Title

EVOLUTION AND INHERITANCE OF EARLY EMBRYONIC PATTERNING IN DROSOPHILA SIMULANS AND D. SECHELLIA

Journal

Evolution, 65(5)

ISSN

0014-3820

Authors

Lott, Susan E
Ludwig, Michael Z
Kreitman, Martin

Publication Date

2011-05-01

DOI

10.1111/j.1558-5646.2010.01206.x

Peer reviewed

Published in final edited form as:
Evolution. 2011 May ; 65(5): 1388–1399. doi:10.1111/j.1558-5646.2010.01206.x.

Evolution and inheritance of early embryonic patterning in *D. simulans* and *D. sechellia*

Susan E. Lott[1,*], **Michael Z. Ludwig**, and **Martin Kreitman**
Committee on Genetics and the Department of Ecology and Evolution, University of Chicago, 1101 E 57th St., Chicago, IL, 60637, USA.

Abstract

Pattern formation in *Drosophila* is a widely studied example of a robust developmental system. Such robust systems pose a challenge to adaptive evolution, as they mask variation which selection may otherwise act upon. Yet we find variation in the localization of expression domains (henceforth 'stripe allometry') in the pattern formation pathway. Specifically, we characterize differences in the gap genes *giant* and *Kruppel*, and the pair-rule gene *even-skipped*, which differ between the sibling species *D. simulans* and *D. sechellia*. In a double-backcross experiment, stripe allometry is consistent with maternal inheritance of stripe positioning and multiple genetic factors, with a distinct genetic basis from embryo length. Embryos produced by F1 and F2 backcross mothers exhibit novel spatial patterns of gene expression relative to the parental species, with no measurable increase in positional variance among individuals. Buffering of novel spatial patterns in the backcross genotypes suggests that robustness need not be disrupted in order for the trait to evolve, and perhaps the system is incapable of evolving to prevent the expression of all genetic variation. This limitation, and the ability of natural selection to act on minute genetic differences that are within the "margin of error" for the buffering mechanism, indicates that developmentally buffered traits can evolve without disruption of robustness

Keywords

segmentation patterning; robustness; scaling

Introduction

Segmentation in *Drosophila* involves the expression of a spatially and temporally robust system of landmarks along the anterior-posterior (A-P) axis in the blastoderm embryo that will precisely demarcate segmental boundaries in the developing fly. This process in *D. melanogaster* has been shown to be robust to stochastic noise (He et al. 2008; Gregor et al. 2007), environmental perturbation (Lucchetta et al. 2005), and genetic variation (Namba et al. 1997; Houchmandzadeh et al. 2002), three hallmarks of a genetically robust system (de Visser et al. 2003).

Within *D. melanogaster*, spatial patterning scales with embryo length (Houchmandzadeh et al. 2002), and is robust to genetic variation in embryo size (Gregor et al. 2005; Lott et al. 2007), despite egg size variation posing a challenge for the patterning system. For example, crosses between small- and large-embryo *D. melanogaster* lines produced F3 embryos varying by approximately 25% in length, but exhibiting no measurable differences among

[*]corresponding author: slott@uchicago.edu.
[1]Present Address: Department of Molecular and Cell Biology, 387 Stanley Hall #3220, University of California, Berkeley, CA 94720, USA.

individuals in the relative positioning of gap or pair-rule gene expression patterns (Lott et al. 2007), a feature we will refer to as “stripe allometry”, suggesting either the lack or the suppression of genetic variation for this trait. Yet strong artificial selection for egg size differences in *D. melanogaster* produces allometric shifts in pattern position, suggesting that there is some genetic variation present for this trait (Miles et al. 2010). The presence of some genetic variation in this trait is important because robust traits are generally believed to be capable of building up large stores of ‘cryptic’ genetic variation (Gibson and Dworkin 2004). If released by either environmental perturbation or change in the genetic background, this variation can give rise to rapid adaptive phenotypic evolution (Le Rouzic and Carlborg 2008).

Stripe allometry has evolved over a short span of evolutionary time between the sibling species *D. simulans* and *D. sechellia* (Lott et al. 2007). *D. sechellia* is an island endemic species, recently derived from the cosmopolitan *D. simulans*, approximately 250,000–500,000 years ago (Kliman et al. 2000; McDermott and Kliman 2008). *D. sechellia* has very little polymorphism (Cariou et al. 1990; Hey and Kliman 1993; Kliman et al. 2000; Morton et al. 2004), likely indicating a small population size throughout its history (Cariou et al. 1990; Kliman and Hey 1993; Legrand et al. 2009). In a short period of time *D. sechellia* has adapted to breeding on the fruit of *Morinda citrifolia*, which is toxic to *D. simulans* (R'Kha et al. 1991; Legal et al. 1992; Moreteau et al. 1994; Amlou et al. 1998), and has also evolved a very large egg size (Lott et al. 2007), reduced ovariole number (Lachaise et al. 1986; Louis and David 1986; Coyne et al. 1991; R'Kha et al. 1991; R'Kha et al. 1997; Orgogozo et al. 2006), as well as a number of other traits (Orgogozo and Stern 2009). With respect to pattern formation, we found several scaled stripes of early *even-skipped* (*eve*) pair-rule gene expression to be displaced posteriorly in *D. sechellia* by as much as several cells relative to scaled *eve* stripe positions in *D. simulans* (Lott et al. 2007, Fig. 1).

Previous work suggested maternal inheritance of patterning allometry (Lott et al. 2007), and as the segmentation network involves many genes, perhaps the pattern differences between species would involve many factors. Does the mechanism providing robustness break down as these genomes are recombined? Such a breakdown would be detectable by the presence of transgressive stripe allometries (more extreme phenotype than either parental species), or as an increase in the variance of the patterning allometry in backcross embryos. Alternatively, the buffering mechanism might be left intact, producing one phenotype, either intermediate or like one of the parental species, but with little or no increased inter-embryo variance.

Here we investigate the evolved pattern differences in *eve*, and the gap genes *giant* (*gt*) and *Kruppel* (*Kr*) in hybrids of *D. simulans* and *D. sechellia*. Hybrid male sterility in the cross between *D. sechellia* and *D. simulans* necessitated the use of a double-backcross genetic design to investigate the basis of the stripe allometry difference. We found that stripe allometry follows a maternal pattern of inheritance, necessitating the production of F3 backcross embryos to analyze properties of pattern formation in embryos with varying genomic contributions of the two parental species. We also analyzed the segregation of 27 genetic markers in these crosses to investigate the genetic architecture of pattern formation and embryo length. We discuss evidence for independent inheritance of pattern formation and embryo length, and its implications for models of scaling and precision in embryonic patterning. Our findings lead us to question whether robust developmental systems are capable of completely suppressing additive genetic variation and suggest they may not be impediments to phenotypic evolution.

Materials and Methods

CROSSES

Virgin females from a *D. simulans* (Florida City) line were mated to males of a *D. sechellia* (Robertson) line to produce a F1 generation, which has fertile females and sterile males. Virgin F1 females were then backcrossed in both directions (to both *D. simulans* and *D. sechellia*) to produce two classes of F2s. Virgin F2 females were again backcrossed to the same parental line to produce two different F3s lines, referred to as F3bk *sim* and F3bk *sech* (indicating the direction of the double-backcross). A diagram of the crossing scheme is presented in Figure 2. 1 to 5 hour old embryos were collected from mated 5–10 day old females at each stage in the cross. Embryos were collected from cages containing thousands of females and males, so it is unlikely that multiple embryos were collected from a single mother.

VISUALIZATION AND MEASUREMENT OF EVE, GT, AND KR EXPRESSION

Embryos were collected, *in situ* hybridization performed, and nuclei visualized as previously described (Lott et al. 2007). Embryos were selected by stage and orientation to identify >60 mitotic cycle 14 embryos for analysis for each line and cross (~120 embryos for each of the two F3s). Each selected embryo was assigned to one of five age classes within mitotic cycle 14 as previously described (Lott et al. 2007; based on substages 4–8 on the Fly-Ex web Site, http://flyex.ams.sunysb.edu/flyex; Myasnikova et al. 2001). Developmental age was also estimated using a measurement of the extent of cellularization; this measure of age performed almost as well, but results reported here use the pattern based staging method as it explained slightly more of the variance, though results obtained using either staging method were similar.

Optical Z-sectioning (0.8-μm per step) was carried out with an Axioplan2 microscope (Zeiss, Thornwood, NY). To separate and prevent overlap of the DAPI, Alexa 546, and Alexa 647 signals, we used the following filters from Chroma Technology (Rockingham, VT): 31000v2 for DAPI, 41002b for Alexa 546, and 31023 for Alexa 647. Photographs were taken with an Orca C4742–95 camera (Hamamatsu, Hamamatsu City, Japan) and Openlab software, version 3.1.7. The measurements of embryo length and stripe position were performed in Openlab as previously described (Lott et al. 2007).

STRIPE AND LENGTH ANALYSIS

A linear model ANOVA was used to account for the differences in stripe position that are due to the normal progression and refinement of these expression patterns during development. The ANOVA was fitted with Proc GLM in SAS (version 8.2) with the main effect of line, and age (the inferred age of development, see Lott et al. 2007) was used as a covariate. The significance was estimated with restricted maximum likelihood, and the denominator degrees of freedom were determined with the Satterthwaite method. Least-square means were calculated with the Lsmeans option:

$$Y=\mu+\text{line}+\text{age}+\text{line}\times\text{age}+\varepsilon$$

To provide evidence for the maternal inheritance of pattern allometry, pairwise t-tests of means were conducted using the residuals from the linear model.

TESTING CONFOUNDING EFFECTS OF BETWEEN-SPECIES CROSS

Markers based on insertion/deletion (indel) differences between *D. simulans* and *D. sechellia* were created using available sequence alignments

(http://www.biostat.wisc.edu/~cdewey/fly_CAF1/). Five indels of 50–100 bp were identified approximately 1/5 of the way along each major chromosome arm (plus one for chromosome 4, and a primer to a region of the male fertility factor kl5 on the Y chromosome for sex determination), for a total of 27 markers across the genome. PCR primers spanning the deletions were designed using Primer 3 (Rozen and Skaletsky 2000), and verified in each species and the F1 hybrids. Table S4 lists primers that were used, genomic position in *D. simulans*, and the size and direction of the indel.

DNA from the same individual embryos that had been imaged was extracted using a modification of the methods of Ludwig et al. (2005) combined with whole-genome amplification using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare, Piscataway, NJ). Protocol is available upon request.

The indel marker products were designed to be varying lengths between 1 kb and 100 bp so that 5 markers could be run at one time; six PCR reactions were performed to cover the genome (1 each for X, 2L, 2R, 3L, 3R, and 1 for Y and 4). Reactions were carried out using the Qiagen Multiplex PCR kit (Qiagen, Inc., Valencia, CA) following standard protocol, and run on 1.7% agarose gels to separate amplified fragments. Genotypes were called based on visual analysis.

Genotyping analysis was performed in R (R Development Core Team, 2006). To determine if the distributions of genotypes at single loci in the F3s were significantly different than Mendelian expectations, the expected and observed counts for each genotype were compared with a chi-squared test. A few loci were identified as deviating significantly from expectations under simple Mendelian segregation (Figure 5). In order to examine if these deviations from Mendelian expectations had an effect on length or expression pattern position, phenotypic means (corrected for the effect of age) were calculated for each genotype at each marker, and multiplied by the frequencies expected for simple Mendelian inheritance of single loci. These phenotypic means with the “rebalanced” Mendelian frequencies were compared to the observed mean phenotype (see Figure S2).

QTL ANALYSIS

Single markers were tested for associations between genotype and phenotype. Because of the maternal phenotype- zygotic genotype offset (gene expression positions are strongly predicted by the maternal genotype, whereas the genotype is that of the embryo, see discussion), interval mapping was not possible with available programs. As we are required to use single marker tests, there may also be some reduction in power compared to multi-marker tests, but this drop is unlikely to be appreciable (Coffman et al. 2003). Single-marker analysis was performed in R.

Two fixed-effects ANOVAs were used to determine the relationship of genotype and phenotype. The first was used to investigate effects of genotype in each F3bc separately:

$$Y=\mu+\text{age}+\text{marker}+\text{age}\times\text{marker}+\varepsilon.$$

The phenotype “Y” is either embryo length or the position of an individual stripe boundary, in either the F3bc *sim* or F3bc *sech* background; μ is the mean for all individuals for this phenotype; “age” is an inferred developmental age class (described above); “marker” is the deviation given by the genotype at a given marker; and ε is the unaccounted deviation of the measured phenotype from the predicted value. This procedure is effective in removing measurement variability due to the dynamic shift of gene expression boundary positions in an embryo as it ages, thus allowing a better assessment of the variance in the phenotype that

is due to marker genotype alone (Lott et al. 2007). The results from this model are represented in Figure 6 A–C and Figure S4 A–C, and ANOVA tables for length, absolute stripe position, and relative stripe position can be found in Tables S5, S7, and S9.

The other model used to examine the association of a genotype with the phenotype of interest pools F3bc *sim* and F3bc *sech*:

$$Y=\mu+\text{age}+\text{BC line}+\text{marker}+\text{age}\times\text{BC line}+\text{age}\times\text{marker}+\text{BC line}\times\text{marker}+\text{age}\times\text{BC line}\times\text{marker}+\varepsilon$$

The variables “Y”, “μ“, “age”, “marker”, and “ε” are the same as in the previous model. To this model we added an additional term: “BC line”, which is an indicator variable for the background line of the genotype, F3bc *sim* or F3bc *sech*. Using this model, we can find QTL that have effects in either or both the F3s, providing additional power when QTL have the effects in both F3 backgrounds. A portion of the results from this model are represented in Figure 6 D–F and Figure S3, and ANOVA tables for length, absolute stripe position, and relative stripe position can be found in Tables S6, S8, and S10.

To determine thresholds for identifying significant associations between any of the markers and phenotype and address multiple testing concerns, we established null probability distributions for association by permuting the data. Specifically, for embryo length we permuted the length phenotype across individual embryos within a cross and kept a record of the most significant p-value for any marker for each permutation. We then calculated the 5% cutoff from this empirical null distribution of the minimum p-value. This threshold was very similar in F3bc *sim* and F3bc *sech*, so one line is drawn on the plots in Figure 6 and Figure S4. Similarly, for each gene expression stripe position, we performed permutations individually by stripe to produce an empirical threshold of significant association. The p-value cutoffs were similar both across stripes and between the two backcross lines, so again one significance threshold is presented in Figure 6 and Figure S4. We do not present p-values adjusted for multiple testing of the 14 *eve* stripe borders for two reasons. First, stripe boundary positions are strongly correlated (Lott et al. 2007), indicating that shifts in position are not independently controlled. Second, the markers that have significant association beyond the threshold (uncorrected for multiple testing) are significant for nearly all stripes (see Tables S11 and S12), thus reinforcing the evidence for association.

Power calculations and proportioning of variance were performed according to standard methods (Lynch and Walsh 1997). All proportions of variance attributed to marker genotypes listed (tables S5–S10) refer to the proportion of variance explained after the removal of the effect of developmental age on stripe position. While the effect of developmental age varied over stripe and somewhat by the type of model and measurement tested (Table S13), on average, developmental age accounts for ~30% of the variance in stripe positioning.

Results

DIFFERENCES IN PATTERN ALLOMETRY BETWEEN SPECIES

Expression boundaries of the pair-rule gene *even-skipped (eve)*, and the gap genes giant *(gt)* and *Kruppel* (*Kr*) were visualized in *D. simulans* and *D. sechellia* embryos through *in situ* hybridization, and boundary positions measured as a percentage of embryo length (%EL). We confirm, as previously reported (Lott et al. 2007), a difference in pattern localization of all of these factors (Figure 1, Figure 3B, Figure S1). As illustrated for *eve* in Figure 1 and Figure 3B, difference in placement of the seven *eve* stripes is most evident for the middle stripes, which are placed more posteriorly in *D. sechellia* than in *D. simulans*.

INHERITANCE OF PATTERN ALLOMETRY

To conduct a genetic analysis, the two species were crossed and then backcrossed twice, to produce two distinct F3 classes (F3bc *sim* and F3bc *sech*), as diagrammed in Figure 2. Embryo length and expression boundary positions were measured (Figure 3A). Mean stripe boundary positions for each stage in the cross were estimated using a linear model, which by including developmental age as a covariate allowed us to account for its effect on patterning dynamics (see Materials and Methods). Positions of *eve* stripes are portrayed in Figure 3B as deviations from the mean values for all the lines, the similar results for *gt* and *Kr* are represented in Figure S1. F1 stripe placement is indistinguishable from maternal parent, *D. simulans* (no p-value <0.05 for any stripe, pairwise t-test, accounting for developmental age, see Table S1). This is evidence that these zygotically expressed factors (*eve*, *gt*, *Kr*) are expressed at the same scaled positions along the A-P axis as in the maternal line, consistent with maternal inheritance, as previously hypothesized (Lott et al. 2007). Mean stripe positions in the two F2 backcross classes are intermediate to *D. simulans* and *D. sechellia*, and are likewise indistinguishable from one another (no p-value <0.05 for any stripe, pairwise t-test, accounting for developmental age, see Table S1), as would be expected under maternal inheritance and a genetic model with additivity, given the heterozygous genotypes of their F1 mothers. In the two F3 backcrosses (to *D. simulans* or to *D. sechellia*), mean stripe positions shift towards the backcross species, although not equally so in the two F3 classes. *gt* and *Kr* (Figure S1) expression domains show similar shifts, with maternal inheritance and deviation between the F3 backcross classes.

PATTERN ALLOMETRY AND EMBRYO LENGTH

Genetic analysis of expression patterns in F3 backcross classes shows that stripe localization (Figure 3B) is genetically separable from embryo length (Figure 3C), and gives rise to changes in the allometry of the segmentation gene patterning. F3bc *sim* embryos (orange in Figures 3B and 3C) have a *D. simulans*-like mean length, but intermediate, F2-like relative stripe positions. In contrast, F3bc *sech* (purple in Figures 3B and 3C) have a *D. sechellia*-like stripe position, but an intermediate, F2-like length. The difference in stripe allometry between the species, therefore, does not involve co-segregation of factors controlling both scaling and embryo size. To the contrary, the data provide unambiguous evidence for the independence of factors responsible for size and stripe allometry differences between the species.

VARIANCE IN PATTERN ALLOMETRY INHERITANCE

As expected, F2 embryos do not exhibit greater inter-individual variability in stripe positions than the parental lines, since all of their mothers were genetically identical F1 hybrids. F3 embryos, in contrast, were produced by genetically variable F2 mothers differing in their combinations of *D. simulans* and *D. sechellia* alleles. These embryos should vary with respect to the parental origin of stripe allometry alleles, and as a consequence should exhibit a commensurate increase in the variance in stripe positioning compared to F2 or F1 embryos. This expected increase in genetic variance will be largest if a small number of "large effect" alleles are responsible for patterning differences between the species. The buffering system can also come into play, either increasing the variance in patterning if combinations of inter-species alleles in F2 mothers are dysgenic for developmental buffering (*i.e.*, produce transgressive offspring), or decreasing the variance by suppressing this segregating genetic variation.

No increase in the variance in stripe positioning was observed within either of the two F3 backcrosses, despite a clear difference in the average stripe allometry of the two F3 classes. Bootstrap resampling of embryos within a line confirmed that the F3 variances are not larger than earlier stages in the cross (see Table S2). Thus, despite the shifts in their mean positions

to novel locations, stripes are positioned in F3 embryos with no loss of accuracy. This accuracy in stripe position is constant along the range of embryo sizes within a F3 backcross line, as can be seen in Figure 4, which plots relative *eve* stripe boundary positions as a function of individual egg length. The distributions of F3 egg lengths differ for the two backcrosses, reflecting egg size differences between the two species, and the mean position of relative stripe boundaries also differ (Figure 4, green triangles), reflecting the novel scaling allometries produced by these crosses. Yet within each cross, *eve* stripes are placed in the same relative positions across individuals. The slopes for the regression of stripe positions against embryo length (from a linear model, with age) are not significantly different from zero ($p<0.05$), and can be found in Table S3. The lack of a measurable variance increase within an F3 backcross class either relative to the previous stages in the cross or embryo length suggests the contribution of a large number of loci of small effect to the allometry difference between the two species, and/or genotype-independent maintenance of developmental buffering in this phenotype.

RULING OUT CONFOUNDING EFFECTS IN A BETWEEN-SPECIES CROSS

Given the early developmental stage and functional importance of phenotypes that differ between these species, a lack of increased phenotypic variance within the F3 backcrosses could occur if extreme phenotypes are especially unfit (embryos die earlier in development than we are sampling) and were, as a consequence, underrepresented in our measured samples of embryos. To test this hypothesis, we developed a method to genotype individual embryos at a number of markers after they were stained and imaged (see Materials and Methods, also Ludwig et al. 2005). Twenty-seven genotyping markers (five per chromosome arm for the X chromosome plus autosomes, a Y and a chromosome 4 marker, see Table S4) were developed to distinguish between *D. simulans* and *D. sechellia* chromosome segments in each individual F3 embryo measured in our study. From the analysis of segregation patterns at these 27 loci, we found that all genotypes (additionally all pairwise combinations of genotypes) were present in our sample, and there was no sex bias. We did observe, however, some significant deviation from Mendelian segregation (Figure 5), where heterozygotes were overrepresented in the F3bc *sech* line relative to *D. sechellia* homozygotes, for most of the X chromosome and for a part of chromosome 2L. To investigate the magnitude of the effect of the observed deviation from Mendelian segregation on embryo length or stripe position (Figure 5), we calculated the mean length or stripe position for each genotype at each marker (Figure S2). From this we calculated expected length and stripe position that would be observed if there was perfect Mendelian segregation at each locus. The expected lengths and stripe positions were not significantly different from the observed lengths and stripe positions (Figure S2). Therefore, for both of the backcross F3 classes, we can rule out the possibility of natural selection eliminating certain genotypic combinations as the cause of their invariant stripe allometries.

GENETIC MAPPING OF STRIPE ALLOMETRY AND EMBRYO LENGTH

Genotypic and phenotypic measurements for embryos in our genetic crosses allowed us to conduct a low-resolution QTL analysis of pattern allometry and embryo length differences between the two species. With maternal inheritance of these two traits, we would ideally want the genotype of a mother and the phenotypes of her embryos for this analysis. However, genotyping the mothers was not feasible, as this would require collecting individual eggs from single F2 females, aging them to the correct developmental stage, and fixing these individual F3 embryos, in large enough samples of mothers and eggs for statistical significance.

Instead, we examined whether genotyping F3 embryos and mapping associations of zygotic genotype to zygotic phenotype could identify genetic determinants of stripe allometry and

embryo length despite the genotype-phenotype offset. In our cross, a heterozygous F3 embryo must have a heterozygous mother, and a homozygous F3 embryo can have either a heterozygous or homozygous mother with equal probability. This will make the mean stripe position (or embryo length) of the homozyogous zygote appear to be closer to that of the heterozygote, thus decreasing statistical power to detect differences between the mean. This makes the genetic mapping experiment pursued here conservative, but does not increase the probability of false positive associations. Since the maternal phenotype – zygotic genotype offset violates assumptions in genetic models used for interval mapping, this prevented us from applying widely used programs for interval mapping. Single-marker association tests are a reasonable alternative, as there is no underlying genetic model of inheritance (Coffman et al. 2003). Two statistical models were used to examine genotype-phenotype associations. Both included the inferred developmental age of an embryo as a covariate (see Materials and Methods), as this has a significant effect on stripe position. One model examines the effect of genotype in both F3bc lines separately (F3s separate model), and the other pools the F3bc lines to examine effect of genotype in both genetic backgrounds (F3s pooled model).

Despite the maternal-zygotic offset, power calculations (using a conservative Bonferroni corrected $p = 0.05$ significance threshold) predict that after accounting for the effect of developmental age on stripe position, we will be able to identify effects explaining upward of ~8% of the variance in length and stripe position when the two F3bc classes are analyzed separately (F3-separate model), and those that explain upward of ~4% when they are pooled (F3-pooled model). Simulations of a contributing genetic factor including the genotype-phenotype offset show that these thresholds (8% and 4%) correspond to ~20–25% and 15%, respectively, of variance explained by the genetic factor were there not an offset between genotype and phenotype. Losses in percent of variance explained due to incomplete marker coverage were minimal compared to the loss due to the genotype-phenotype offset. Primarily, this indicates that any associations we are able to detect will be genetic factors with moderately large contributions to the phenotype.

With this caveat in mind, we restricted our attention to answering the following two questions: (1) Can length and patterning differences be explained by a single major-effect QTL, or are there likely to be multiple loci involved in the traits? Results presented in the previous section found no evidence for an increase in the variance in F3 phenotypes relative to F2 or parental generation, suggesting a genetic model with a large number of loci of small effect on stripe allometry. (2) Do the chromosomal regions containing a QTL for pattern allometry encompass known maternal segmentation determinants? The maternal basis of these phenotypes, and similar effects on both gap and a pair-rule gene raises the possibility that maternal segmentation factor(s) may underlie patterning allometry. Evolved differences in the expression profile of any of these transcription factors could plausibly lead to a shift in the allometry of downstream gap or pair-rule expression.

Is there a single major-effect QTL? Significant associations (and general patterns) emerge from mapping the association of genotype with embryo length, absolute stripe position (uncorrected for embryo length), and relative stripe position (corrected for embryo length), (Figure 6, Figure S4). Generally, QTL for embryo length are also significant for absolute stripe position, but not for relative stripe position, indicating that stripe position scales to embryo length. Those QTL that are present in absolute stripe position, but not embryo length, are the ones that we identify as significant in relative stripe position. QTL that affect one stripe boundary affect most of the other boundaries, both across patterns (*eve*, *gt*, *Kr*) and along the length of the embryo (See Tables S11 and S12). There are two QTL of interest for embryo length (one at the centromeric portion of 2L, the other, which does not reach significance, in the middle of chromosome 3), and two for relative stripe position (2L, closer to the telomere than the length QTL on this arm, and the X chromosome, primarily in F3bc

sim). Figure S3 shows a representation of the mean position by genotype for each phenotype, and ANOVA tables for each model and phenotype, with proportions of variance attributable to each marker, can be found in Tables S5–S10.

Of the significant associations in the F3s pooled model, the QTL on 2L (nearer the centromere) for length and absolute stripe position and the QTL on the X and 2L (nearer the telomere) for relative stripe position each explain roughly 5 % of the segregating variance in these traits. From simulations, we estimate this 5% to be ~18% of the variance if we were able to directly measure the maternal genotype. For both length and relative stripe position, there is a significant proportion of variance remaining to be accounted for, so length is likely to have more than the one genetic determinant we found, and relative stripe position is probably attributable to more than the two loci identified here. Indeed, each locus uncovered here may also contain more than one gene that contributes to the phenotype. So, despite our reduced power in this study, we were able to identify causal loci for both length and stripe allometry between these two species, and provide support for a multigenic basis for both of these traits.

Do maternal patterning factors overlap with QTL? The genomic regions identified as having a significant effect on a phenotype contain few maternal or gap genes, even though they entail large segments of chromosome arms. The difference in relative stripe position between the species is not due solely to maternal expression differences at the *bicoid (bcd)* or *hunchback (hb)* loci, as these genes are located in the proximal region of 3R, and are not in one of the regions where there is evidence for an association with relative stripe position. Also, *bcd* mRNA localization factors, such as *exuperantia*, *swallow*, and *staufen*, can be similarly discounted as the major causes of relative stripe position differences between species, as the locations of these factors in the genome differ from that of the QTL identified here. The posterior maternal genes *caudal* and *nanos* (on 2L close to the centromeric end of the arm, and 3R close to the telomeric end of the arm, respectively) are also not in regions implicated in being involved in these phenotypes. Additionally, while we can probably rule out the genes we measured - *eve*, *gt*, and *Kr* - as being responsible for the stripe differences between species because they are zygotically (not maternally) activated, we can add additional evidence from the finding that none of these genes overlap with genomic regions with significant effects on the phenotype. Even *gt*, which is on the X chromosome (one of the regions showing a significant effect on stripe position in F3bc *sim*), is on the opposite end of the chromosome from the region of significant effect. This does not rule out the involvement of these factors in the difference in stripe allometry between species, but does indicate that these factors are not the major causes of these differences.

Discussion

STRIPE ALLOMETRY IS MATERNALLY DETERMINED

Stripe allometry is determined by the genotype of the mother, even though the genes measured in this study are zygotically expressed. The F1s have the same phenotypes as their *D. simulans* mothers, and both F2 backcross lines (who share the same F2 mothers, but have different fathers) have the same intermediate phenotypes. Interestingly, this implies that the zygotic genome, and therefore the *cis* aspect of regulation of these genes, is not responsible for the differences in patterning positioning or embryo length between species. Could *a trans*-acting transcription factor instead be responsible?

For the pattern allometry phenotype, the maternal morphogen Bcd is an obvious candidate as a factor responsible for the differences in stripe position between species. But several lines of evidence make this unlikely. When *bcd* gene dosage is experimentally manipulated, all of the *eve* stripes undergo a coordinate shift in positions (Driever and Nusslein-Volhard

1988; Busturia and Lawrence 1994). Additionally, in embryos artificially selected for embryo size, *eve* stripes were found to shift allometrically, with the most significant differences being found in the posterior of the embryo (Miles et al. 2010). In contrast, patterning differences we observe between *D. simulans* and *D. sechellia* are not uniformly different across the entire embryo, but rather are greatest in the middle of the embryo and smallest at either the anterior- and posterior-most stripes (Fig. 2B). We also noted in a previous study that deviations in neighboring stripe boundary positions in individual embryos are strongly correlated, with correlations decreasing with increasing physical distance between stripes (Lott et al. 2007); this is also true between species. Both of these factors point to local spatial control of patterning differences, whereas changes in *bcd* gene dosage produce an effect on the entire embryo. Additionally, we now observe the lack of overlap of the *bcd* locus with the genomic regions associated with differences in stripe allometry, as well as with Bcd localization factors or other maternal factors. We take these findings as additional evidence against differences in stripe positioning being determined by maternal factors in the segmentation pathway. This does not, however, rule out other factors that may be acting on Bcd in *trans* that we did not test. Also, this exclusion does not mean that these factors have no effect on the evolution of stripe allometry between the two species – merely that they do not have discernibly large effects on stripe position differences that we can detect with our limited methods. We do, however, detect loci with significant effects on stripe allometry between species, which would indicate that even if our power is too low to detect effects of *bcd* or related factors, there are other factors that contribute more to this phenotype, and are therefore worthy of examination.

RELATIONSHIP BETWEEN STRIPE POSITION AND EMBRYO LENGTH

In the genetic analysis of stripe position and embryo length, we find that the two traits are genetically separable. Yet for the embryos of a given cross, stripe position scales with embryo length. That length QTL also appear as regions with a significant effect on absolute – but not relative – stripe position, is an additional line of evidence that relative stripe position is buffered against genetic variation in embryo length. In every case, correcting stripe measurements for embryo length removes significant associations at the length QTL, which means that the major genetic determinants for length differences do not affect relative stripe position. Was this not the case (*i.e.*, failure in buffering to embryo length), the genomic regions associated with length would appear to affect relative stripe position. This is consistent with stripe patterning being buffered to length within each of the two F3 backcross lines (F3bc *sim* and F3bc *sech*). While F3bc *sim* and F3bc *sech* differ in length and have their stripes in different places relative to one other, the difference in stripe position is not due to genetic differences in length. Each line has relative stripe positions that are independent of genetic variation in length within that line.

PATTERN ALLOMETRY VARIANCE AND EVOLUTION

Despite the changes in pattern allometry in our crosses, we are unable to measure any increase in variation in pattern allometry in the F3 embryos, either relative to earlier stages in the cross (where maternal genotypes are invariant), or, as previously discussed, relative to embryo length. We have been unable to disrupt the much-investigated precision of the segmentation system, while we have changed the scaling of expression boundaries to embryo length. The genetic underpinning(s) of the precision of stripe placement is unknown, but it is clearly genetically separable from the scaling property of the system. One possibility is that the segmentation network itself assures precision and scaling, either directly, through evolved features such as cross-regulation of transcription factors (Manu et al. 2009; Vakulenko et al. 2009), or indirectly, as an ‘emergent’ property of the network (Bergman and Siegal 2003). The specific spatio-temporal location of segmentation gene expression, in contrast, may be influenced by fine-scale features of cellular architecture, which differ

between the species. In this way, precision and scaling could be maintained even as natural selection or genetic drift acts to shift the specific locations of patterning gene expression.

EVOLUTION OF ROBUST TRAITS

The robustness of traits is thought to be a barrier to phenotypic evolution, as robustness is thought to mask the kind of variation upon which selection may act. A related prediction from this definition of robustness is that in order to evolve, the mechanisms conferring robustness must be disrupted or exceeded, resulting in a release of previously cryptic variation, which selection may then act upon (for a review, see de Visser et al. 2003; Flatt 2005; Felix and Wagner 2008). In the example presented here, novel phenotypes are generated in hybrids in a robust system without any increase in phenotypic variance. It is unclear whether the lack of increase in variance in novel stripe allometries documented here is due to undisrupted buffering, or to the segregation of a large number of loci of small effect which determine this trait. We are able to identify a few loci, for both embryo length and stripe position, with moderately large effects (up to ~20% of the variance), which can be viewed as evidence against the latter possibility. In either case, the explosion of cryptic variation hypothesized to be necessary in order for robust traits to evolve is not observed. In this way, robust traits can evolve without disruption of robustness or the mechanism(s) that confer it.

It might be argued that we've only explored a limited region of the multidimensional genotypic space of mutations that could expose cryptic variation. Our crosses were not intended to address this broader question, but rather to explore the special region of this space that includes intermediate genotypes, likely to be similar to ones that gave rise to the evolved differences between the species. These intermediate genotypes did not exhibit any transgressive increase in the phenotypic variance, indicating that the patterning trait may have evolved between the species without disrupting the mechanism(s) conferring robustness. Changes in the particular positioning of the stripe boundary positions between species are moreover consistent with the classical genetic model of a complex trait: many alleles with small individual effects on phenotype. We are unable, however, to exclude other possible mechanisms releasing variance when pattern allometry differences evolved between these species. Yet, even if we have failed to model the historical intermediate phenotypes in the evolution of these species, we have still observed novelty in this robust phenotype without evidence for the release of cryptic variation in an artificial laboratory condition.

If this study does tentatively rule out the necessity of a release of variance to create novel phenotypes, how then might robust traits evolve? Perhaps the variability-reducing mechanisms in this system are not sufficient to exclude the expression of additive genetic variability. In this case, evolution in these traits can occur from selection on polymorphism of small phenotypic effect (Weber 1992). The first example of this kind of variation for stripe allometry segregating within a species was recently reported by Miles, *et al.* (Miles et al. 2010), where lines of *D. melanogaster* were artificially selected for egg size, and found to differ in their *eve* stripe positions.

The specific genetic changes that are responsible for patterning difference between *D. simulans* and *D. sechellia* have yet to be identified, but small changes of this nature are unlikely to be highly deleterious, in contrast to revealing previously unexpressed genetic variation due to a breakdown of developmental buffering (Mieiklejohn and Hartl 2002). The availability of genetic variation that does not damage the robustness of a phenotype, together with the efficacy of natural selection, we suggest, can produce rapid shifts in the precise phenotypes of developmentally buffered traits, and can account for the diversity observed in many developmental phenotypes.

Of course, what remains unexplained is the remarkable functional conservation of the segmentation gene network itself, including preservation of individual stripe enhancer accuracy despite extensive sequence divergence and hundreds of millions of years of evolution (Ludwig et al. 2005; Hare et al. 2008). Perhaps the hallmark of robust developmental systems is not in their ability to completely suppress genetic variation or to prevent rapid or short-term evolution, but rather in their ability to maintain the system within strongly confined boundaries.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank J. A. Coyne for fly stocks, C. L. Williams for help with DNA preparation for genotyping, and C. Miles for careful reading and helpful comments, B. He, E.L. Ferguson, and U. Schmidt-Ott for discussion. This work was supported by National Institutes of Health Grant 5R01GM078381 (to M.K. and M.Z.L.) and U.S. Department of Education GAANN P200A030043 (to S.E.L.).

Literature Cited

Amlou M, Moreteau B, David JR. Genetic analysis of *Drosophila sechellia* specialization: oviposition behavior toward the major aliphatic acids of its host plant. Behav Genet. 1998; 28:455–464. [PubMed: 9926614]

Bergman A, Siegal ML. Evolutionary capacitance as a general feature of complex gene networks. Nature. 2003; 424:549–552. [PubMed: 12891357]

Busturia A, Lawrence PA. Regulation of cell number in *Drosophila*. Nature. 1994; 370:561–563. [PubMed: 7914352]

Cariou ML, Solignac M, Monnerot M, David JR. Low allozyme and mtDNA variability in the island endemic species *Drosophila sechellia* (*D. melanogaster* complex). Experientia. 1990; 46:101–104. [PubMed: 2153569]

Coffman CJ, Doerge RW, Wayne ML, McIntyre LM. Intersection tests for single marker QTL analysis can be more powerful than two marker QTL analysis. BMC Genet. 2003; 4:10. [PubMed: 12816551]

Coyne JA, Rux J, David JR. Genetics of morphological differences and hybrid sterility between *Drosophila sechellia* and its relatives. Genet Res. 1991; 57:113–122. [PubMed: 2055454]

de Visser JA, Hermisson J, Wagner GP, Ancel Meyers L, Bagheri-Chaichian H, Blanchard JL, Chao L, Cheverud JM, Elena SF, Fontana W, Gibson G, Hansen TF, Krakauer D, Lewontin RC, Ofria C, Rice SH, von Dassow G, Wagner A, Whitlock MC. Perspective: Evolution and detection of genetic robustness. Evolution Int J Org Evolution. 2003; 57:1959–1972.

Driever W, Nusslein-Volhard C. The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. Cell. 1988; 54:95–104. [PubMed: 3383245]

Felix MA, Wagner A. Robustness and evolution: concepts, insights and challenges from a developmental model system. Heredity. 2008; 100:132–140. [PubMed: 17167519]

Flatt T. The evolutionary genetics of canalization. Q Rev Biol. 2005; 80:287–316. [PubMed: 16250465]

Gibson G, Dworkin I. Uncovering cryptic genetic variation. Nat Rev Genet. 2004; 5:681–690. [PubMed: 15372091]

Gregor T, Bialek W, de Ruyter van Steveninck RR, Tank DW, Wieschaus EF. Diffusion and scaling during early embryonic pattern formation. Proc Natl Acad Sci U S A. 2005; 102:18403–18407. [PubMed: 16352710]

Gregor T, Wieschaus EF, McGregor AP, Bialek W, Tank DW. Stability and nuclear dynamics of the Bicoid morphogen gradient. Cell. 2007; 130:141–152. [PubMed: 17632061]

Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB. Sepsid *even-skipped* enhancers are functionally conserved in *Drosophila* despite lack of sequence conservation. PLoS Genet. 2008; 4:e1000106. [PubMed: 18584029]

He F, Wen Y, Deng J, Lin X, Lu LJ, Jiao R, Ma J. Probing intrinsic properties of a robust morphogen gradient in *Drosophila*. Dev Cell. 2008; 15:558–567. [PubMed: 18854140]

Hey J, Kliman RM. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol Biol Evol. 1993; 10:804–822. [PubMed: 8355601]

Houchmandzadeh B, Wieschaus E, Leibler S. Establishment of developmental precision and proportions in the early *Drosophila* embryo. Nature. 2002; 415:798–802. [PubMed: 11845210]

Kliman RM, Andolfatto P, Coyne JA, Depaulis F, Kreitman M, Berry AJ, McCarter J, Wakeley J, Hey J. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics. 2000; 156:1913–1931. [PubMed: 11102384]

Kliman RM, Hey J. DNA sequence variation at the period locus within and among species of the *Drosophila melanogaster* complex. Genetics. 1993; 133:375–387. [PubMed: 8436278]

Lachaise D, David JR, Lemeunier F, Tsacas L, Ashburner M. The reproductive relationships of *Drosophila sechellia* with *Drosophila mauritiana*, *Drosophila simulans* and *Drosophila melanogaster* from the Afrotropical region. Evolution. 1986; 40:262–271.

Le Rouzic A, Carlborg O. Evolutionary potential of hidden genetic variation. Trends Ecol Evol. 2008; 23:33–37. [PubMed: 18079017]

Legal L, David JR, Jallon JM. Toxicity and attraction effects produced by Morinda citrifolia fruits on the *Drosophila melangaster* complex of species. Chemoecology. 1992; 3:125–129.

Legrand D, Tenaillon M, Matyot P, Gerlach J, Lachaise Deceased D, Cariou ML. Species-wide Genetic Variation and Demographic History of *Drosophila sechellia*, a Species Lacking Population Structure. Genetics. 2009

Lott SE, Kreitman M, Palsson A, Alekseeva E, Ludwig MZ. Canalization of segmentation and its evolution in *Drosophila*. Proc Natl Acad Sci U S A. 2007; 104:10926–10931. [PubMed: 17569783]

Louis J, David JR. Ecological specialization in the Drosophila melanogaster species subgroup: A case study of *Drosophila sechellia*. Acta Oecologica. Oecologia generalis. 1986; 7:215–229.

Lucchetta EM, Lee JH, Fu LA, Patel NH, Ismagilov RF. Dynamics of Drosophila embryonic patterning network perturbed in space and time using microfluidics. Nature. 2005; 434:1134–1138. [PubMed: 15858575]

Ludwig MZ, Palsson A, Alekseeva E, Bergman CM, Nathan J, Kreitman M. Functional evolution of a cis-regulatory module. PLoS Biol. 2005; 3:e93. [PubMed: 15757364]

Lynch, M.; Walsh, B. Genetics and Analysis of Quantitative Traits. Sunderland, MA: Sinauer Associates; 1997.

Manu, Surkova S, Spirov AV, Gursky VV, Janssens H, Kim A-R, Radulescu O, Vanario-Alonso CE, Sharp DH, Samsonova M, Reinitz J. Canalization of Gene Expression in the Drosophila Blastoderm by Gap Gene Cross Regulation. PLoS Biology. 2009; 7:e49.

McDermott SR, Kliman RM. Estimation of isolation times of the island species in the *Drosophila simulans* complex from multilocus DNA sequence data. PLoS ONE. 2008; 3:e2442. [PubMed: 18560591]

Mieiklejohn CD, Hartl DL. A single mode of canalization. Trends in Ecology and Evolution. 2002; 17:468–473.

Miles CM, Lott SE, Luengo Hendriks CL, Ludwig MZ, Manu, Williams CL, Kreitman M. Artificial selection on egg size perturbs early pattern formation in *Drosophila melanogaster*. Evolution. 2010

Moreteau B, R'Kha S, David JR. Genetics of a nonoptimal behavior: oviposition preference of *Drosophila mauritiana* for a toxic resource. Behav Genet. 1994; 24:433–441. [PubMed: 7993320]

Morton RA, Choudhary M, Cariou ML, Singh RS. A reanalysis of protein polymorphism in *Drosophila melanogaster*,*D. simulans*,*D. sechellia* and *D. mauritiana*: effects of population size and selection. Genetica. 2004; 120:101–114. [PubMed: 15088651]

Myasnikova E, Samsonova A, Kozlov K, Samsonova M, Reinitz J. Registration of the expression patterns of *Drosophila* segmentation genes by two independent methods. Bioinformatics. 2001; 17:3–12. [PubMed: 11222257]

Namba R, Pazdera TM, Cerrone RL, Minden JS. *Drosophila* embryonic pattern repair: how embryos respond to *bicoid* dosage alteration. Development. 1997; 124:1393–1403. [PubMed: 9118810]

Orgogozo V, Broman KW, Stern DL. High-resolution quantitative trait locus mapping reveals sign epistasis controlling ovariole number between two Drosophila species. Genetics. 2006; 173:197–205. [PubMed: 16489225]

Orgogozo V, Stern DL. A compilation of phenotypic differences between species of the *D. melanogaster* subgroup. 2009

R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R foundation for Statistical Computing; 2006.

R'Kha S, Capy P, David JR. Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis. Proc Natl Acad Sci U S A. 1991; 88:1835–1839. [PubMed: 1900368]

R'Kha S, Moreteau B, Coyne JA, David JR. Evolution of a lesser fitness trait: egg production in the specialist *Drosophila sechellia*. Genet Res. 1997; 69:17–23. [PubMed: 9164172]

Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000; 132:365–386. [PubMed: 10547847]

Vakulenko S, Manu, Reinitz J, Radulescu O. Size regulation in the segmentation of *Drosophila*: interacting interfaces between localized domains of gene expression ensure robust spatial patterning. Phys Rev Lett. 2009; 103:168102. [PubMed: 19911861]

Weber KE. How small are the smallest selectable domains of form? Genetics. 1992; 130:345–353. [PubMed: 1541394]

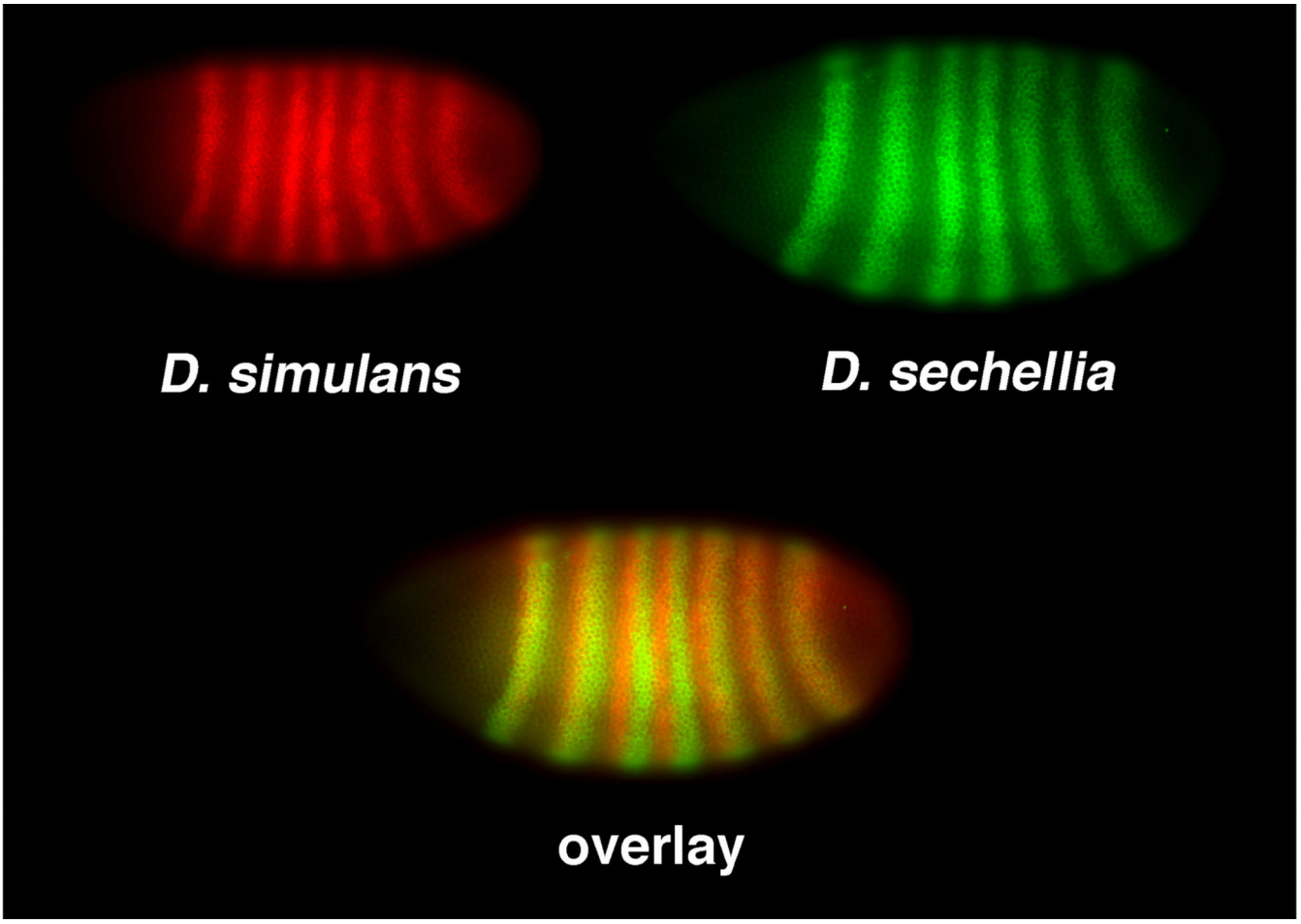

Figure 1.
Length and *eve* patterning differences in the small embryo of *D. simulans* and the large embryo of *D. sechellia*. Embryos pictured were chosen for having *eve* stripe pattern localization closest to the mean for each species, and are shown in the same dorsal-ventral orientation. For the overlay, images were scaled to produce embryos of the same length.

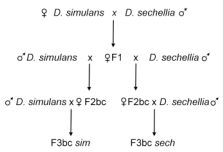

Figure 2.

Crossing scheme. A double backcross (bc) design, producing two different F3 lines, F3bc *sim* from the double backcross to *D. simulans*, and F3bc *sech* from the double backcross to *D. sechellia*.

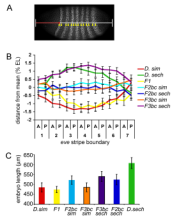

Figure 3.

Inheritance of *eve* stripe position reveals novelty in between-species crosses. (A) Illustration of *eve* expression pattern boundary measurements in a typical embryo labeled by fluorescent in situ hybridization. Stripe expression boundaries are measured relative to the anterior pole of the embryo, and standardized by embryo length. (B) Mean *eve* expression boundaries (anterior and posterior boundaries indicated with A and P) plotted as a deviation from the mean stripe position for all lines, in % embryo length, +/− standard error. Values in the positive direction indicate a posterior shift, negative values reflect an anterior shift, relative to the mean. The *sim* or *sech* designation following the backcross generation number indicates the direction of the backcross. (C) Mean embryo lengths for each line in the cross, +/− standard error.

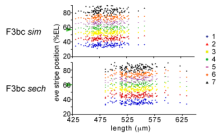

Figure 4.

F3 backcross lines precisely scale to embryo length within a line. Relative positions (in % embryo length) for *eve* stripes (anterior boundaries only for clarity, 1–7), for each of the F3bc *sim* and the F3bc *sech* embryos, plotted against the length of that embryo. The effect of age on stripe positions has been removed (by a linear model). The two lines have different mean stripe positions (as illustrated for stripe 4 by green triangles, at 58% and 60% embryo length for F3bc *sim* and F3bc *sech* respectively) across the entire distribution of embryo length.

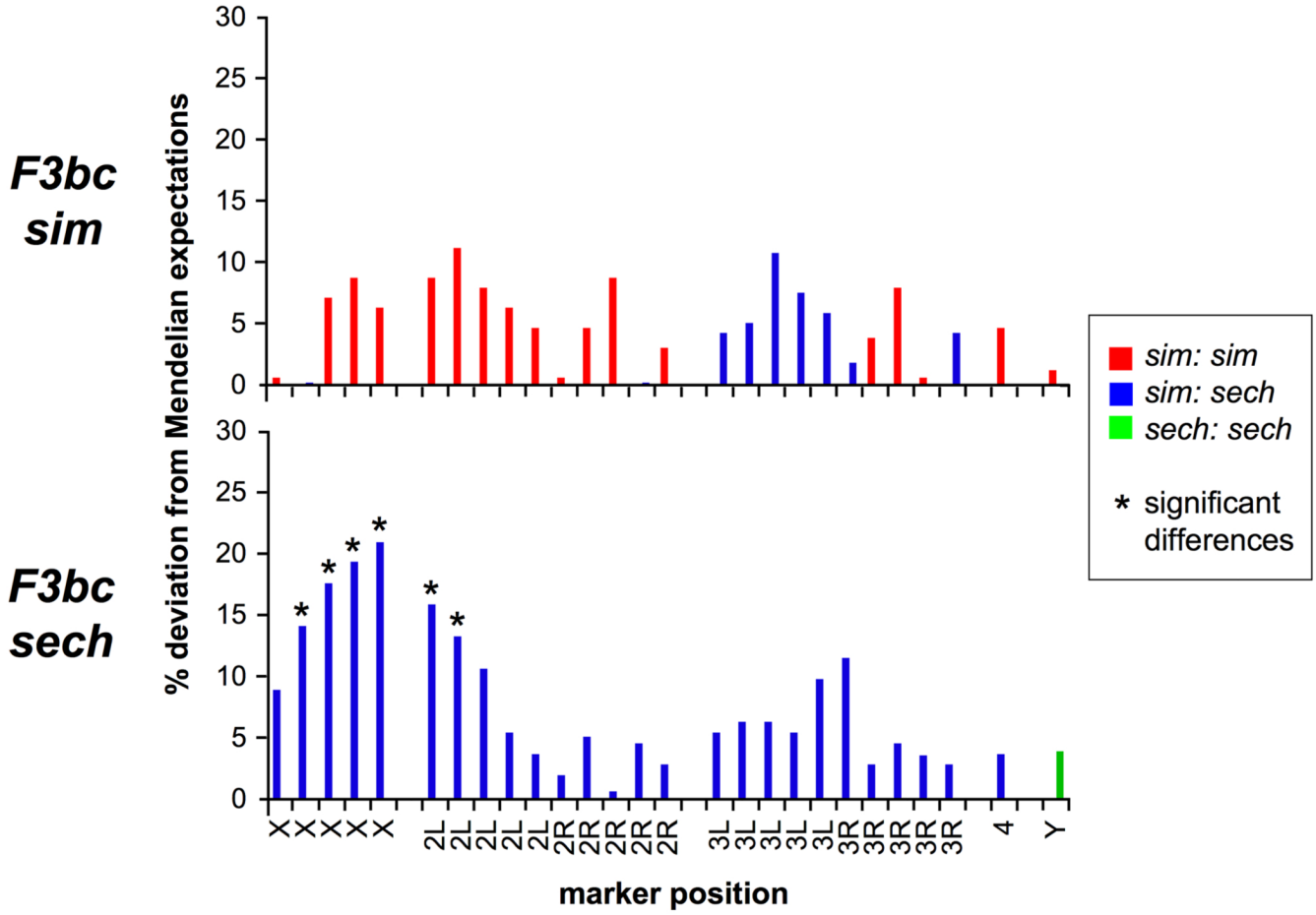

Figure 5.

Departures from Mendelian inheritance. Counts of genotypes were compiled for each marker, and the distribution of these genotypes is compared to the expected frequencies of single loci under Mendelian inheritance. The genotype in excess is plotted by the percent of that excess from Mendelian expectations, at each marker across the genome. These deviations are plotted separately by F3 backcross line, and significant differences ($p < 0.05$, by a two-sided t-test) are marked with asterisks.

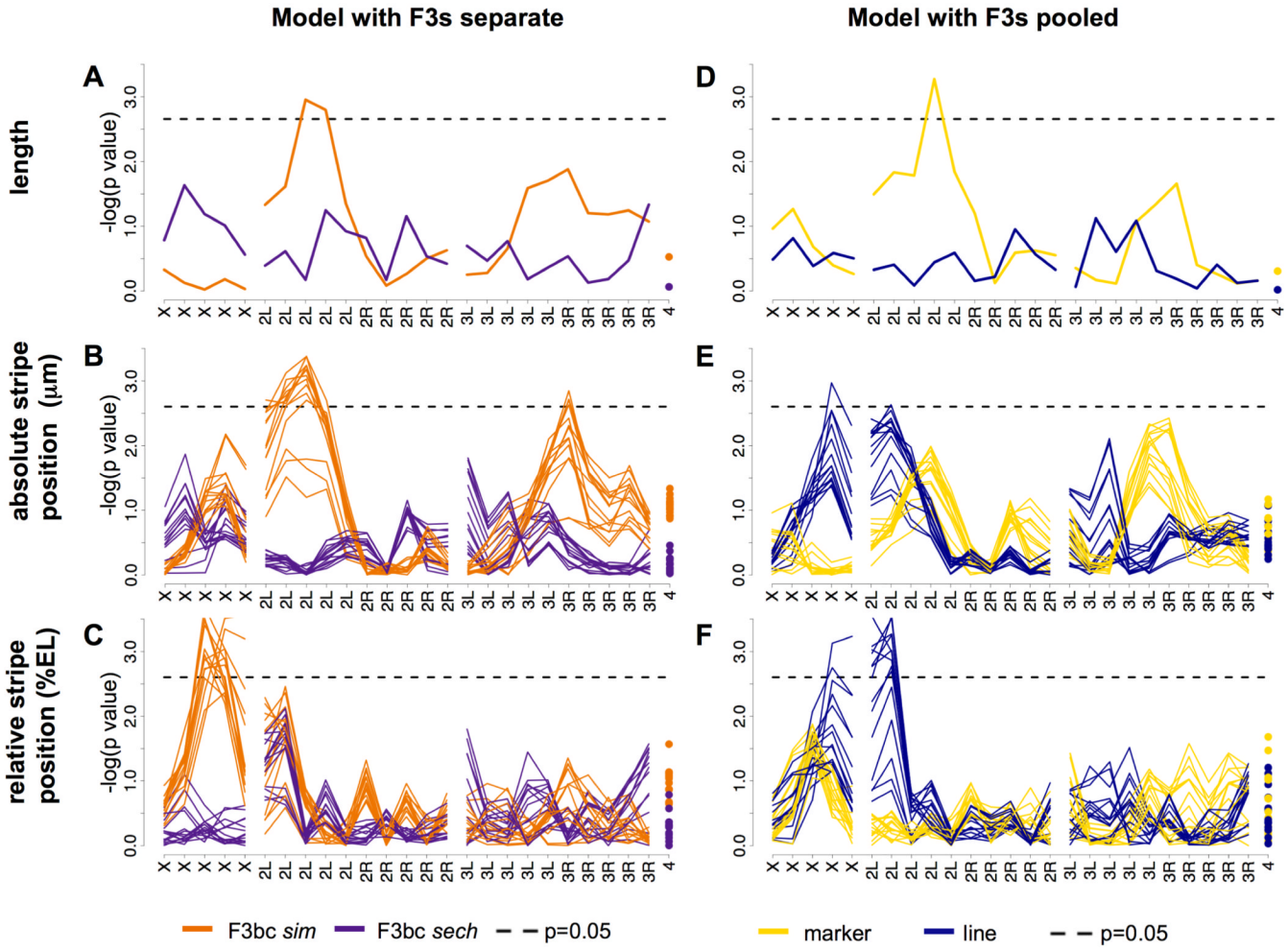

Figure 6.

Significance of genotype-phenotype associations. The significance as -log (p value) of genotype-phenotype associations is plotted over all markers across the genome. For stripe positions, each individual line is for an individual stripe boundary, 14 total for each of the anterior and posterior boundaries of seven *eve* stripes (see Supplemental Figure 4 for *gt* and *Kr*). The threshold of significance over all markers at p=0.05 was determined through permutation testing, see Methods. Significance of effects are shown for length (A, D), stripe position uncorrected for variation in embryo length ("absolute stripe position", B, E), and stripe position corrected for embryo length variation ("relative stripe position", C, F) for two different models, one where the effect of genotype on phenotype are examined in each F3bc line separately (A–C), and the other where all F3s are pooled together, and line (F3bc *sim* or F3bc *sech*) is an effect in the model (D–F).