDR corrected)	Gene	FlyBase ID	Distance from ORF (bp)
2R:17491270	0.026	*NT5E-2*	FBgn0050104	0
2R17498059	0.034	*CG30103*	FBgn0050103	2061
2R:17515133	0.022	*CG4853*	FBgn0034230	0
2R: 20537878	0.013	*CG13423*	FBgn0034513	0
2R:23601278	0.005	*CG10332*	FBgn0260455	0
2R:23601278	0.005	*IM18*	FBgn0067903	0
2R:23613785	0.013	*Eglp4*	FBgn0034885	0
2R:23613785	0.013	*Eglp2*	FBgn0034883	0
2R: 23646252	0.016	*retn*	FBgn0004795	0
3L:12831924	0.005	*CG10960*	FBgn0036316	0
3L: 20999119	0.022	*skd*	FBgn0003415	0
3R: 21523866	0.013	*CG7956*	FBgn0038890	0
3R: 2559549	0.011	*Pzl*	FBgn0267430	0
X: 14891220	0.013	*Flo2*	FBgn0264078	0
X: 14891220	0.013	*CG9514*	FBgn0030592	0
X:16039731	0.017	*Muc14a*	FBgn0052580	0
X: 793052	0.011	*CG16989*	FBgn0025621	95
X: 9448676	0.034	*mgl*	FBgn0261260	0



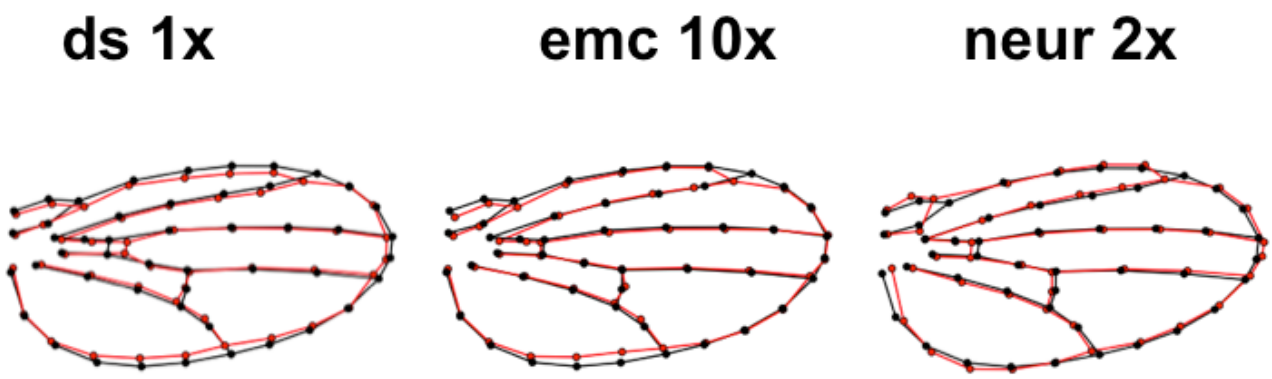
Supplemental Figure 1. Allele Frequency spectra demonstrate more variation at *ds* compared with *emc* in the synthetic outbred population used for artificial selection. Figures show estimated alternate allele frequencies at *ds* and *emc*. Alternate allele frequencies are estimated using parental strain genotype data and assuming an equal contribution from each parent to the founding population. Note the different y axis scale between *ds* and *emc*.



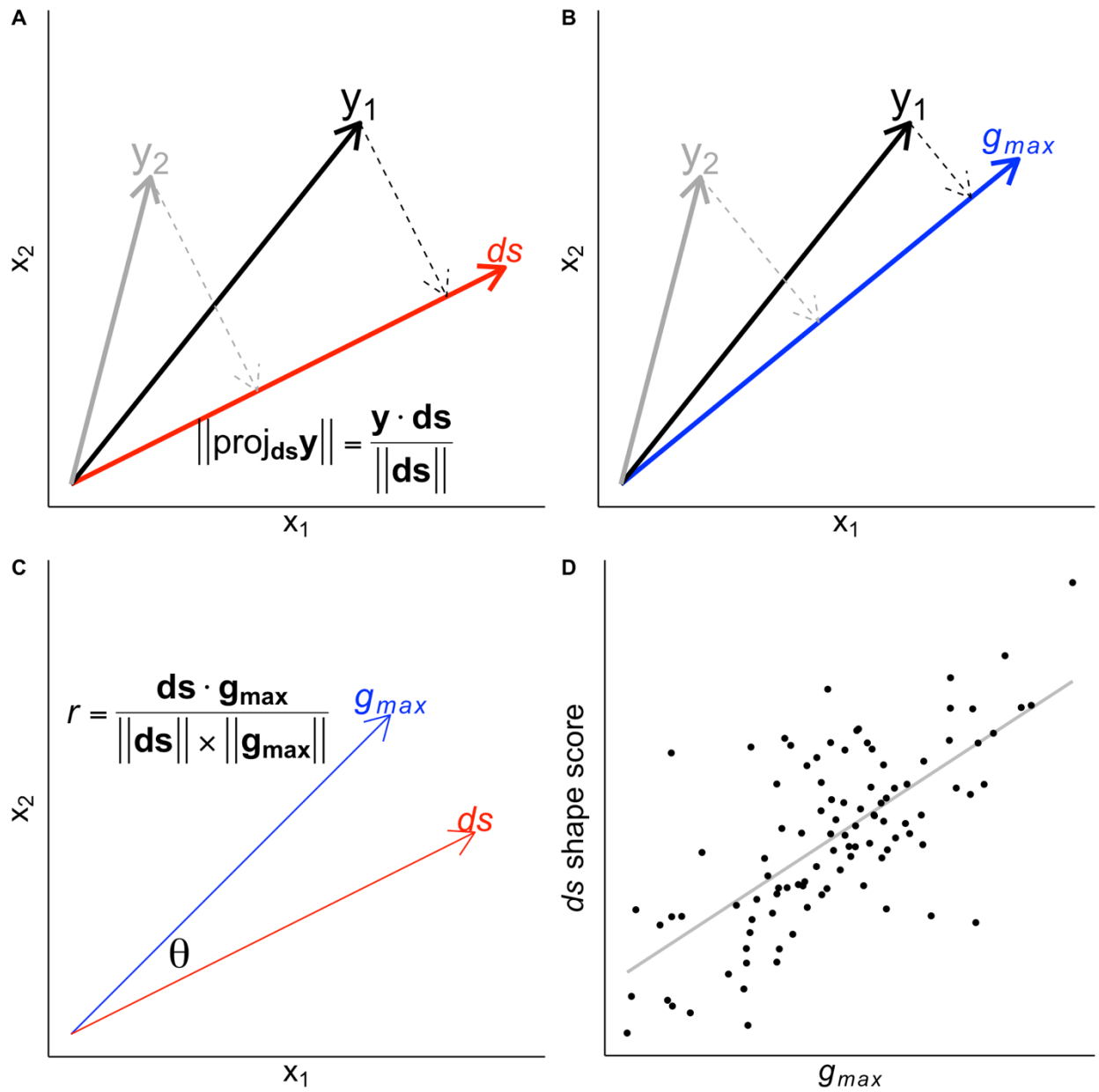
Supplemental Figure 2. Projection of FVW14 wings onto *ds* shape change vector shows clear distinction between *D. melanogaster* and *D. simulans.* Clear separation between the FVW14 samples (black) and *D. simulans* data (gold) indicates that *D. melanogaster* females were accurately identified.



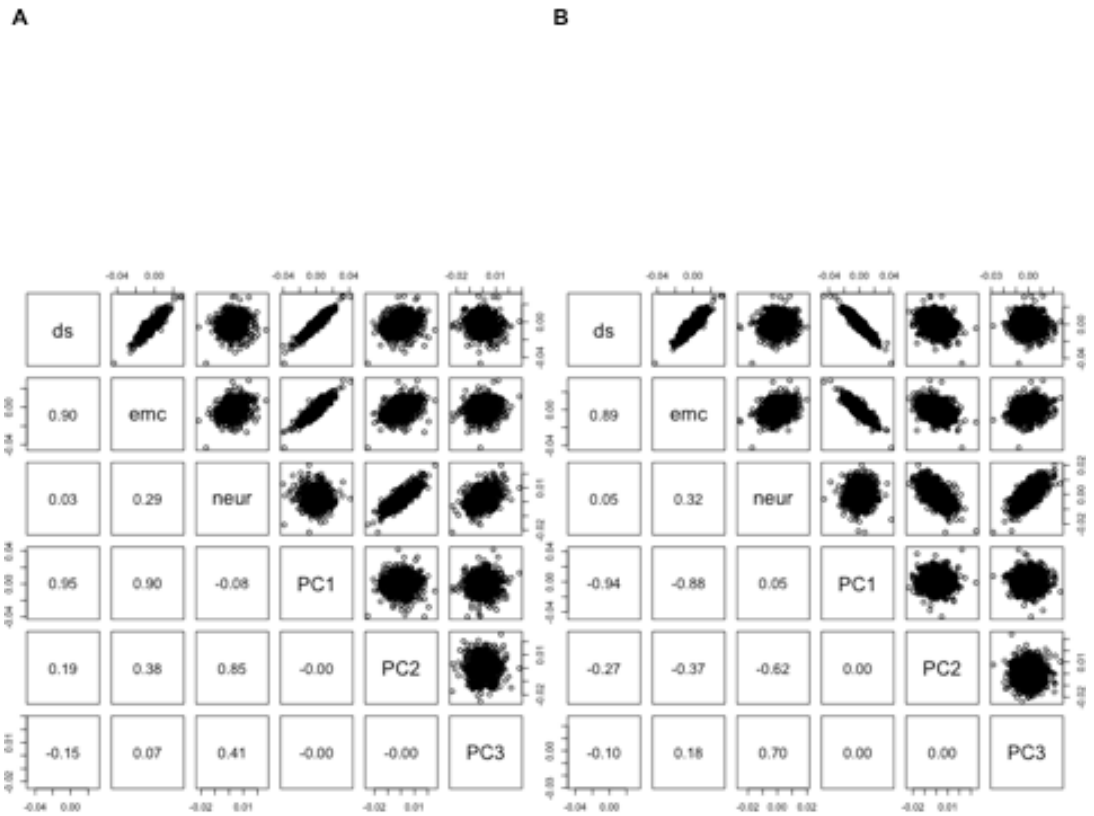
Supplemental Figure 3. Principal component analysis of shape variation within and among populations for wild collected *Drosophila*. First four axes from the PCA for shape variation are shown. (A) and (B) PCA includes all shape variation. (C) and (D) use the 'allometry corrected' landmarks, (residuals from a model regressing landmarks onto centroid size of wings).



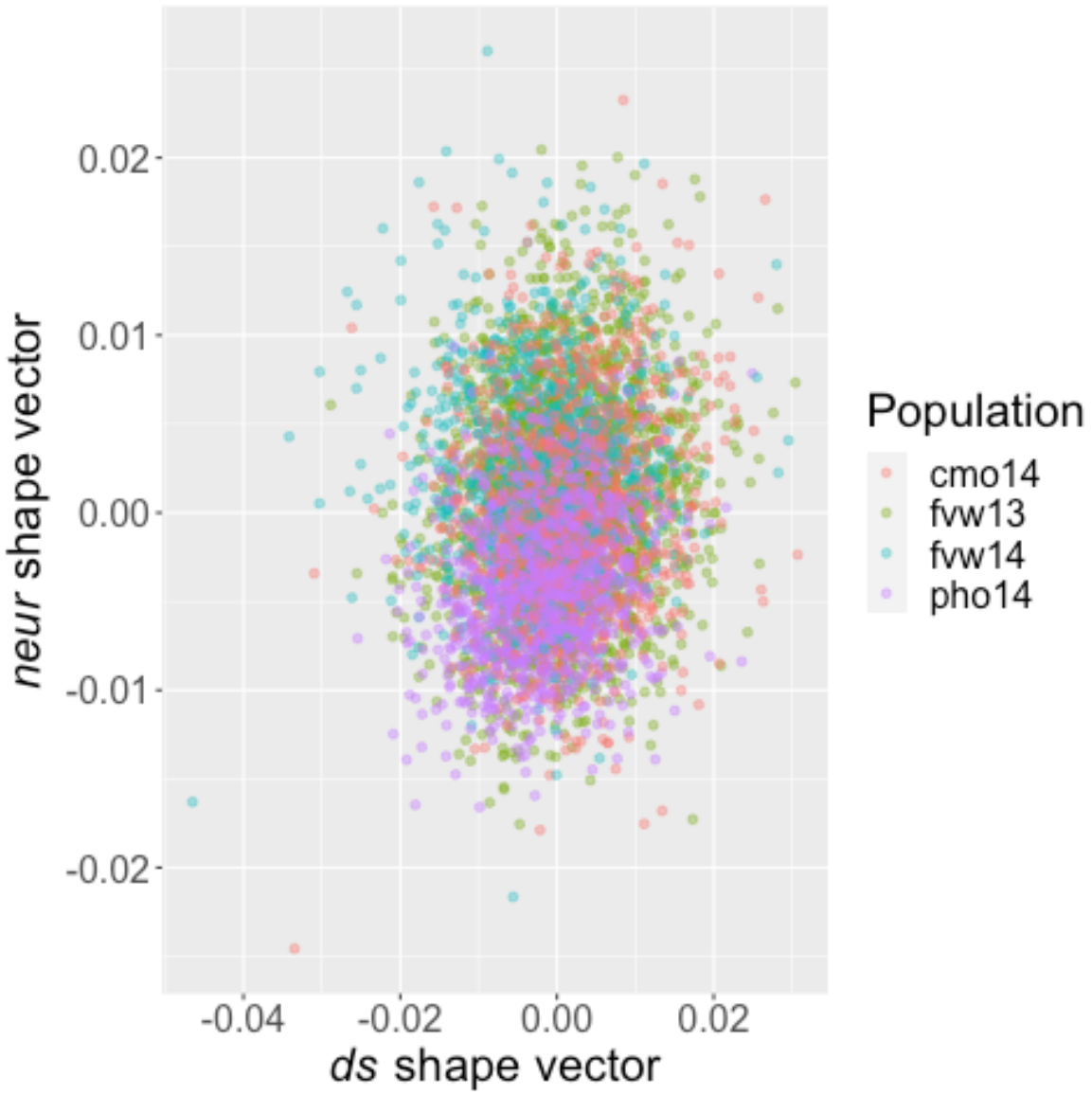
Supplemental Figure 4. Shape change effects due to RNAi knockdown of *ds*, *emc* and *neur.* Scaling of effects provided to aid in visualization of shape changes. Different magnifications are provided to account for the disparate magnitudes for estimated shape change vectors: *ds* = 5.5, *emc* = 0.44, *neur* = 2.8. The vectors from these analyses were used for projections in this study.



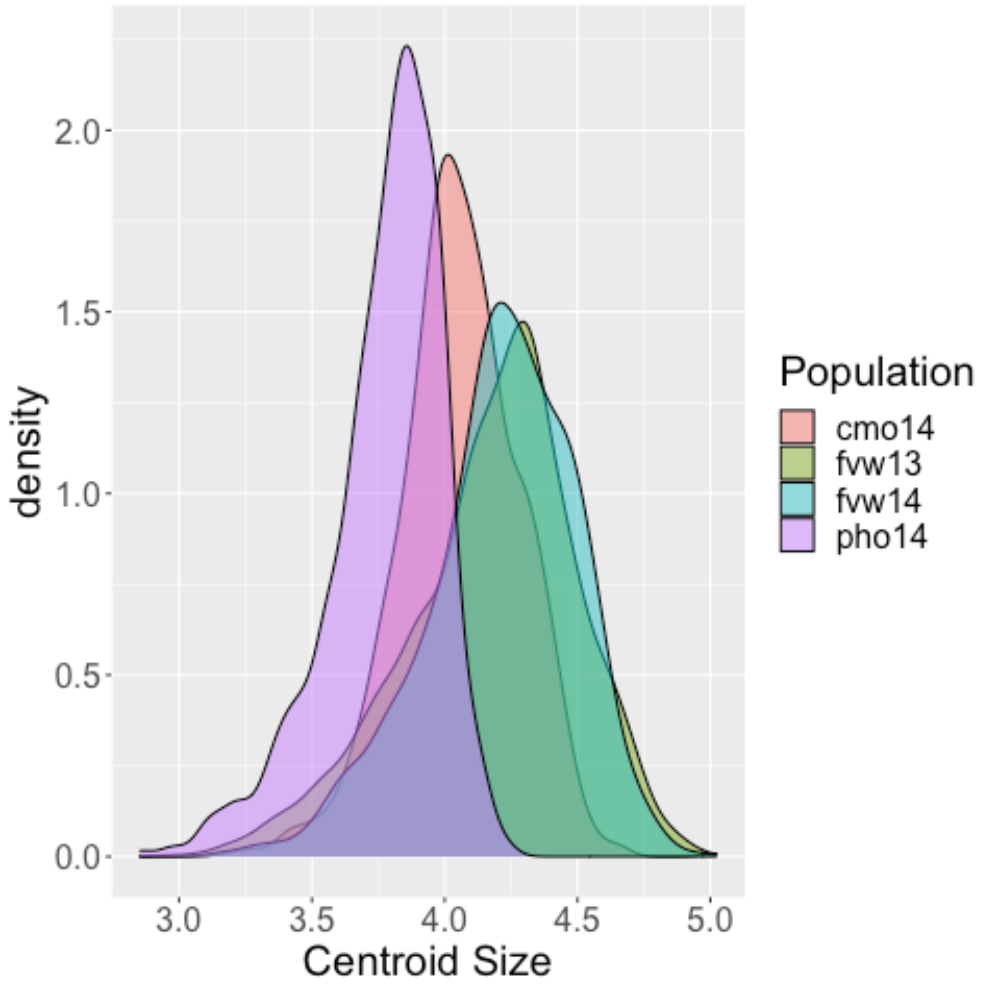
Supplemental Figure 5. Illustration of projections onto shape vectors to generate shape scores used in this study. The *ds* shape change vector is used for demonstration. (A) Calculating shape score using projection. y_1 and y_2 represent vectors of landmarks for two representative individuals. Dotted arrows represent projection of y_1 and y_2 onto the shape change vector defined by *ds* RNAi to generate the *ds* shape score. (B) Similarly, y_1 and y_2 are projected onto the $\boldsymbol{g_{max}}$ (PC1) vector, representing the direction of maximum genetic variation in the genetic variance-covariance matrix, ***G***. (C) Comparing $\boldsymbol{g_{max}}$ and shape change vectors, using the correlation (*r*) of vectors directly, or via the angle θ, between vectors. (D) Hypothetical relationship between *ds* shape scores and $\boldsymbol{g_{max}}$, indicating a relationship between directions of *ds* induced shape change and the direction of genetic variation in shape variation.



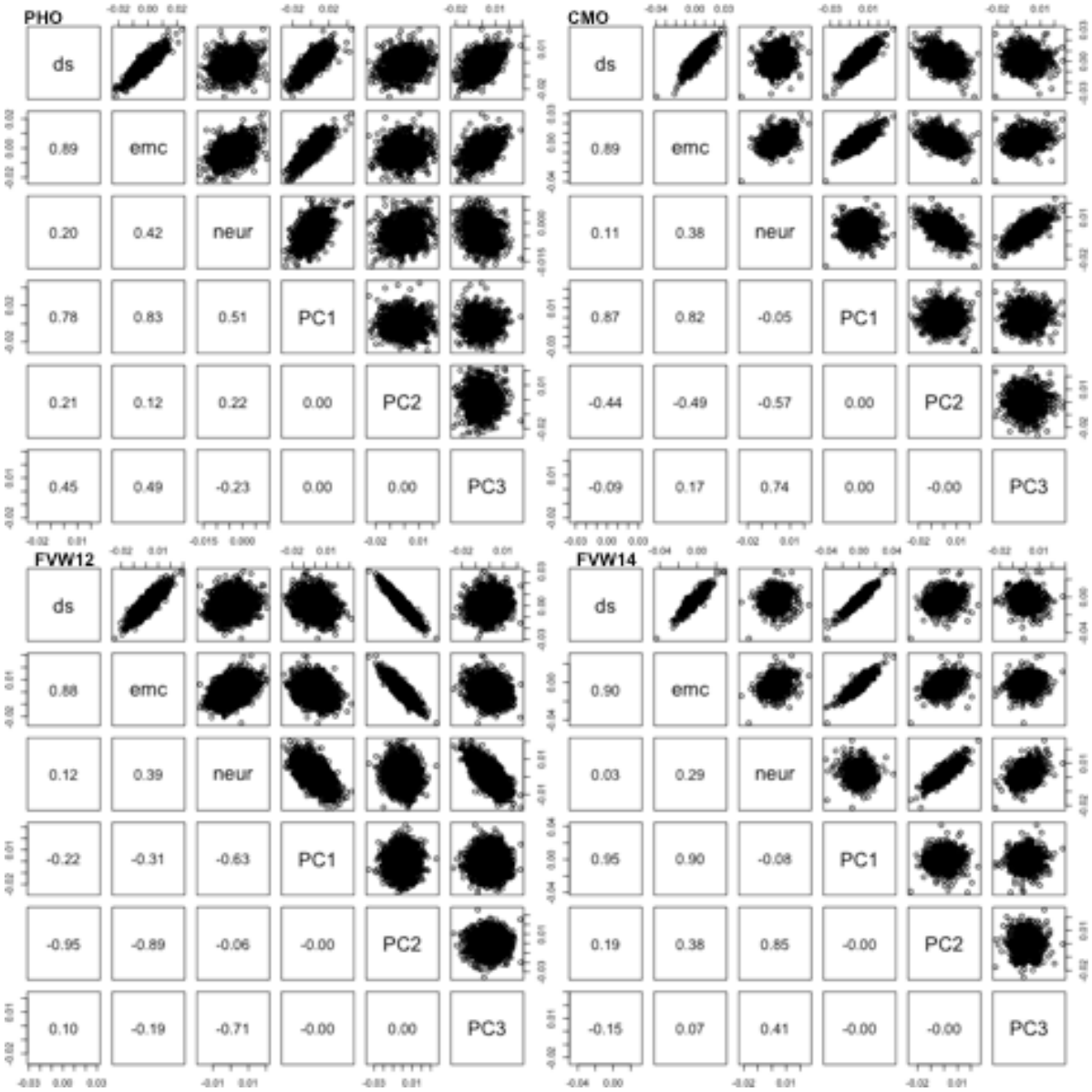
Supplemental Figure 6. Including wild-caught females in the analyses does not change interpretation for FVW14 data. Correlation between projection of shape data from FVW14 population onto *ds, emc* and *neur* shape change vectors and the first three PCs are calculated from shape data from all wings in the FVW14 population. (A) male only data, (B) females and male data. Inclusion of female data does not change the conclusions drawn from the relationships between directions shape change vectors and PCs. Note that the "flipped" directions of the PCs between (A) and (B) represents the arbitrary sign of the set of eigenvectors for the PCA (i.e. the set of eigenvectors can equivalently be multiplied by -1).



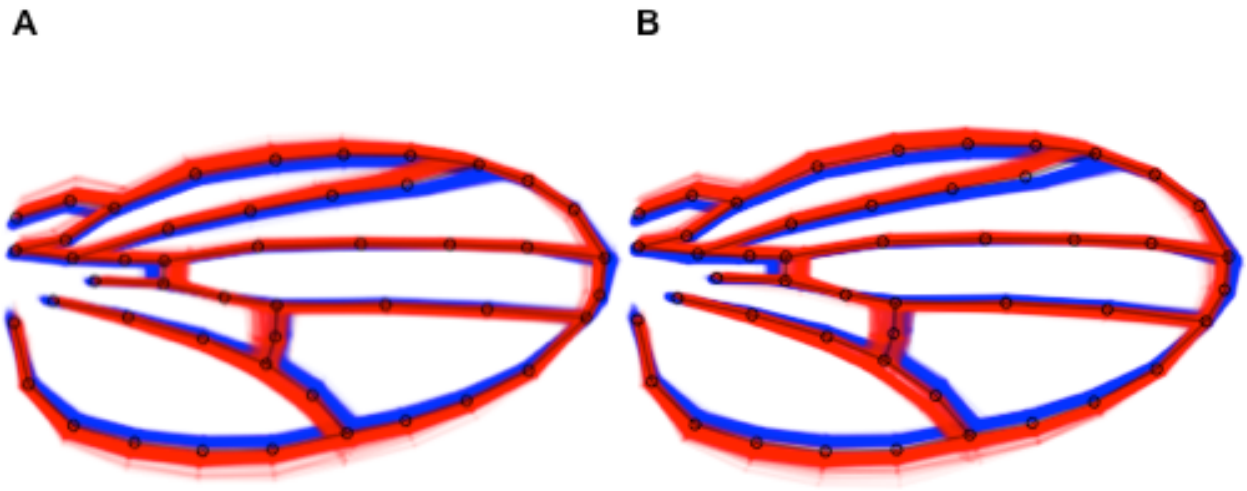
Supplemental Figure 7. Shape variation in field collected samples by projecting individual shape data onto *ds* and *neur* shape change vectors. Projections were performed used size "adjusted" landmark data onto *ds* and *neur* shape change vectors. The correlation between the ds and neur shape scores is r = 0.12. In linear models with shape scores (*ds* and *neur* respectively) regressed onto population and wing size, the partial R^2 for population effects is 0.040 (*ds*) and 0.18 (*neur*).



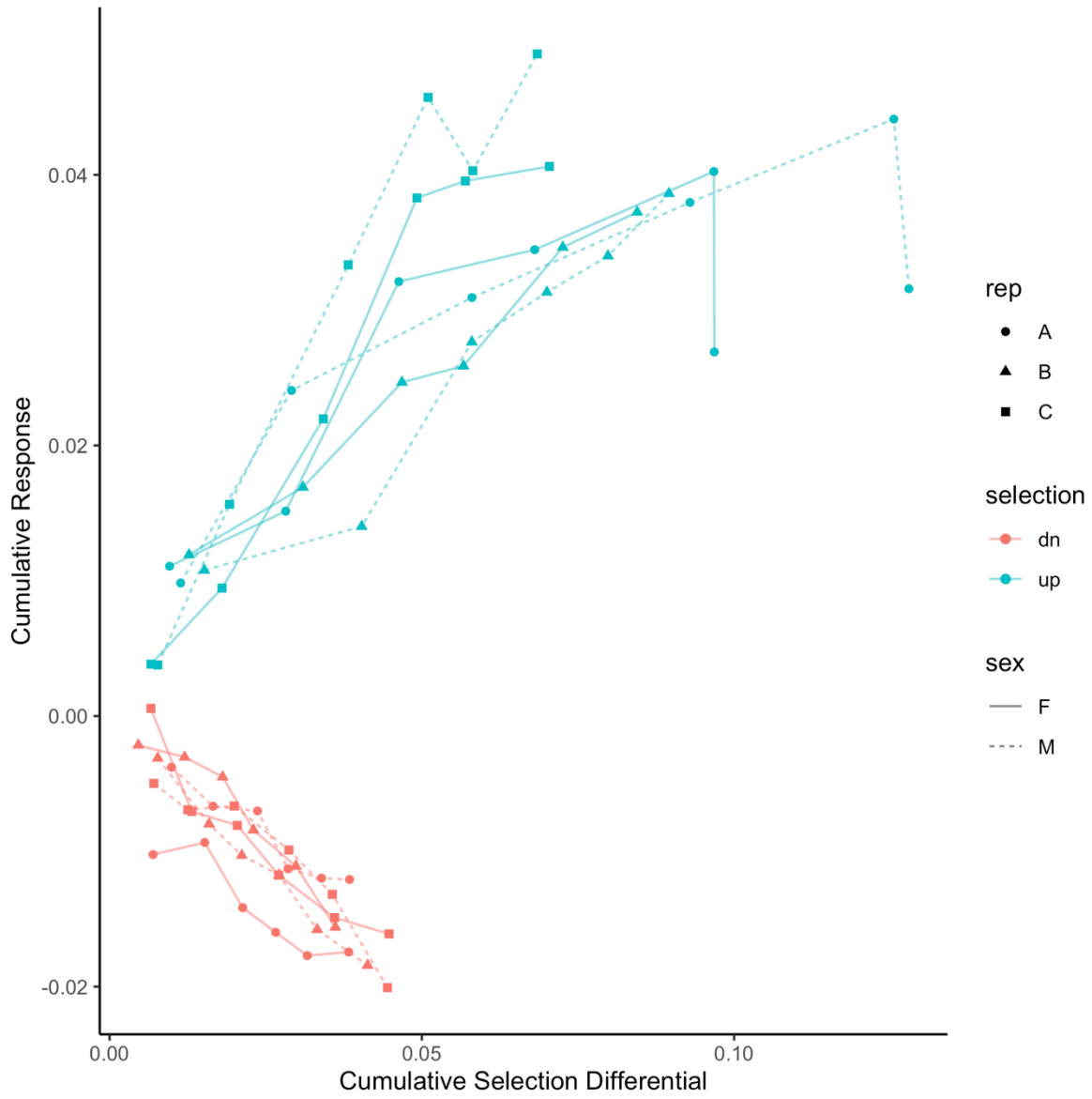
Supplemental Figure 8. Distribution of wing sizes (males only) in wild caught cohorts. When size is linearly regressed onto population, model $R^2 = 0.27$.



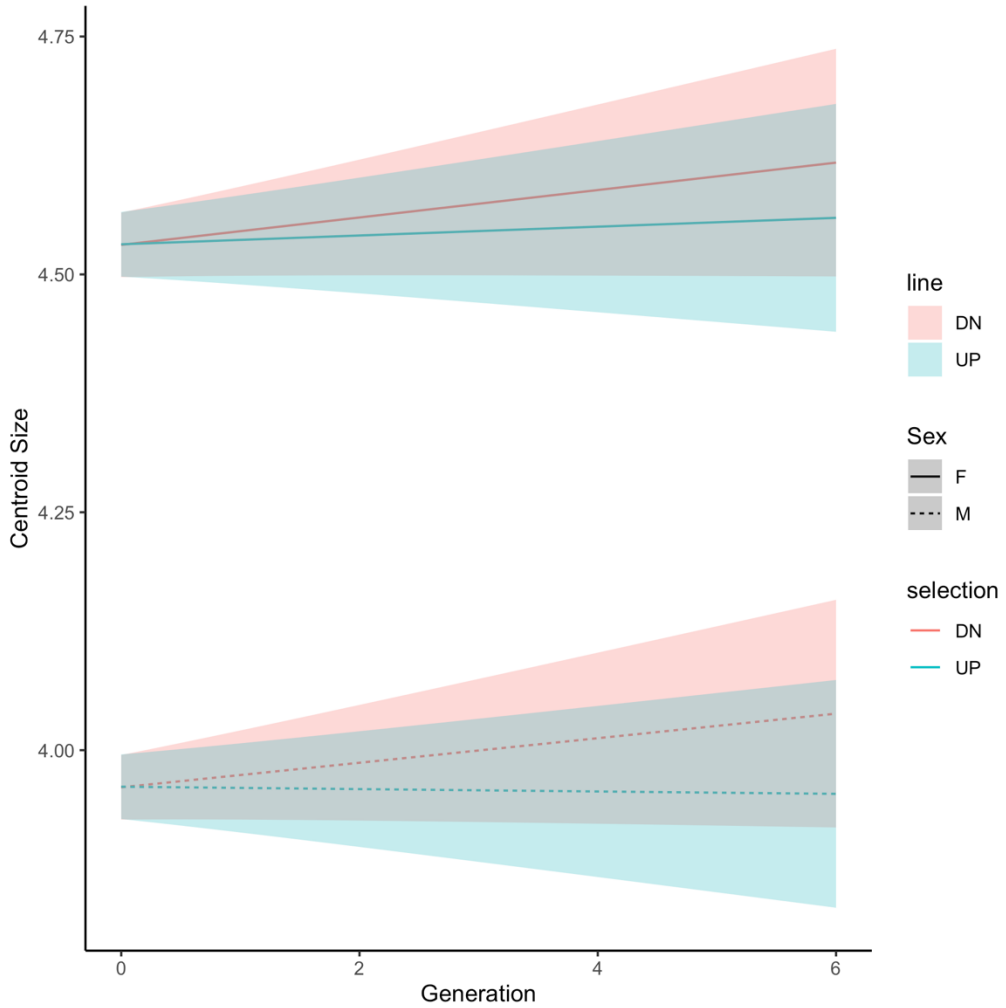
Supplemental Figure 9. **Projections of data onto RNAi shape change vectors are correlated with major axes of shape variation in wild-caught *Drosophila* from each population**. Correlation between projection of shape data from wild cohorts onto *ds, emc* and *neur* shape change vectors and the first three PCs are calculated from shape data from all wings in each cohort independently. Note that the "flipped" directions of the PCs for FVW12 represents the arbitrary sign of the set of eigenvectors for the PCA (i.e. the set of eigenvectors can equivalently be multiplied by -1). Note that the eigenvectors representing PC1 and PC2 in FVW12 have "swapped" because of similarities in variance accounted for.



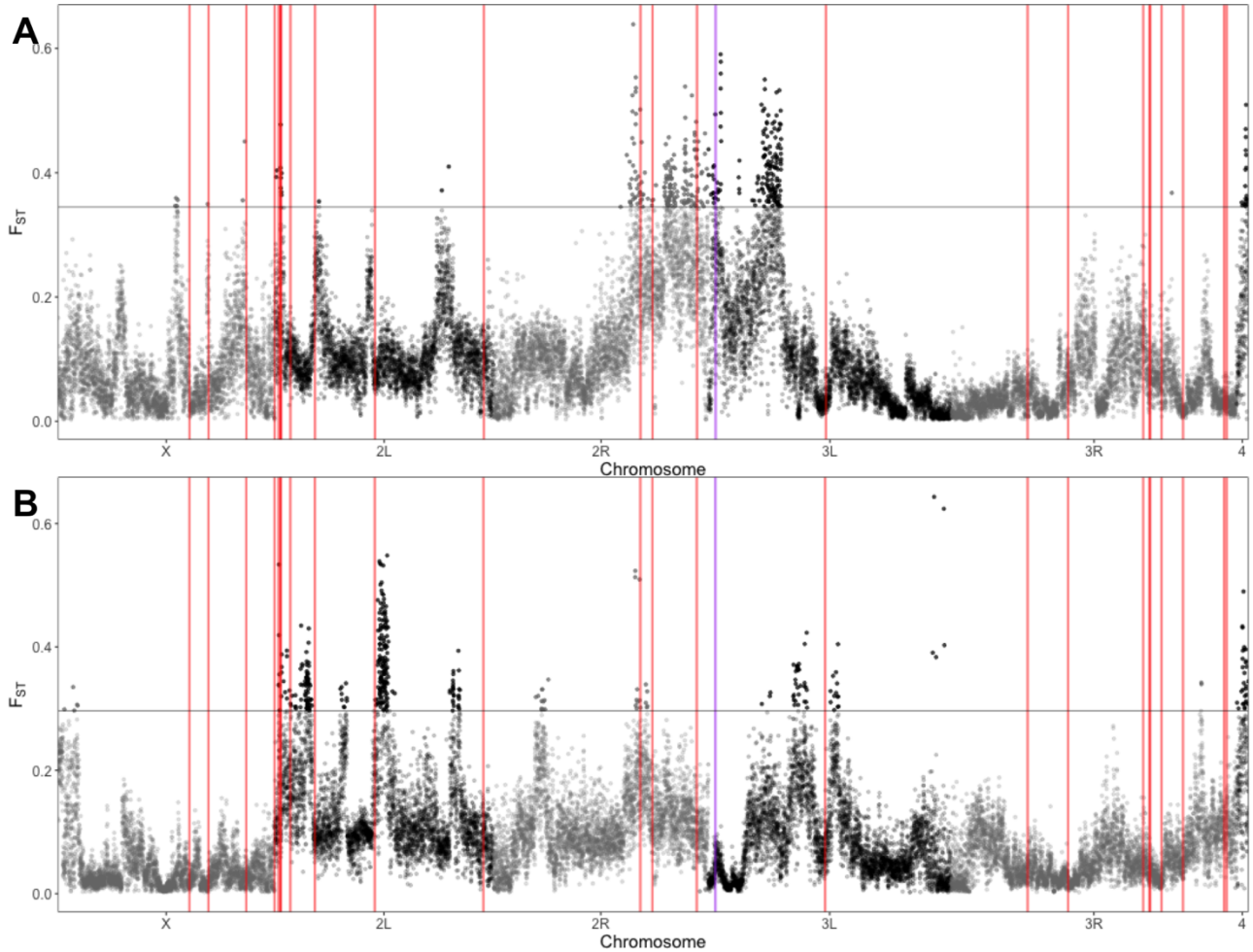
Supplemental Figure 10. Variation in wing shape among individuals from artificial selection along the *ds* shape change vector. All wings from generation 7 of up (red) and down (blue) selection lineages plotted, with females in (A) and males in (B). Black wireframe is mean shape from the experiment.



Supplemental Figure 11. Response to artificial selection along the *ds* shape change vector. Regression of cumulative selection differential onto cumulative response was used to estimate realized heritability in both treatments ("up" & "down") independently.



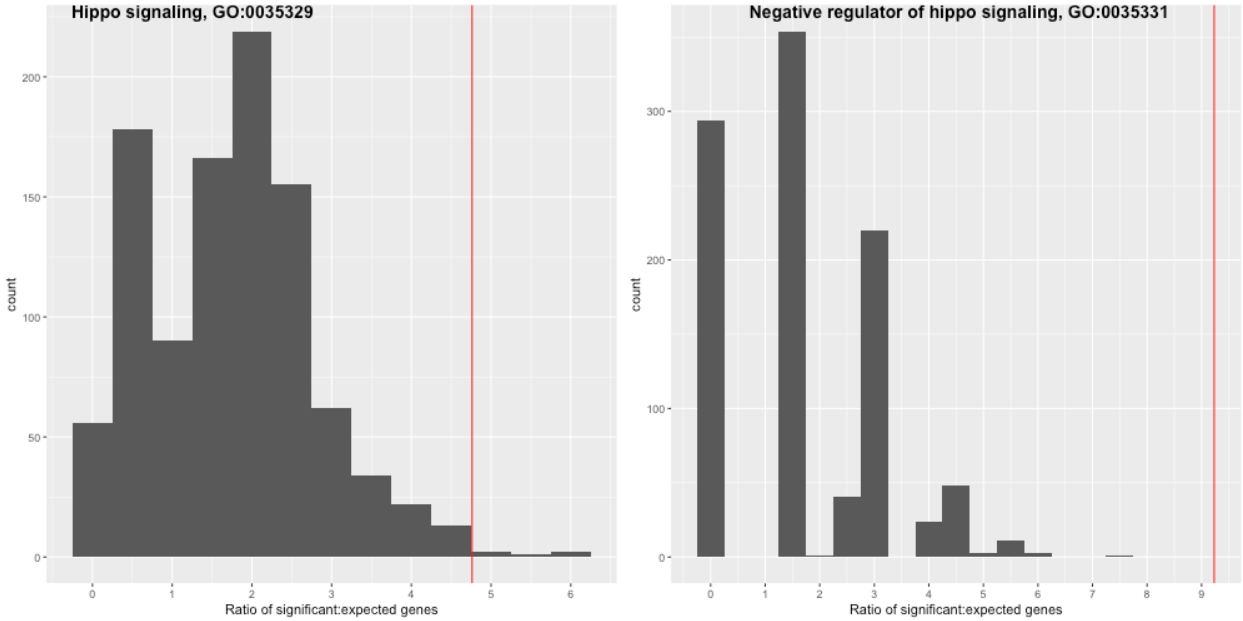
Supplemental Figure 12. Modest (and not significant) changes in wing size following artificial selection based on *ds* shape change. Mean size estimated for up and down *ds* selection lineages estimated from a linear mixed model with replicate lineage fitted as a random effect. Shaded regions represent 95% confidence bands for the correlated response to selection.



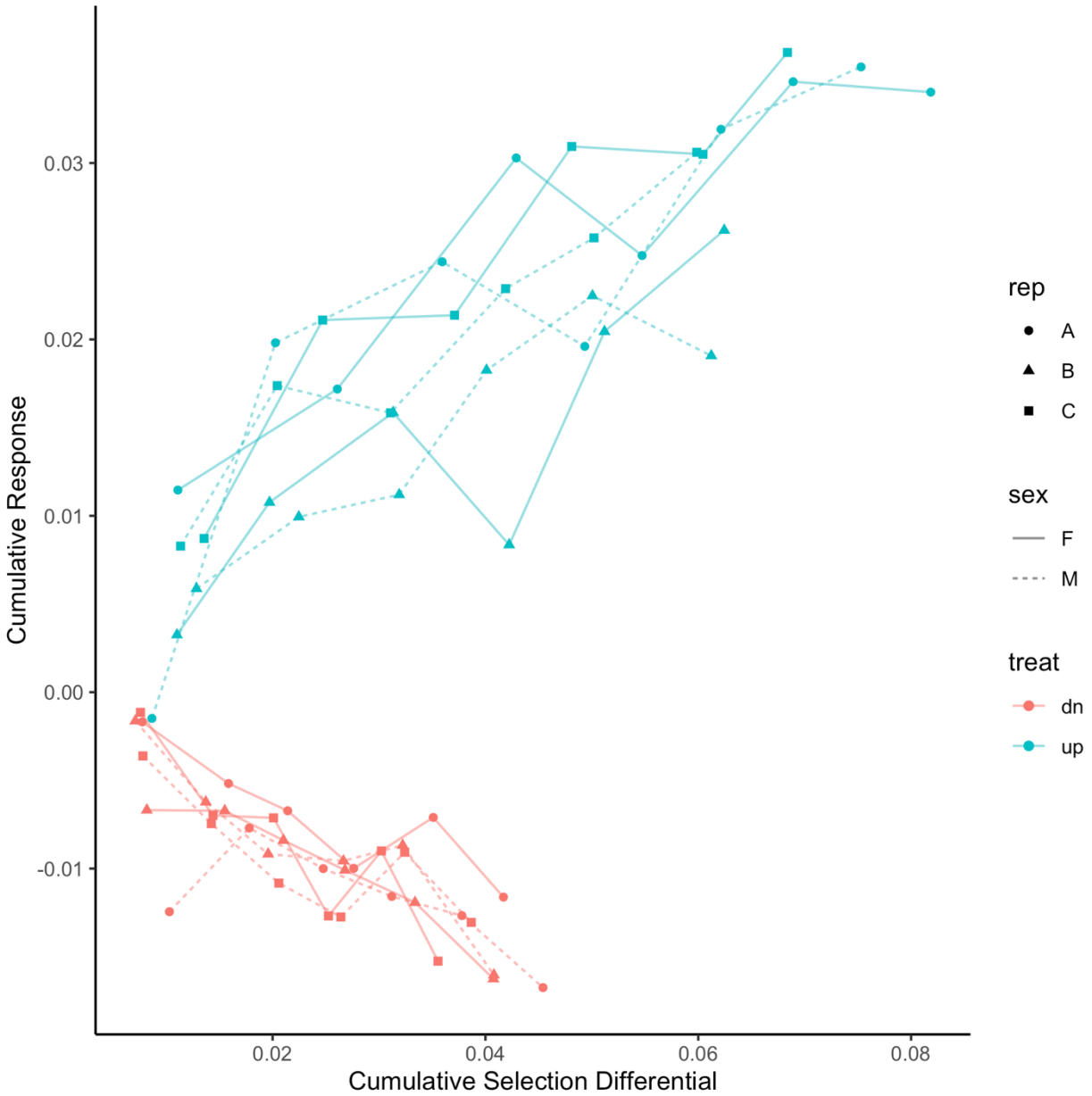
Supplemental Figure 13. Genetic differentiation between artificial selection pools (Figure 3B), with core hippo signaling loci (red) and *emc* (purple) marked. Genomic differentiation between up and down selection lineages (F_{ST}) measured in 5000bp windows for the artificial selection along *ds* (A) and *emc* shape change vectors (B). Horizonal grey lines represents 3sd from mean F_{ST}.

```
195     GGATTATATGGGGTTTACTAGAAGTGTATATGACCCATCTTAGTTGGCATATTTGTGCTT   108
28      GGATNATATGGGGTTTACTAGAAGTGTATATGACCCATCTTAGTTGGCATATTTGTGCTT   106
96      GGATNATATGGGGTTTACTAGAAGTGTATATGACCCATCTTAGTTGGCATATTTGTGCTT   106
48      GGATNATATGGGGTTTACTAGAAGTGTATATGACCCATCTTAGTTGGCATATTTGTGCTT   106
59      GGATTATATGGGGTTTACTAGAAGTGTATATGACCCATCTTAGTTGGCATATTTGTGCTT   109
801     GGATTATATGGGGTTTACTAGAAGTGTATATGACCCATCTTAGTTGGCATATTTGTGCTT   108
ref     ggattatatggggtgtacaagaagtgtatatgactcatcttag-----------------   163
129     GGATTATATGGGGTGTACTAGAAGTGTATATGACTTATCTTAG-----------------   102
301     GGATTATATGGGGTGTACTAGAAGTGTATATGACTTATCTTAG-----------------   100
69      GGATTATATGGGGTGTACTAGAAGTGTATATGACTTATCTTAG-----------------   102
385     GGATTATATGGGGTGTACTAGAAGTGTATATGACTTATCTTAG-----------------   100
75      GGATTATATGGGGTGTACTAGAAGTGTATATGACTCATCTTAG-----------------   101
83      GGATTATATGGGGTGTACTAGAAGTGTATATGACTTATCTTAG-----------------   100
491     GGATTATATGGGGTGTACTAGAAGTGTATATGACTCATCTTAG-----------------   101
34      GGATTATATGGGGTGTACAAGAAGTGTATATGACTCATCTTAG-----------------   99
774     GGATTATATGGGGTGTACAAGAAGTGTATATGACTCATCTTAG-----------------   102
        **** ********* *** ***************  *******

195     AGGTGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGCCAAGAAAATTTCAAATT   168
28      AGGTGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGCCAAGAAAATTTCAAATT   166
96      AGGTGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGCCAAGAAAATTTCAAATT   166
48      AGGTGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGCCAAGAAAATTTCAAATT   166
59      AGGTGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGCCAAGAAAATTTCAAATT   169
801     AGGTGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGCCAAGAAAATTTCAAATT   168
ref     -ttggcatatttctaacaaaaatccgtaactattttgagtgtccaagaaaatttcaaa-t   221
129     -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAAAT-T   159
301     -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAAAT-T   157
69      -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAAAT-T   159
385     -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAATT-T   157
75      -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAAAT-T   158
83      -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAAAT-T   157
491     -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGGTAGAAA-ATTTCAAAT-T   158
34      -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGTCCAAGAAAATTTCAAA-T   157
774     -TTGGCATATTTCTAACAAAAATCCGTAACTATTTTGAGTGTCCAAGAAAATTTCAAA-T   160
            ************************************    * * * **  *   *

195     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTTGCATTGTTTGGCTAATCTACGTAAATCT   228
28      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTTGCATTGTTTGGCTAATCTACGTAAATCT   226
96      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTTGCATTGTTTGGCTAATCTACGTAAATCT   226
48      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTTGCATTGTTTGGCTAATCTACGTAAATCT   226
59      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTTGCATTGTTTGGCTAATCTACGTAAATCT   229
801     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTTGCATTGTTTGGCTAATCTACGTAAATCT   228
ref     ttttttgcgtgtgcttggatgttgtaatttcgcattgtttggctaatctacgtaaatct   281
129     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   219
301     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   217
69      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   219
385     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   217
75      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   218
83      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   217
491     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   218
34      TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   217
774     TTTTTTTGCGTGTGCTTGGATGTTGTAATTTCGCATTGTTTGGCTAATCTACGTAAATCT   220
        ****************************** ***************************
```

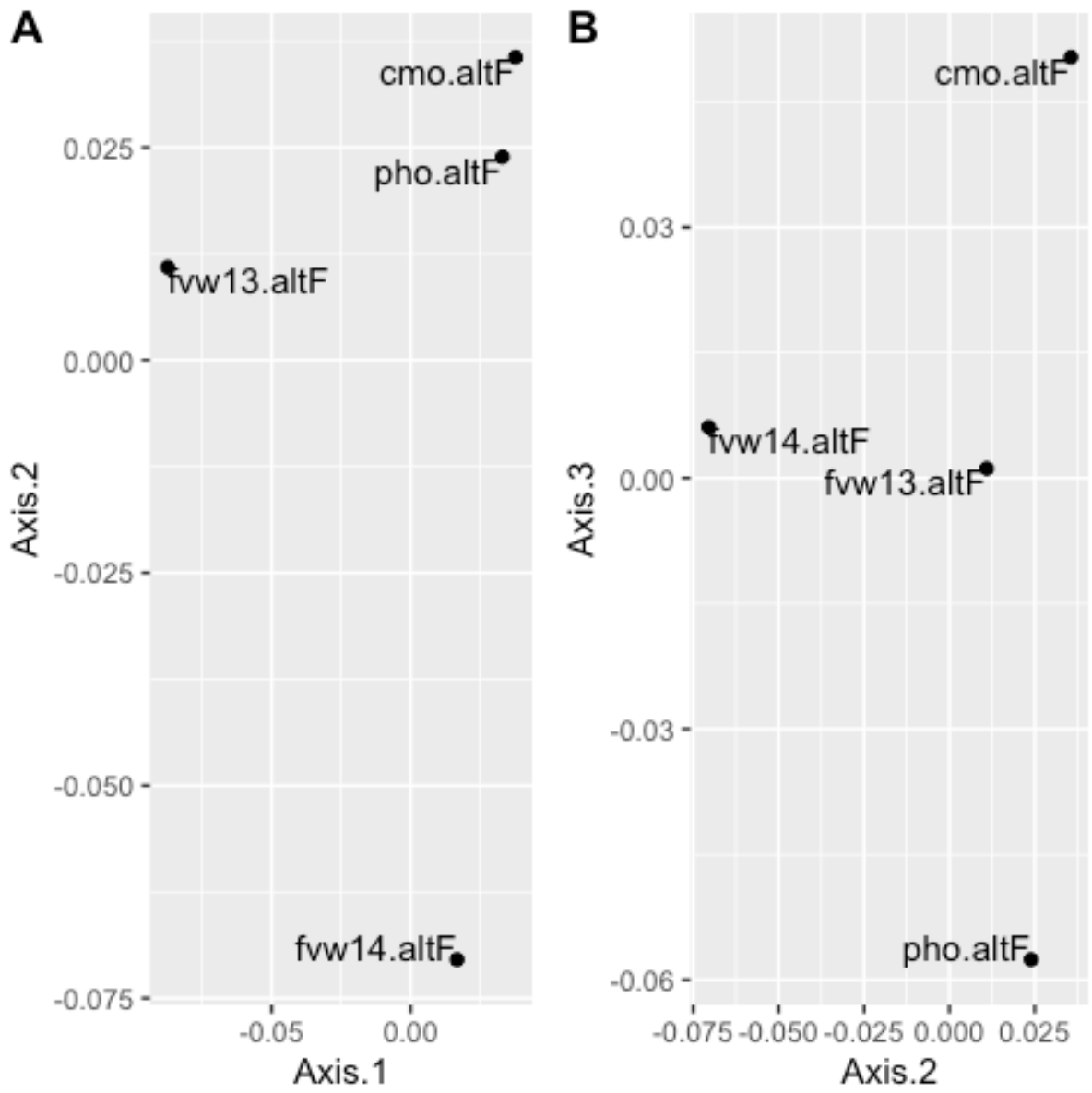
Supplemental Figure 14. Alignment of sanger sequencing of the region containing the *ds* polymorphism from several DGRP lines (line numbers indicated at line starts) to reference sequence (from *Drosophila melanogaster* genome). DGRP lines are indicated with DGRP 195, 28, 96, 48, 59 and 801 predicted to have the insertion, and, DGRP 129, 301, 69, 385, 75, 83, 491, 34 and 774 without. In addition to the insertion in these lines, several other associated SNPs are found linked to the insertion.



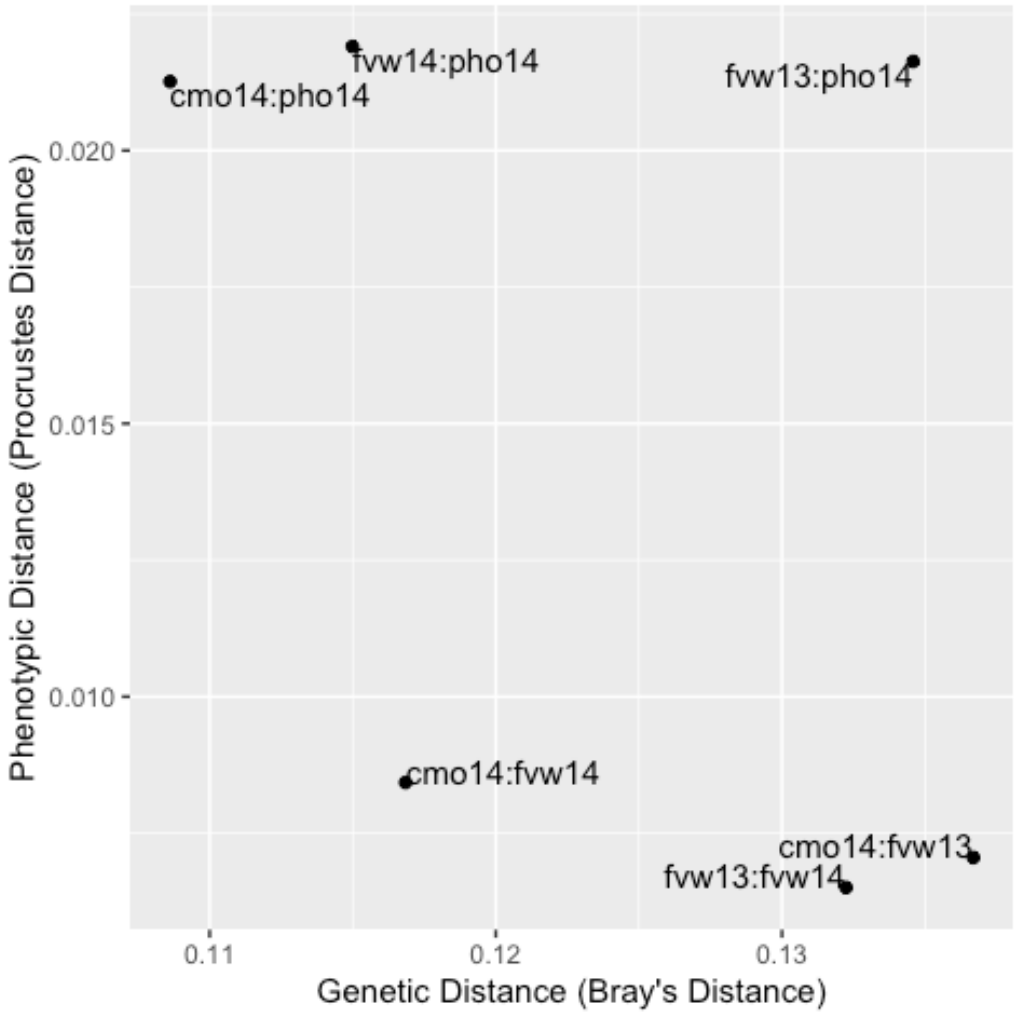
Supplemental Figure 15: Permutation test for over representation of hippo signaling terms in outlier regions. The red line represents the observed ratio of significant to expected genes in outlier windows (F_{ST} greater than 3 standard deviations from the mean). The permutation test selected random windows from the genome, equal in number to those identified as outliers. For each random draw of loci, the ratio of significant to expected genes in the term of interest was calculated for 1000 permutations.



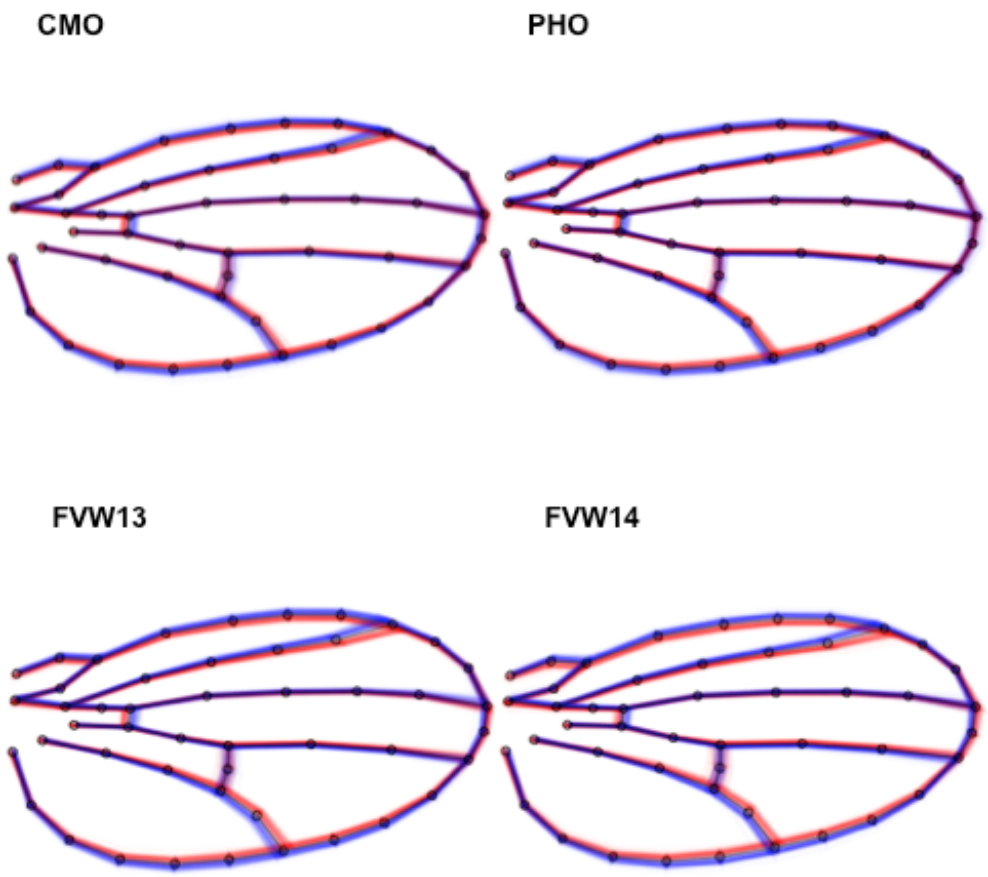
Supplemental Figure 16. Response to selection based on projections onto *emc* shape change vector. Regression of cumulative selection differential onto cumulative response was used to estimate realized heritability in both “up” and “down” selection regimes independently.



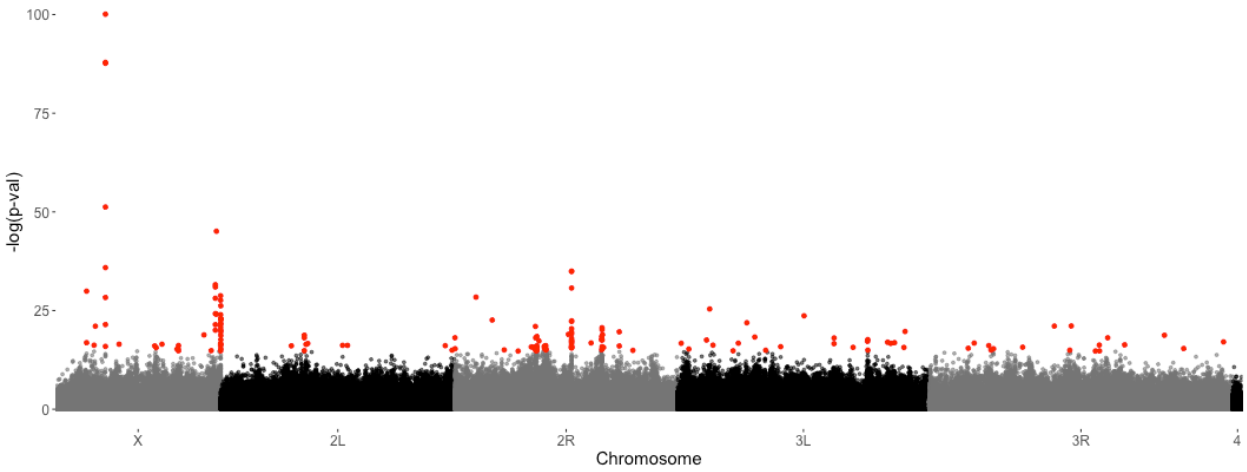
Supplemental Figure 17. Minimal genetic structure among wild populations used in this study. Principal coordinate plot calculated from Bray's distances estimated from allele frequency data (altF), between the four wild cohorts used in this analysis. Axis 1-3 explains 45%, 30% and 25% of the variance in genetic distances respectively.



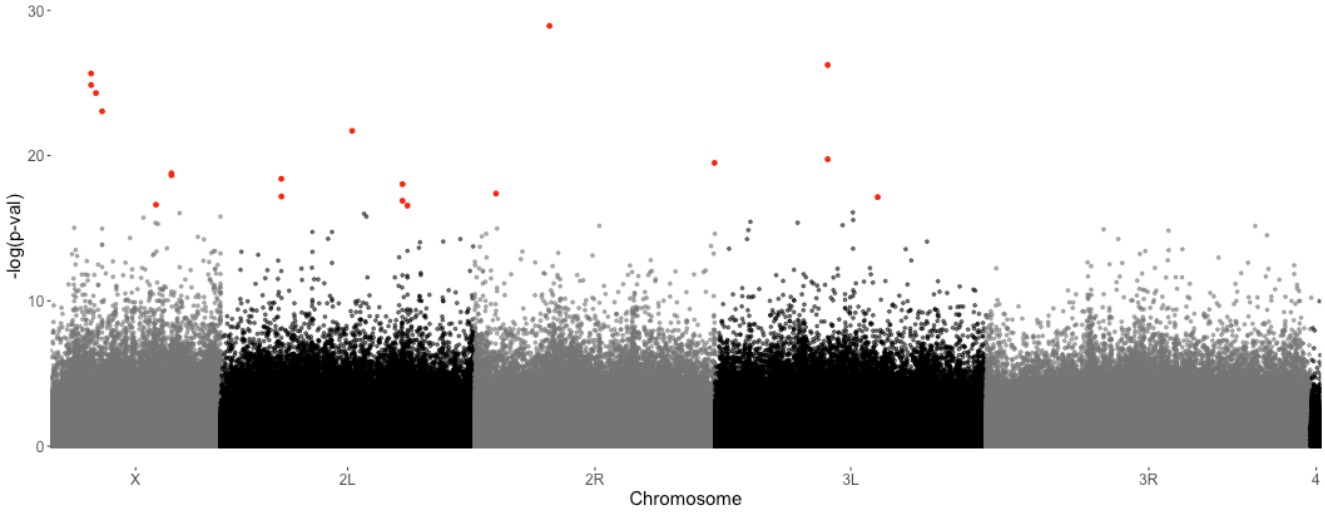
Supplemental Figure 18. Genetic distance is not correlated with phenotypic distance among wild cohorts used in this study. Phenotypic distances are the Euclidian distances from model estimated shape vectors for each population from a multivariate regression of shape onto for population and wing centroid size.



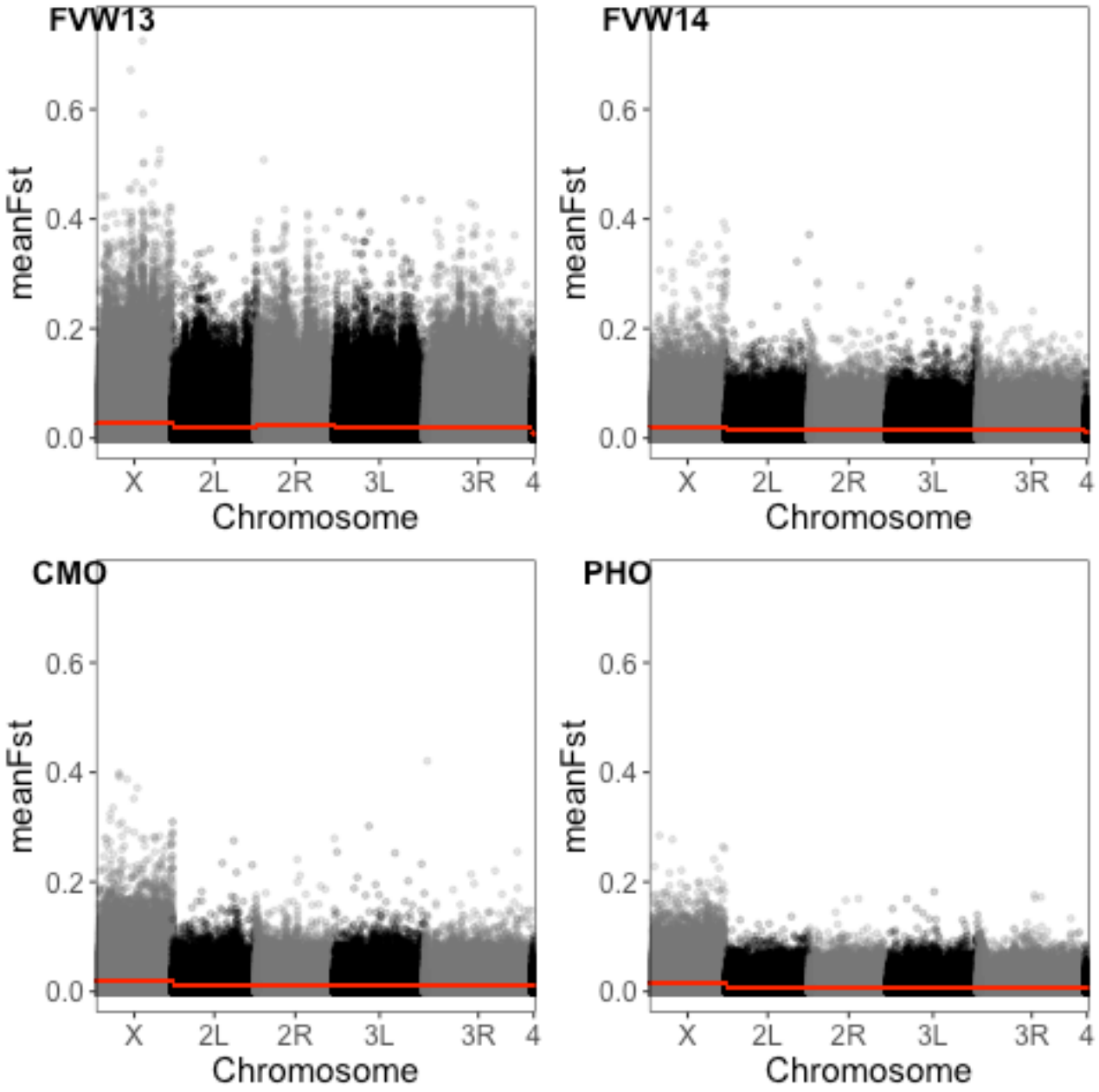
Supplemental Figure 19. Shape variation within *ds* selected pools for bulk segregant analysis by population. Wings within selected pools (one red and one blue, representing the “up” and “down” pools) are plotted to show the phenotypic extremes for the shape scores used for selecting individuals within and between pools. Black line indicates mean shape between pools.



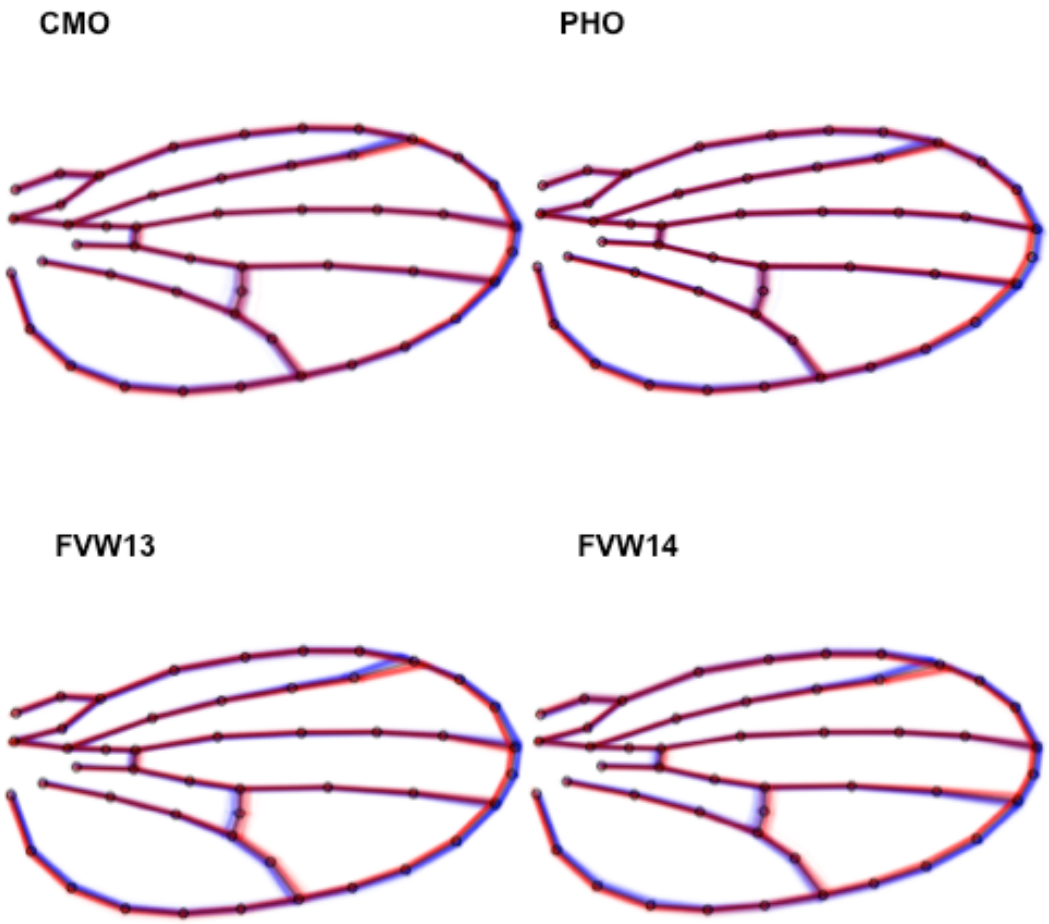
Supplemental Figure 20. Removing PHO from the BSA analysis increases the number of differentiated sites between *ds* shape change pools. Genome-wide scan for differentiated loci between pools selected based on *ds* shape change vector using the CMH test using ACER. Variants in *ds* are still not implicated in this analysis. Points in red indicate sites with significant differentiation based with a FDR of 0.05.



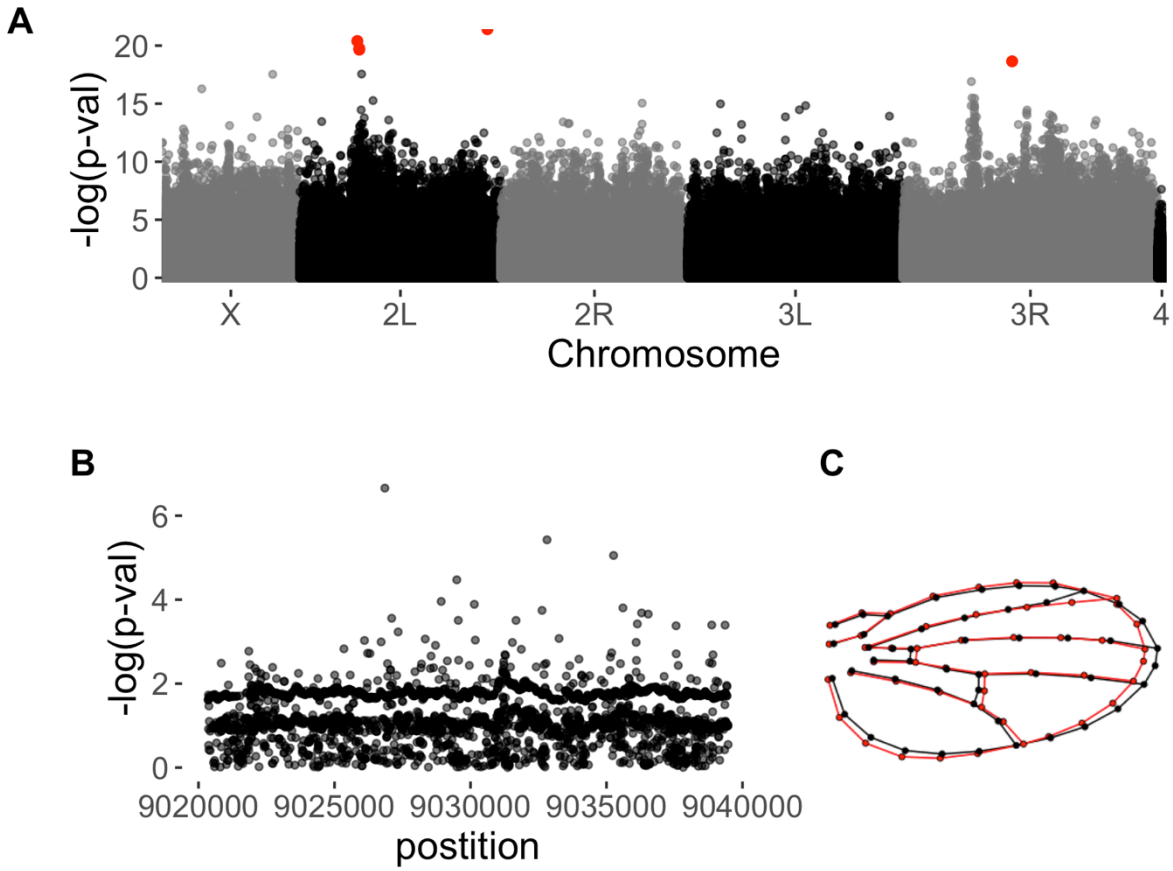
Supplemental Figure 21. Down sampling genome coverage does not impact results for CMH tests substantially. When all pools are sampled to a coverage depth of 75 reads, while preserving allele frequency, we do not find an increase in number of differentiated sites. Although the sites identified as significantly differentiated do change somewhat, in part due to sampling procedure as any sites without a minimum depth of 75 reads in each of the 4 populations was dropped from this analysis. Variants in *ds* are still not implicated in this analysis.



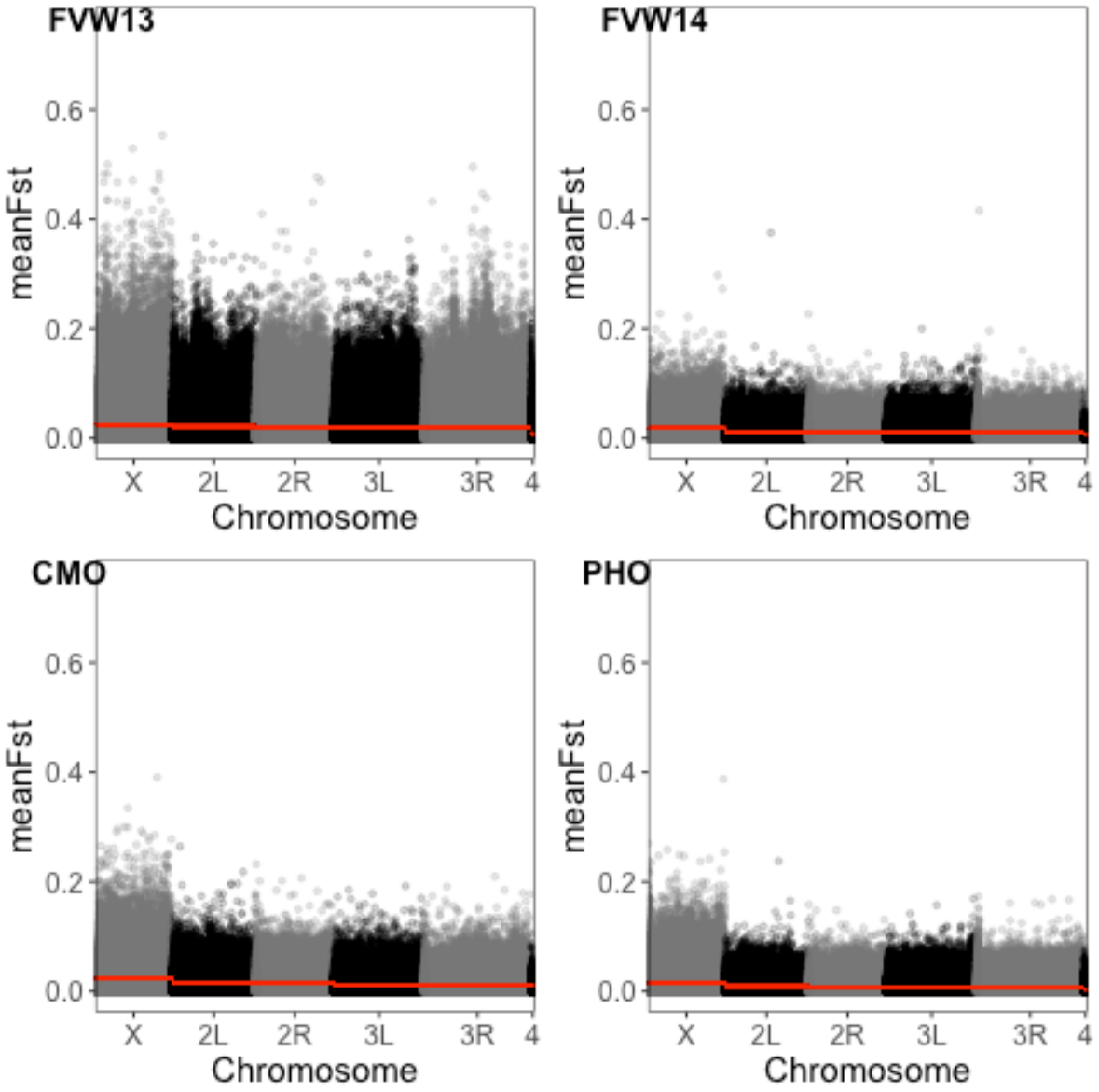
Supplemental Figure 22. F_{ST} between pools of individuals selected along the *ds* shape change axis within each population. Calculated in 100 bp windows using PoPoolation2 program. Elevated F_{ST} on the X chromosome is due to sampling of fewer X chromosomes, relative to autosomes as most pools, with the exception of FVW14, consist of only males. Red line indicates the mean F_{ST} for the chromosome, which as expected is very low.



Supplemental Figure 23. Shape variation within *neur* selected pools for bulk segregant analysis by population. Wings within selected pools (one red and one blue, representing “up” and “down” extreme pools respectively) are plotted.



Supplemental Figure 24. Genome-wide scan for differentiated loci between pools selected based on *neur* shape change vector using the CMH test implemented with ACER. (A) Whole genome scan for differentiation. Points in red indicate sites with significant differentiation based with a FDR of 0.05. (B) No significantly differentiated sites within *neur.* (C) shape change between selected pools based on *neur* shape change vector, effect size is multiplied by 2 for visualization.



Supplemental Figure 25. F_{ST} between pools of individuals selected along the *neur* shape change axis within each population. Calculated in 100 bp windows using PoPoolation2 program. Elevated F_{ST} on the X chromosome is due to sampling of fewer X chromosomes, relative to autosomes as most pools, with the exception of FVW14, consist of only males. Red line indicates the mean F_{ST} for the chromosome.

Supplemental Table 1. Number of individuals used for BSA from each wild-caught cohort

Collection	Females Phenotyped	Males Phenotyped	Females – "up" pool	Females – "down" pool	Males – "up" pool	Males – "down" pool
FVW13	0	2184	0	0	75	75
FVW14	403	797	59	0	16	75
PHO14	0	1232	0	0	75	75
CMO14	0	1001	0	0	75	75

Supplemental Table 2. Directions of major axes of variation differ somewhat between the among line (DGRP) covariance matrix and the wild caught cohorts, likely due to a combination of differences in allele frequencies and the major contribution of environmental variation (and potentially GxE) in the wild caught individuals. Correlations for the first three eigenvectors computed from among the DGRP, wild-cohorts (all populations) and CMO shape variation. |Estimates| < 10^{-15} are treated as 0.

	dgrp.PC1	dgrp.PC2	dgrp.PC3	wild.PC1	wild.PC2	wild.PC3	cmo.PC1	cmo.PC2	cmo.PC3
dgrp.PC1	1	0	0	-0.17	-0.46	0.26	-0.35	-0.54	0.066
dgrp.PC2		1	0	-0.18	-0.21	0.61	-0.10	-0.60	-0.27
dgrp.PC3			1	0.56	-0.12	0.40	-0.32	0.27	-0.71
wild.PC1				1	0	0	-0.39	0.68	-0.29
wild.PC2					1	0	0.90	0.31	-0.061
wild.PC3						1	0.056	-0.55	-0.74
cmo.PC1							1	0	0
cmo.PC2								1	0
cmo.PC3									1

Supplemental Table 3. Top 50 enriched GO terms of linked differentiated sites following artificial selection based on *ds* shape change
See: ds_artSelGOtop50.csv

Supplemental Table 4. Top 50 enriched GO terms of linked differentiated sites following artificial selection based on *emc* shape change
See: emc_artSelGOtop50_3sd.csv

Supplemental Table 5. Pairwise Procrustes (Euclidian approximation) distances between mean shapes across wild cohorts. Pairwise comparisons between population means estimated from a model with fixed effects for centroid size and population. Z-score and p-value are calculated using permutation of residuals as implemented in the RRPP package using a null model with only the effect of centroid size. Comparisons to the PHO samples, which show the largest pairwise differences are in bold.

Population Comparison	Distance	95% permutation threshold	Z-score	p
CMO-FVW13	0.0070	0.0011	9.70	0.001
CMO-FVW14	0.0084	0.0014	11.0	0.001
CMO-PHO	**0.021**	**0.0013**	**15.2**	**0.001**
FVW13-FVW14	0.0065	0.0012	7.25	0.001
FVW13-PHO	**0.022**	**0.0012**	**13.0**	**0.001**
FVW13-PHO	**0.022**	**0.0014**	**13.5**	**0.001**

Supplemental table 6. Significantly differentiated sites when PHO population is left out of CMH test.
See: ds_wild_significantSites_zeroGen_genes_fdr_noPho.csv

Supplemental table 7. GO analysis of significantly differentiated sites found in CMH test when PHO population is left out of analysis.
See: GOanalysisTop50_cmhSites_fdr_noPHO_ds_wild.csv

Supplemental Table 8. Significantly differentiated variants for *ds* shape change (CMH test) from the wild-caught cohorts (BSA) when all data is down-sampled to 75X coverage per site.

Location	CMH p-value (FDR corrected)	Gene	FlyBase ID	Distance from ORF (bp)
X:14833408	5.42×10^{-5}	*CG14411*	FBgn0030582	0
X:14833412	3.29×10^{-5}	*CG14411*	FBgn0030582	0
X:15381676	7.54×10^{-5}	*CG9164*	FBgn0030634	0
X:16037404	2.21×10^{-4}	*Muc14A*	FBgn0052580	0
X:22956575	4.58×10^{-2}	*lncRNA:CR45502*	FBgn0267058	1501
X:4118816	9.49×10^{-3}	*tyf*	FBgn0026083	0
X:4130993	9.84×10^{-3}	*GlcAT-I*	FBgn0066114	218
2L:14793860	1.15×10^{-2}	*CG18420*	FBgn0028866	656
2L:1480390	3.03×10^{-2}	*CG31928*	FBgn0051928	11905
2L:20626829	7.29×10^{-4}	*lncRNA:CR44909*	FBgn0266214	623
2L:4481772	1.54×10^{-2}	*dpy*	FBgn0053196	0
2L:4481788	3.70×10^{-2}	*dpy*	FBgn0053196	0
2L:4843588	4.62×10^{-2}	*mxt*	FBgn0031637	0
2R:11876761	2.73×10^{-2}	*CG13185*	FBgn0033661	0
2R:16405045	3.70×10^{-6}	*Sema2b*	FBgn0264273	0
2R:9943212	5.15×10^{-3}	*Mef2*	FBgn0011656	0
3L:19082031	4.49×10^{-3}	*Mkp3*	FBgn0036844	0
3L:19082032	2.73×10^{-3}	*Mkp3*	FBgn0036844	0
3L:23237904	3.05×10^{-2}	*nAChRalpha4*	FBgn0266347	0

Supplemental table 9. Significantly differentiated variants for *neur* shape change from the wild-caught cohorts (BSA).

Location	CMH p-value (FDR corrected)	Gene	FB ID	Distance from ORF (bp)
2L:15775767	0.026	*CG43760*	FBgn0264260	675
2L:15967220	0.017	*Beat-Ib*	FBgn0028645	4785
2L:15967222	0.017	*Beat-Ib*	FBgn0028645	4787
2L:8651597	0.015	*Sema1a*	FBgn0011259	0
3R: 21898160	0.024	*CG6678*	FBgn0038917	0