Whole-Genome Expression Plasticity across Tropical and Temperate Drosophila melanogaster Populations from Eastern Australia

Mia T. Levine,*† 1 Melissa L. Eckert,1 and David J. Begun1

1 Department of Evolution and Ecology, University of California, Davis
† Present address: Fred Hutchinson Cancer Research Center, Division of Basic Sciences, 1100 Fairview Avenue, Seattle, Washington.
* Corresponding author: E-mail: mlevine@fhcrc.org.

Associate editor: John True

Abstract

The genotypic signature of spatially varying selection is ubiquitous across the Drosophila melanogaster genome. Spatially structured adaptive phenotypic differences are also commonly found, particularly along New World and Australian latitudinal gradients. However, investigation of gene expression variation in one or multiple environments across these well-studied populations is surprisingly limited. Here, we report genome-wide transcript levels of tropical and temperate eastern Australian populations reared at two temperatures. As expected, a large number of genes exhibit geographic origin-dependent expression plasticity. Less expected was evidence for an enrichment of down-regulated genes in both temperate and tropical populations when lines were reared at the temperature less commonly encountered in the native range; that is, evidence for significant differences in a “directionality” of plasticity across these two climatic regions. We also report evidence of small scale “neighborhood effects” around those genes significant for geographic origin-dependent plasticity, a result consistent with the evolution of high level, likely chromatin based gene regulation during range expansion in D. melanogaster populations.

Key words: clinal variation, expression plasticity, chromatin, temperature, Drosophila.

Introduction

Population genetic processes such as mutation–selection balance, genetic drift, and balancing selection contribute to the abundant genetic variation observed in natural populations. One particularly pervasive form of balancing selection, spatially varying selection, can maintain variation across heterogeneous landscapes if fitness variation is sufficiently great to overcome the homogenizing effects of gene flow (Felsenstein 1976; Endler 1977). The resultant adaptive phenotypic differences across populations often exhibit plasticity in response to different environmental conditions (Robinson and Partridge 2001; Gilchrist and Huey 2004; Swindell et al. 2007; Liefting et al. 2009).

Genetic variation for phenotypic plasticity, formally referred to as genotype × environment interactions (GEI), is widespread (DeWitt and Scheiner 2004). In Drosophila pseudoobscura, for example, as rearing temperature shifts from 14 to 21 °C, abdominal bristle number increases in some genotypes but decreases in others (Gupta and Lewontin 1982). In principle, both neutral and nonneutral evolutionary forces could maintain variants exhibiting GEI within and between natural populations. However, clinal variation for GEI, particularly along environmental gradients where gene flow is high, represents a strong case for selection contributing to geographically structured genetic variation for phenotypic plasticity.

New World and Australian latitudinal gradients have emerged as classic ecological contexts for exploring clinal variation, especially in the model system D. melanogaster (e.g., Stalker 1976; Oakeshott et al. 1982; Berry and Kreitman 1993; Verrelli and Eanes 2001; Turner et al. 2008). Microsatellite loci show virtually no spatial structure (Gockel et al. 2001; Kennington et al. 2003), yet many single nucleotide polymorphisms (Verrelli and Eanes 2001; Sezgin et al. 2004; Turner et al. 2008), inversions (Stalker 1976; Knibb et al. 1981; Umima et al. 2005), and phenotypes (e.g., Gockel et al. 2001; Hoffmann et al. 2002; Schmidt et al. 2005) vary clinally across latitude, implicating the action of spatially varying selection. Moreover, several clinal phenotypes exhibit genetic variation for plasticity (Robinson and Partridge 2001; Gilchrist and Huey 2004; Liefting et al. 2009).

Although there are several examples of clinal gene expression variation (Whitehead and Crawford 2006; Swindell et al. 2007), there are no published investigations of clinal variation for expression plasticity across D. melanogaster populations derived from these well-studied geographic regions, on a single gene or a whole-genome scale (i.e., clinal GEI, where “genotype” refers more precisely to “genotype of geographic origin”). Given that transcript abundance often exhibits plasticity across environmental conditions (Causton et al. 2001; Kreps et al. 2002; Bochdanovits et al. 2003), one would expect many genes to exhibit expression plasticity correlated with locally adapted genotypes.

If such variation exists, it may be attributable to variation at local regulatory sequence or at regulatory proteins, such
as transcription factors (Li et al. 2006; Smith and Kruglyak 2008). Genes regulated by chromatin remodeling factors may harbor a distinct transcriptional signature, namely a "neighborhood effect" of correlated expression change (Sproul et al. 2005; Batada et al. 2007). The physical scale of this effect (generally greater than five genes) renders unlikely a role for variation at local features, such as bidirectional promoters or read-through transcription (Drosophila: Boutanaev et al. 2002; Spellman and Rubin 2002; Kalmykova et al. 2005; Blanco et al. 2008; Mezey et al. 2008; humans: Caron et al. 2001; zebrafish: Ng et al. 2009; mouse: Li et al. 2005). Moreover, several recent studies have documented an association between chromatin structure and physical coexpression. For example, genomic regions harboring elevated nucleosome occupancy are associated with elevated coexpression in yeast cells (Batada et al. 2007). In fruit flies, the genomic region surrounding a cluster of nonhomologous, male-germline coexpressed genes is sensitive to DNase I in germ cells but not sensitive in somatic cells (Kalmykova et al. 2005). Recent evidence that chromatin remodeling factors show clinal variation (Harr et al. 2002; Levine and Begun 2008) suggests that chromatin regulation may contribute to adaptive expression variation along latitudinal gradients. If so, clinal expression plasticity may span physically clustered gene neighborhoods.

Here, using natural populations of D. melanogaster from the tips of the well-studied eastern Australia latitudinal cline, we investigate clinal expression plasticity by testing for an interaction between genotype and rearing temperature, where genotype refers to either tropical or temperate origin. We also characterize genotypic variation for expression plasticity at physically clustered unrelated genes to test for evidence of adaptive chromatin domain organization.

Materials and Methods

D. melanogaster isofemale lines were collected from temperate Australia (Cygnet, Southern Tasmania, latitude 43°09'12") and tropical Australia (Innisfail, Queensland 17°31'20") during the summer of 2008 (provided by A. Hoffmann). The long-term monthly averages for the temperate region range from 11.9 to 21.9 °C and the equivalent averages for the tropical region range from 24.1 to 30.8 °C (www.weatherzone.com.au).

Fifteen isofemale lines from each population were reared at 25 °C on standard medium for more than five generations before two replicate single pair matings per isofemale line were aspirated into two separate vials and placed immediately into either an 18 °C or a 30 °C incubator with 12 h day/night cycles. We therefore had four treatment combinations represented by 15 single pair matings per treatment: tropical at 18 °C, temperate at 18 °C, tropical at 30 °C, and temperate at 30 °C. The choice of temperatures was based on both the native range of temperatures experienced by these populations in nature (see above) and the known fertility limits of natural populations of D. melanogaster (Chakir et al. 2002).

Three pools of 15 males were generated for each treatment combination by collecting three F1 adult males from each vial (using light CO2) 24–48 h posteclosion, and adding each male to one of three pools. These pools therefore served as both technical and biological replication (the isofemale lines are highly heterozygous). Prior to freezing down each pool of 15 males in liquid nitrogen, flies were allowed to acclimate at room temperature for 120 min to enrich for (though not rigorously account for) expression plasticity specifically associated with the temperature at which development occurred.

Total RNA was extracted using Trizol (Invitrogen), followed by cDNA synthesis, amplification, and labeling using the Affymetrix 3'IVT Express Kit (Affymetrix). Three replicate Affymetrix GeneChip 2.0 expression arrays were used for each population sample at each of two temperatures for a total of 12 arrays. Arrays were probed, stained, washed, and scanned at the UC Davis Microarray Core Facility following Affymetrix guidelines.

Probesets called absent by the Affy GCOS software on at least 6 of the 12 arrays were removed from the analysis, leaving 12,560 of the original 18,952 probesets. All chips were normalized and background corrected using rma from the Bioconductor affy package (Gentleman et al. 2004) followed by calculation of the mean log2 expression intensity. The genefilter package was used to remove all probesets below an interquartile range of 0.5 across arrays (Scholten and von Heydebreck 2005), leaving 1,760 probesets for the analysis (see supplementary fig. 1a–c, Supplementary Material online for quality assessment figures). These 1,760 probesets therefore represent only those exhibiting substantial variation across arrays. Using the limma package, three a priori contrasts were set up: 1) tropical at 18 °C — tropical 30 °C; 2) temperate at 18 °C — temperate at 30 °C; and 3) (tropical at 18 °C — temperate at 30 °C) — (tropical at 18 °C — tropical at 30 °C). The first two contrasts test for plasticity across the two temperatures for each genotype pool separately, whereas the third contrast effectively tests for the interaction of temperature and geographic origin (“GEI contrast” hereafter). Although traditional GEI cannot be tested for directly using pooled samples (rather than individual genotypes), plasticity differences between the tropical and temperate genotype pools should reflect, on average, differences between these pools that are greater than differences within pools. Note that all of our contrasts use within-region comparisons and so are not confounded by intensity differences that may result from sequence differentiation between populations (due to differential hybridization of alternative alleles). False discovery rates (FDR) were generated following (Benjamini and Hochberg 1995). Gene ontology analyses were conducted using GeneMerge, which uses hypergeometric distributions to generate probabilities associated with overrepresentation in particular categories (Castillo-Davis and Hartl 2003). We report the Bonferroni corrected P values ("e-values") generated by GeneMerge.

To investigate whether geographic origin-dependent expression plasticity (the GEI contrast) is manifest at the level...
of chromosomal domain, we asked whether GEI in the neighborhood of significant genes is significantly different from randomly assembly neighborhoods. We positioned windows of varying sizes (7, 11, 15, . . . , 61 genes) centered on each of the top 50 genes (FDR < 0.09) from the GEI contrast and then calculated the median t-statistic from the “unfiltered genes,” excluding the significant gene. Windows containing a potential paralog of the significant gene were removed from the analysis. We inferred paralogy if a BLAST analysis (Altschul et al. 1990) using all coding sequences of significant genes as both the query and the database returned an e-value of 1e−5 or less (a very conservative cutoff) and overlapped >25% of coding sequence of the significant gene. The distribution of median t-statistics for each window size was compared with an empirically derived null distribution generated in two different ways. In the first, we randomly picked genes (there were no observed chromosome-specific effects) in numbers corresponding to those used to calculate observed medians (6, 10, 14, . . . , genes) for windows centered on the top 50 genes. This was repeated 1,000 times to generate P values. In the second, we randomly picked 50 genes, generated local windows of sizes corresponding to those used in the analysis of the actual data and calculated the median plasticity of these randomly placed windows. This was repeated 1,000 times. The results of the two approaches were qualitatively equivalent, so only results from the former approach are presented. All statistical analyses were conducted in R.

Results

Within-Region Expression Differences across Temperatures

At a FDR of 0.10, 11 genes were expressed differently at 18°C versus 30°C for the temperate population sample and 44 genes for the tropical population sample (supplementary table S1a and S1b, Supplementary Material online). At an FDR of 0.125, the corresponding numbers of significant genes for the temperate and tropical population samples were 48 and 341, respectively. The associated gene lists with log-fold change, t-statistic, probability value, FDR, gene name, and location can be found in supplementary table S1a and S1b, Supplementary Material online.

We observed a strong statistical signature of “directionality” in the plasticity data, with a consistent excess of significant genes characterized by higher transcript abundance in the rearing temperature more commonly encountered in the geographic origin of the population pool. Our a priori contrast design predicts that if such “home” versus “away” dynamics hold, we would observe an excess of positive t-statistics in the temperate population (contrast with home italicized: temperate at 18 °C — temperate at 30 °C) and an excess of negative t-statistics in the tropical population (tropical at 18 °C — tropical at 30 °C). We indeed observed a significant excess of positive t-statistics in the temperate population sample (genes with higher 18 °C expression) and a significant excess of negative t-statistics in the tropical population sample (genes with higher expression at 30 °C) for both the 25 and the 50 most significant genes (fig. 1; supplementary table S1a and S1b, Supplementary Material online). The next 50 most significant genes exhibited the same pattern, although only the tropical region was significantly enriched (fig. 1). These results suggest that for a gene exhibiting highly significant expression plasticity in our experiment, transcript abundance tends to be higher at the rearing temperature most frequently occurring in the population of origin (see Methods).

Gene Ontology analyses for the 50 most significant genes exhibiting plasticity across temperatures in the temperate pool (FDR ~ 0.13) revealed that categories “galactose binding” (e = 0.019), “monophenol monooxygenase” (e = 0.003), “oxygen transport” (e = 0.040), and “defense response” (e = 0.008) are significantly overrepresented, whereas in the tropical pool (FDR ~ 0.11), “lipase activity” (e = 0.002), “steroid metabolism” (e = 0.022), and “chitin metabolism” (e = 0.044) are overrepresented. In the temperate zone only, multiple odorant-binding protein genes exhibited plasticity (Obp19a, Obp8a), whereas in the tropical zone, many more metabolic genes emerged as particularly plastic (e.g., CG6776, CG6012, and CG13325). Both lipase- and carbohydrate metabolism are known to show clinal variation (reviewed in Eanes 1999; de Jong and Bochdanovits 2003), and several odorant-binding proteins are known to vary clinally along the Australian cline (Turner et al. 2008), though not the two described here. Frost, a gene associated with cold shock recovery (Sinclair et al. 2007), was expressed at significantly lower levels in the flies raised at 18°C (supplementary table S1b, Supplementary Material online) in the tropical pool. Finally, eight genes significant in this analysis (FDR < 0.125) are also significant in a recently published data set of up/down regulation across the D. melanogaster genome in response to artificial selection for increased resistance to chill coma (Telonis-Scott et al. 2009; supplemental table S2, Supplementary Material online). These experimental lines were derived from the same region of Australia as the tropical pool sampled in this analysis, and seven of the eight are significant for expression plasticity only in the tropical pool.

Geographic Origin-Dependent Plasticity

Our interest in GEI motivated a test for geographic origin-dependent plasticity differences (see Methods for our definition of “GEI” using pooled samples) by evaluating the contrast: (temperate at 18 °C — temperate at 30 °C) — (tropical at 18 °C — tropical at 30 °C). At an FDR of 0.10, 56 genes showed significant GEI, whereas at an FDR 0.125, 64 genes were significant. The associated gene list with log-fold change, t-statistic, probability value, and FDR can be found in supplementary table S1c, Supplementary Material online.

The directional biases observed for within-tropical and within-temperate plasticity contrasts predict the possibility of directional bias in this GEI contrast, but only if the same genes exhibit opposing, geographic origin-specific,
directional biases. In other words, if a given gene exhibits a positive $t$-statistic in the temperate data (higher $18^\circ\text{C}$ expression) and/or a negative $t$-statistic in the tropical data (higher $30^\circ\text{C}$ expression), then we predict a positive $t$-statistic for the combined contrast (home expression italicized: $[\text{temperate at } 18^\circ\text{C} - \text{temperate at } 30^\circ\text{C}] - [\text{tropical at } 18^\circ\text{C} - \text{tropical at } 30^\circ\text{C}]$). We indeed observed a significant excess of positive $t$-statistics for the GEI contrast. In the top 25 genes (FDR, $0.06$), there were 17 positive $t$-statistics (binomial probability $50.05$) and in the top 50 (FDR, $0.09$), 34 positive $t$-statistics (binomial probability $= 0.007$; see supplementary table S1c, Supplementary Material online). The six most significant GEI genes, five of which exhibit this home versus away pattern, provide a striking illustration of this enrichment (fig. 2).

Gene Ontology analysis of the 50 most significant GEI genes (FDR $< 0.09$) from this analysis suggests enrichment in the categories galactose binding ($e = 0.045$), "triacylglycerol lipase activity" ($e = 0.018$), "oxidoreductase activity" ($e = 0.014$), "monoxygenase activity" ($e = 0.019$), and "heme binding" ($e = 0.042$). Interestingly, 20 of these top 50 genes overlap the 50 genes most significantly diverged for transcript level in a study comparing African and European populations (Hutter et al. 2008; supplementary table S2, Supplementary Material online). Finally, two of these 50 most significant "GEI" genes are also found in regions significant for sequence differentiation in Australia (Cyp6a17, CG1304) (Turner et al. 2008).

In summary, a considerable number of $D.\ melanogaster$ genes exhibit evidence of transcript abundance plasticity in samples collected from the endpoints of the Australian latitudinal gradient, particularly in the tropical genotype pool. There appears to be geographic structure to expression plasticity variation, with tropical genotypes raised at $30^\circ\text{C}$ and temperate genotypes raised at $18^\circ\text{C}$ exhibiting higher transcript levels than the corresponding genotypes raised in opposite conditions consistent with higher expression at home and lower expression away.

Chromatin Organization and Geographic Origin-Dependent Expression Plasticity

If spatially varying selection at chromatin remodeling proteins underlies a portion of the expression variation described above, expression plasticity may also vary across genotype pools at a scale greater than individual genes. We tested for such neighborhood effects by testing for elevated GEI in the genes located near the most significant genes in our GEI analysis (see Methods). We found such neighborhood effects for geographic origin-dependent expression plasticity at the scale of 6–10 genes but not at larger physical scales (fig. 3). Interestingly, the distribution of medians around significant genes was shifted toward zero relative to the genome-wide average, suggesting that more small values occur around these significant genes. Any one (or more) of the four components of the GEI contrast may underlie this observation.
Discussion

We have reported genome-wide rearing temperature-dependent transcript levels of temperate and tropical D. melanogaster population samples from the endpoints of a well-studied cline. Although the relationship between transcript abundance and fitness remains unclear (Feder and Walser 2005), genetic differences maintained despite the homogenizing force of gene flow across eastern Australia (Gockel et al. 2001; Kennington et al. 2003) suggest that the gene expression phenotypes documented here may be adaptive. Moreover, our data suggest that a component of the expression variation differences between populations may be explained by changes in chromatin conformation at the scale of 6–10 genes. The fact that we detected significant geographic origin-dependent plasticity for several individual genes and a genomic signature of chromatin regulation in our relatively crude experiment (two temperatures, whole animal transcript abundance) suggests that more detailed analyses will reveal abundant genetic variation for environment-dependent gene expression.

We observed a strong trend toward many more genes showing expression plasticity in the tropical genotype pool than in the temperate genotype pool. This difference could simply be a function of the two particular rearing temperatures used in our experiment, that is, an artifact. For example, temperate genotypes could be similarly (or more) plastic across rearing temperatures than tropical ones, but the 18 °C treatment could have been more stressful for the tropical lines than the 30 °C for the temperate lines, thereby inducing a more dramatic plasticity response. Alternatively, if our results reflect a biologically relevant phenomenon, they suggest that temperate genotypes may be more buffered against temperature variation occurring...
during development. Further experiments are required to distinguish between these two possibilities.

Our results suggest a general pattern of lower transcript abundance in away environments (i.e., temperate genotypes at 30°C and tropical genotypes at 18°C). Environmental stress (oxidative stress, ultraviolet, heat, etc.) in Saccharomyces cerevisiae induces dramatic genome-wide down-regulation of vast numbers of genes that results in suspended cellular processes until the stress is removed (Causton et al. 2001). Thus, the directional plasticity found in our data may reflect a passive stress response in the away treatment rather than adaptive plasticity. Future studies may reveal that down-regulation (of genes unrelated to stress response induction) is generally observed when genotype “mismatches” environment.

Recent evidence of spatially varying selection at several chromatin remodeling factors raises the possibility of high-level regulation of adaptive gene expression variation. The neighborhood effects described here support the idea that chromatin regulation contributes to temperature-dependent genotypic variation in transcript abundance. The observed scale of effect is consistent with findings from other physical coexpression papers in Drosophila (Spellman and Rubin 2002; Mezey et al. 2008). If chromatin remodeling factors are at least in part responsible for elevated buffering found in the temperate pool, we might expect neighborhood effects to be more pronounced in the tropical pool. We tested this post hoc hypothesis by performing a similar neighborhood analysis using the absolute values of the t-statistics from the within-genotype plasticity data. We observed no neighborhood effect for temperate genotype plasticity at any window size but a significant neighborhood effect at six gene-windows for the tropical genotypes (supplementary fig. S2a and S2b, Supplementary Material online). These data are consistent with the idea that temperate genotypes at chromatin factors may buffer environmental variation more effectively than tropical ones or alternatively (again), an artifact. Intriguingly, the population genetic signature of recent strong selection at clinal chromatin remodeling factors in temperate populations (Harr et al. 2002; Levine and Begun 2008) also supports this idea. Moreover, five chromatin remodeling factors putatively targeted by spatially varying selection (Harr et al. 2002; Levine and Begun 2008) are associated with either the Polycomb or the Trithorax group proteins—a class of chromatin remodeling factors that maintain gene expression programs through mitosis at target loci regulating developmental programs and stem cell maintenance (Schwartz and Pirrotta 2007).

The “countergradient variation” model (Levins 1968; Conover and Schultz 1995; Grether 2005) describes the distribution of variation across environmentally heterogeneous habitats such that the effects of the environment “oppose” the effects of selection on the focal phenotype. The mean trait values are consequently more similar across natural populations experiencing different (native) environments than those raised in a single, common environment. Stabilizing selection on phenotypes that are environmentally plastic may result in underlying adaptive evolution across populations, even though mean phenotypes are the same across populations in their native habitat.

Although we have not tested directly for stabilizing selection on transcript level, our data fit the countergradient variation/genetic compensation model well. At a disproportionate number of loci, we observed relatively high levels of gene expression for home temperatures compared with the frequent down-regulation observed for away temperatures. Thus, the expression at a disproportionate number of genes was more similar across home environments than within a common environment. The contrasting responses to the same temperature shift at common loci suggest geographically differentiated genetic variation for plasticity, the regulation of which may be at least partially chromatin based.

These findings are consistent with the idea that many genes are evolving adaptively to maintain expression levels associated with fundamental developmental processes across variable environments. Integration of investigations of clinal developmental phenotypes, chromatin organization, and gene expression may reveal novel chromatin-based mechanisms of adaptive evolution.

Supplementary Material

Supplementary tables S1a–S1c, S2 and figures S1, S2 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors are grateful to A. Hoffmann for generously providing the Australian isofemales lines used in this study. The authors also thank M. Rolston at the UC Davis Microarray facility for sample preparation, hybridization, and scanning, and G. Cooper and J. Haloin for extremely useful discussions about the neighborhood analyses. M. Hahn, G. Lee, and A. Hoffmann read earlier drafts of the manuscript and provided helpful feedback. The manuscript was significantly improved by the critical comments of two anonymous reviewers. This research was supported by a National Science Foundation Doctoral Dissertation Improvement Grant to M.T.L. and D.J.B. and the National Institute of Health grant GM084056 to D.J.B. M.T.L. was supported by a UC Davis Dissertation Year Fellowship.

References

Geographically Structured Expression Plasticity · doi:10.1093/molbev/msq197


