bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Article in the Discoveries section

Evolution of Gene Expression and Splicing in Parallel Cold-Adapted Fly Populations

Yuheng Huang[1*@], Justin B. Lack[1**], Grant T. Hoppel[1], and John E. Pool[1@]

[1]Laboratory of Genetics, University of Wisconsin-Madison, Madison, Wisconsin 53706

[*]Current address: Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, CA 92697

[**]Current address: Advanced Biomedical Computational Science, Frederick National Laboratory for Cancer Research, Frederick, MD 21701

[@]e-mail: yhenh3@uci.edu; jpool@wisc.edu.

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Changes in gene regulation at multiple levels may comprise an important share of the molecular changes underlying adaptive evolution in nature. However, few studies have assayed within- and between-population variation in gene regulatory traits at a transcriptomic scale, and therefore inferences about the characteristics of adaptive regulatory changes have been elusive. Here, we assess quantitative trait differentiation in gene expression levels and alternative splicing (intron usage) between three closely-related pairs of natural populations of *Drosophila melanogaster* from contrasting thermal environments that reflect three separate instances of cold tolerance evolution. The cold-adapted populations were known to show population genetic evidence for parallel evolution at the SNP level, and here we find significant although somewhat limited evidence for parallel expression evolution between them, and less evidence for parallel splicing evolution. We find that genes with mitochondrial functions are particularly enriched among candidates for adaptive expression evolution. We also develop a method to estimate cis- versus trans-encoded contributions to expression or splicing differences that does not rely on the presence of fixed differences between parental strains. Applying this method, we infer important roles of both cis- and trans-regulation among our putatively adaptive expression and splicing differences. The apparent contributions of cis- versus trans-regulation to adaptive evolution vary substantially among population pairs, with an Ethiopian pair showing pervasive trans-effects, suggesting that basic characteristics of regulatory evolution may depend on biological context. These findings expand our knowledge of adaptive gene regulatory evolution and our ability to make inferences about this important and widespread process.

Introduction

Different species or populations often evolve similar phenotypes when adapting to similar environments (Schluter 2000; Losos, 2011). Although such parallel phenotypic evolution can be caused by structural mutations changing amino acids (Hoekstra and Coyne, 2007), there is increasing evidence that regulatory mutations altering gene expression underlie many cases of phenotypic evolution (Wittkopp & Kalay, 2012; Jones et al. 2012; Stern 2013; Sackton et al. 2019). Most studies on gene expression focus on expression abundance (the number of transcripts for a whole gene). However, alternative splicing resulting in the difference of transcript proportions can also contribute to adaptation (Barbosa-Morais et al. 2012; Gamazon and Stranger 2014; Smith et al. 2018). Therefore, understanding the transcriptomic basis of parallel phenotypic evolution requires studies of both expression abundance and alternative splicing, although the latter aspect has rarely been studied.

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The level of parallelism for gene expression abundance changes varies across study systems. In some taxa and natural conditions, significantly more genes show parallel changes (repeatedly up- or down-regulated in one ecotype relative to the other among independent population pairs) than anti-directional changes (Zhao et al. 2015; Hart et al. 2018; Kitano et al. 2018; McGirr and Martin. 2018). However, some other cases did not show significant parallel patterns, or they even show anti-parallel patterns (Derome et al. 2006; Lai et al. 2008; Hanson et al. 2017). The varying degree of parallelism may partly be explained by the level of divergence among ancestors: more closely related ancestors are expected to show a higher degree of parallel genetic evolution underlie similar phenotypic evolution (Conte et al. 2012; Rosenblum et al. 2014).

Furthermore, gene expression evolution can be caused by the same or different molecular underpinnings. Because of the difficulties of mapping expression QTL, a first step is to classify the expression evolution into two regulatory classes. Cis-regulatory changes are caused by local regulatory mutations and result in allele-specific expression difference in a hybrid of divergent parental lines (Cowles et al. 2002; Wittkopp et al. 2004). Trans-regulatory changes are caused by regulatory mutations at other loci. They modify the expression of both alleles in hybrid diploids and do not result in allele-specific expression difference (Gilad et al. 2008). The relative importance of cis- and trans-effects to parallel evolution varies among different studies systems (Wittkopp et al. 2008; McManus et al. 2010; Wittkopp and Kalay 2012; Chen et al. 2015; Osada et al. 2017; Hart et al. 2018; Nandamuri et al. 2018; Verta and Jones 2019). Most previous studies have focused on regulatory evolution between relatively distantly related lineages such as different species, from which population genetic evidence of adaptive evolution may not be available. Hence, the contributions of cis- and trans-effects to the recent adaptive divergence between populations remain mostly unknown.

In part because of the interest in the evolutionary response to climate change, *Drosophila* has been used as a model system to study the genetic basis of thermal adaptation (Hoffmann et al. 2003). Because temperature is an important environmental variable along latitudinal clines, clinal populations of *Drosophila melanogaster* have been studied for decades (Adrion et al. 2015). Along these clines, populations exhibit different degrees of cold tolerance in the expected direction, suggesting spatially varying selection related to temperature (Hoffmann and Weeks 2007; Schmidt and Paaby 2008). The recent development of genomics has allowed identification of clinal genomic variants, which are candidates for thermal adaptation (e.g., Kolaczkowski et al. 2011; Fabian et al. 2012;

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bozicevic et al. 2016; Mateo et al. 2018). There is also evidence of parallel evolution at the genomic and transcriptomic level (Reinhardt et al. 2014; Bergland et al. 2015; Machado et al. 2015; Zhao et al. 2015; Juneja et al. 2016; Zhao and Begun 2017). Some of these studies compared clines between species (which may have somewhat distinct biology), while others compared clines between Australia and North America (which both feature primarily European ancestry with clinally variable African admixture). Other transcriptomic studies have identified genes showing differential expression between sub-Saharan African and European populations (e.g., Catalan et al. 2012; Huylmans and Parsch 2014), which are separated by moderately strong neutral genetic differentiation associated with the out-of-Africa bottleneck.

More broadly, populations of *Drosophila melanogaster* from contrasting environments offer an excellent opportunity to study parallel gene regulatory evolution and its underlying mechanisms. Originating from a warm sub-Saharan ancestral range (Lachaise et al. 1988; Pool et al. 2012), D. melanogaster has occupied diverse habitats, including environments with contrasting temperature ranges. There are at least three instances in which the species expanded to cold environments: from Africa into higher latitude regions in Eurasia, from Ethiopia lowland to higher altitudes, and from South Africa lowland to higher altitudes. Populations were collected from these six regions, representing three warm-cold population pairs: Mediterranean pair (Med), collected in Egypt (EG, warm) and France (FR, cold); Ethiopian pair (Eth) collected in Ethiopia lowland (EA, warm) and highland (EF, cold); and South Africa pair (SAf), collect in South Africa lowland (SP, warm) and highland (SD, cold). Importantly, each of these population pairs has the advantage of low genetic differentiation between its warm- and cold-adapted members (Pool et al. 2017). Although the cold populations have invaded colder habitats for only ~1000-2000 years (~15k-30k generations) (Sprengelmeyer et al. 2019), they have shown signals of parallel adaptation for cold tolerance and allele frequency changes (Pool et al. 2017). In the present study, this unique system allows us to assess the degree of parallelism for transcriptomic changes underlying parallel cold tolerance evolution.

Here, we generate RNA sequencing (RNA-seq) data for multiple outbred genotypes from each of the six population samples listed above, from larval, pupal, and adult stages. We estimate gene expression and alternative intron usage levels for each sample, then identify cases of unusually high quantitative trait differentiation between each pair of warm- and cold-adapted populations. We find evidence for parallel evolution for expression abundance at the larval and female adult stage, but less parallel signal for splicing. We further tease out the cis-vs. trans-regulatory effect by sequencing the

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transcriptomics of the parental lines from different populations and their F1 offspring. Applying our resampling approach to study cis- and trans-regulatory effects, we find the relative contribution of cis- vs. trans-effects to adaptive expression differentiation varies notably across population pairs. Finally, we identify several candidate genes with both cis-effects and high F_{ST}, as potential targets of local adaptation.

Results

Phenotypic evolution related to cold adaptation

The cold populations have been shown to have a higher proportion of recovered female adults after prolonged cold exposure than the respective warm populations (Pool et al. 2017). Here for egg-to-adult survival at 15°C, we found the FR and EF populations have significantly higher survival than the ancestral range Zambia ZI population, while at 25°C benign temperature all the populations have relatively high survival (75%). Although SD is not significantly better than ZI at cold temperature for this assay, it follows the same trend. Together the results for survival and adult cold tolerance suggest the cold populations have evolved to adapt to low temperature.

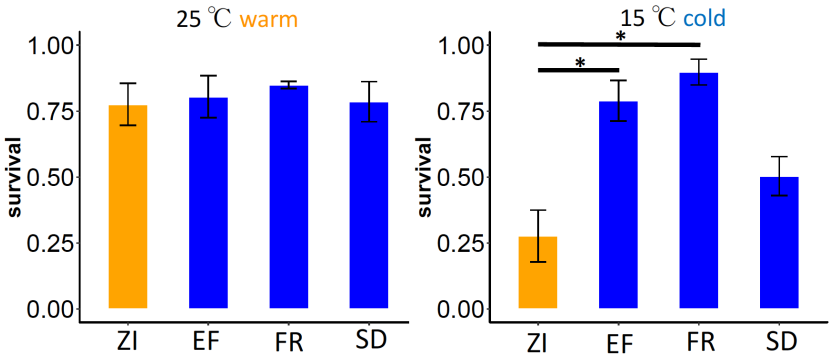

Fig 1. Survival for the ancestral warm population (ZI) and the cold populations (EF, FR and SD) at different temperatures. * indicates the survival is significant difference ($p < 0.05$). Error bar shows the standard error based on three strains.

Co-directional evolution in gene expression between population pairs

To focus on the transcriptomes of outbred genotypes, we generated eight within-population crosses from each population under a derived cold environment (15 °C). We then surveyed the transcriptomes on larvae, pupae and female adults for each cross using high-throughput RNA

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sequencing (RNA-Seq). We used a quantitative genetic index, P_{ST}, to quantify phenotypic differentiation of expression and splicing between populations in each pair. P_{ST}, analogous to F_{ST} for genetic variation, measures the amount of trait variance between populations versus total variance for a phenotype (Merila et al. 1997; Brommer 2011; Leinonen et al. 2013). The genes/introns with highest P_{ST} quantiles are more likely to be under ecological differential selection between populations than those with lower P_{ST} quantiles (Leder et al. 2015).

The numbers of genes that passed the filters for analysis were (same across population pairs): 4699 genes for larva, 5098 genes for pupa and 6786 genes for adult. To study gene expression divergence potentially under ecologically differential selection, we calculated P_{ST} (Materials and Methods). The top 20 P_{ST} outliers for each population/stage for expression and for splicing are listed in Table S1.

We used the upper 5% of P_{ST} quantile as outliers for each population pair. We found signals of parallel expression divergence in all three pairwise comparisons (Med vs. Eth; Med vs. SAf; Eth vs. SAf), where the shared outliers with co-directional changes were more than expected by chance. Across the three developmental stages, adult stage showed the highest level of parallelism (on average 0.34% of outliers were shared and changed consistently).

stage	larva		pupa		adult	
Population pairs	Med	SAf	Med	SAf	Med	SAf
Eth	1.4	1.2	**2.2**	0.79	1.4	**4.3**
SAf	**3.1**		0.16		**2.4**	

Table 1. The proportional enrichment of parallel expression abundance changes in real data relative to expectation. The proportion of parallelism expected by chance is 0.05×0.05×0.5=0.00125. All the proportions for larvae and adult stages were higher than the expectation (proportional enrichment > 1). Those that were significantly greater than the expectation are in bold (binomial one-tailed test, Bonferroni correction for nine tests in total).

To explore the broader patterns of parallel changes, we used the upper 5% P_{ST} outliers in a population pair (*Outlier Pair*) and examined whether the expression for this set of genes changed in the same direction in another pair (*Directional Pair*), regardless of outlier status in the latter pair. There

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

were excesses of co-directional changes in the *Directional Pairs* for the larval stage (Figure 1). However, the patterns were weaker for the adult stage and there were excesses of anti-directional changes for the pupal stage.

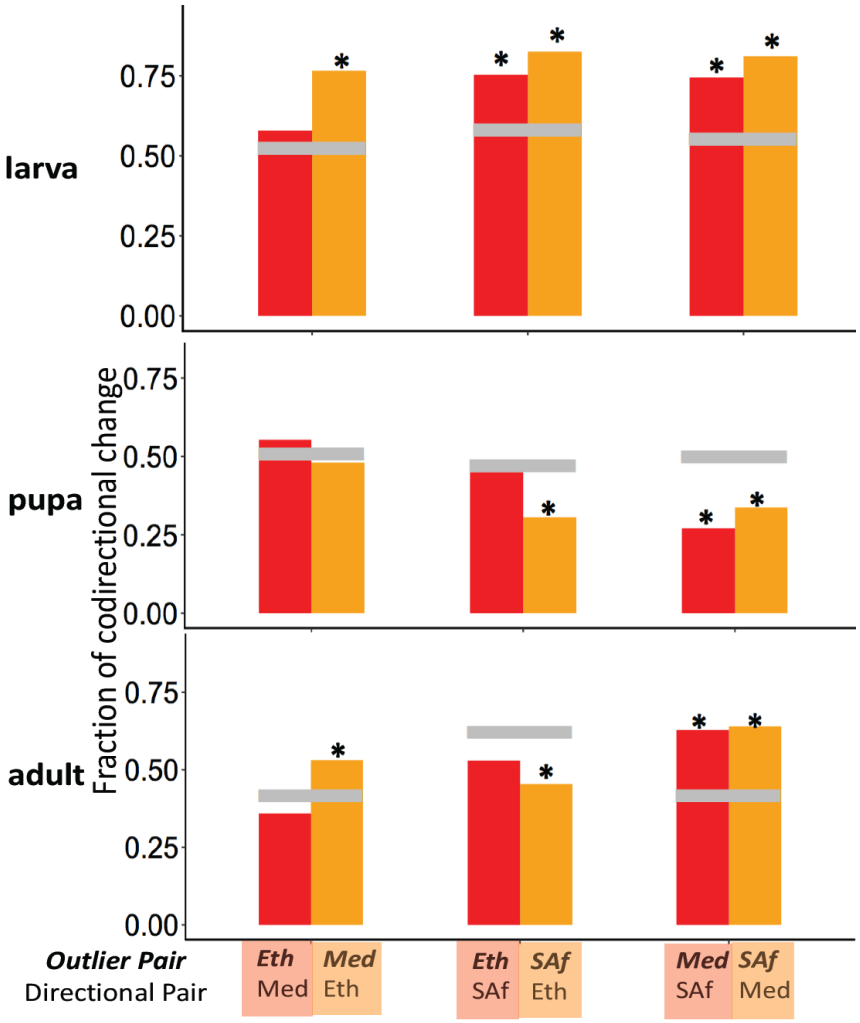

Fig. 2. Fractions of co-directional gene expression changes in *Directional Pairs* for the P_{ST} outliers identified at the *Outlier Pairs*. The grey bar indicates the fraction of co-directional change for the control set of genes. * indicates the fraction for the outliers is significantly different from the control genes after Bonferroni correction. Overall, the Med and SAf pairs showed stronger co-directional changes than the comparisons involved Eth pair.

We also performed a similar analysis for P_{ST} outliers of alternative exon junction usage. The numbers of exon junctions that passed the cutoffs for P_{ST} calculation were 976 for larva, 4604 for pupa and 7059 for adult. The patterns of co-directional changes were qualitatively similar to those for gene

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expression (Fig. S1). The fractions of co-directional changes were still highest for the larvae among the three stages; all of the comparisons except one showed an excess of co-directional changes relative to the control comparisons. For pupae, there was evidence for both co-directional and anti-directional changes. For female adult stages, the major pattern was an excess of anti-directional changes.

Enriched functional categories for the P_{ST} outliers for gene expression and exon usage

Significant Gene Ontology (GO) terms enriched in different sets of P_{ST} outliers for gene expression are listed in Table S2. Among the significant GO terms for different population pairs, we found six terms shared between Med pair and Eth pair at the adult stage. The level of sharing is significantly more than we expect by chance based on permuted outlier sets ($p < 0.001$), suggesting functional convergence for adult development to the cold environment for Med pair and Eth pair. Further, similar GO terms were identified from different pairs at different stages such as terms related to mitochondria, nucleoside metabolic process, and oxidoreductase complex. However, the majority of GO terms were unique for different pairs, suggesting that many functional changes for adaptation to cold environments may be population-specific.

Cis- and trans-acting contributions to differential gene expression abundance

One major goal is to distinguish the contributions of cis- and trans-regulatory effects on expression differentiation. First, we compared the overall strengths of cis- and trans-effects by estimating the absolute values of cis- and trans-effects for all analyzed genes. The magnitudes of trans-effects are significantly larger than the cis-effects in all three population pairs (mean absolute cis effects and trans effects are: Med pair, 0.09 vs. 0.14, $p < 2.2e\text{-}16$; Eth pair, 0.27 vs. 0.32, $p = 1.5e\text{-}14$; SAf pair, 0.14 vs. 0.15, $p < 2.2e\text{-}16$. 'Mann-Whitney' paired test.). Moreover, we found strong negative relationships between cis- and trans-effects within each population pair (Fig. S2), where the cis- and trans-effects are generally in the opposite directions.

Next, we used our conservative permutation approach (see Materials and Methods) to study how many genes show a significant cis-effect, trans-effect or both. Averaged across population pairs, we found that for the expression abundance, 12.6% show cis only regulatory effects while 26.2% show trans only effects, consistent with trans-effects being stronger on average than cis-effects (Table 2). Because we are interested in the regulatory contributions to adaptive evolution of gene expression, we further compared the ratio of trans only to cis only genes between P_{ST} outliers and non-outliers. The

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ratio is significantly lower in P_{ST} outliers than that in non-outliers for Med pair (p = 0.003) but the pattern reverses for Eth pair (p = 8.5e-10; Fig. 3 left; Fig. 4). While the ratio is not different for SAf pair (p = 0.999). Hence, there is not a consistent pattern of greater usage of cis- versus trans-regulatory changes in putatively adaptive expression changes compared with transcriptome-wide differentiation.

On average across population pairs, about 31% of all genes in the analysis showed both effects (Table 2). Among the outlier genes showing both effects (Fig 4), the vast majority (85%) of them were in opposite directions (*i.e.* compensatory). Similarly, most of the control genes with both effects showed apparent compensation (88%), which is consistent with the transcriptome-wide negative relationship between cis- and trans-effects (Fig. S2). Although the pattern can be biologically meaningful, it may also represent an artifact from using the same F1 expression data for allele specific expression (ASE) estimation to infer both cis- and trans-effects. Any measurement error on ASE will introduce artifactual negative correlation between cis- and trans-acting changes (see Discussion below).

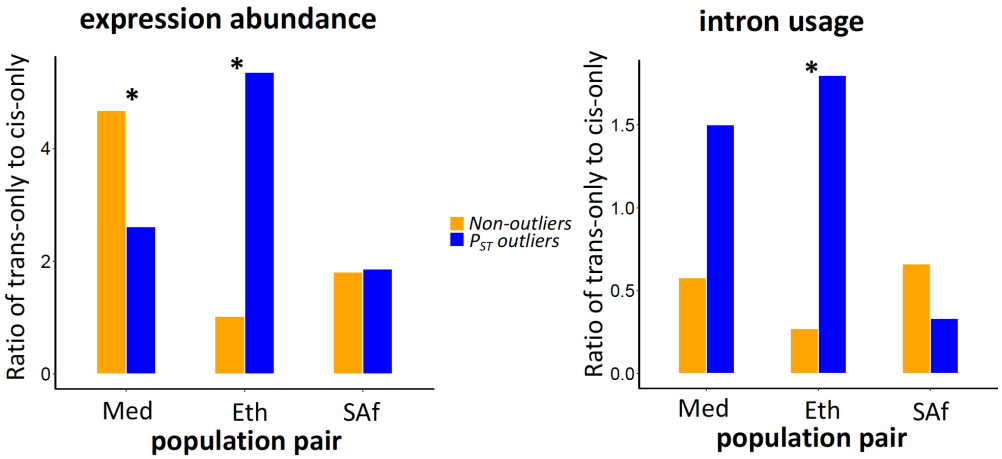

Fig 3. Ratio of genes (left) and intron usage (right) showing trans only regulatory effects to those showing cis only effects for P_{ST} outliers and non-outliers. * indicates the ratio is significantly different between P_{ST} outliers and non-outliers (p < 0.005).

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

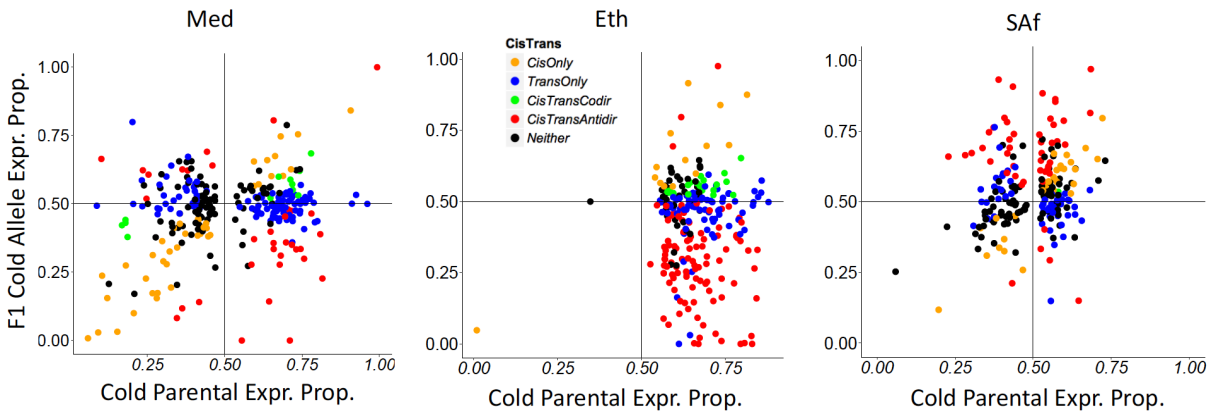

Fig 4. The relative expression proportion of cold alleles in parental and F1 datasets. The colors of the points indicate the regulatory mechanism of the outlier genes inferred by our approach.

Since the cis-regulatory mutations contributing to local adaptation may show differentiation in allele frequency between populations, we examined whether genes with cis-effects (including cis only genes and genes with both cis- and trans-effects) show association with high F_{ST} between the two warm and cold populations – for both window F_{ST} (F_{ST_winmax}) and SNP F_{ST} (F_{ST_SNPmax}). We found that only genes with high F_{ST_winmax} are enriched in cis-regulated genes in Med pair (the proportion for high F_{ST_winmax} is: cis-effect genes, 22.5%; control non-outliers, 11%; p = 0.037). However, there was no significant enrichment for high F_{ST_SNPmax} in cis-regulated genes. Moreover, there was no enrichment for either window or F_{ST_SNPmax} with cis-regulated genes in the other population pairs.

We then focused more narrowly on a set of outlier genes that showed both significant cis-effect only and higher F_{ST} quantile (upper 5%), which could reflect adaptive regulatory evolution targeting the surveyed sequences or nearby sites. For Med pair, there were three cis-genes showing high window F_{ST} (*Ciao1*, *Cyp6a17,* and *NiPp1*) and one cis-gene showing high F_{ST_SNPmax} (*spin*). Interestingly, *Cyp6a17* encodes a cytochrome P450 protein that is required for temperature preference behavior (Kang et al. 2011). *Cyp6a17* variants have also been associated with insecticide resistance (Battlay et al. 2018; Duneau et al. 2018). *Cyp6a17* is impacted by a polymorphic whole-gene deletion with contrasting frequencies between populations (Chakraborty et al. 2018), underscoring its likely role in local adaptation. The *spin* gene is essential for mTOR reactivation and lysosome reformation after starvation and has important effects on nervous system and courtship behavior (Nakano et al. 2001; Rong et al. 2011). For Eth pair, there were two genes with high window F_{ST} (*CG3529* and *mle*) and one with high F_{ST_SNPmax} (*Aldh-III*), which encodes a protein that confers a xenobiotic stress resistance and neutralises the lipid aldehydes formed after the attack of reactive oxygen and radicals (Arthaud et

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

al. 2011; Mateo et al. 2014). For SAf pair, one cis-gene showed both high window F_{ST} and high F_{ST_SNPmax} (*AGO2*) and one showing high F_{ST_SNPmax} (*eca*). *AGO2* is involved with antiviral defense and developmental regulation (Deshpande et al. 2005; Nayak et al. 2010) and was previously found to contain fixed differences between European and African populations (Pool 2015). For the genes showing high F_{ST_SNPmax} (*spin*, *Aldh-III, AGO2,* and *eca*), we plotted the SNP F_{ST} along the gene region to show the sites that are the most likely targets of selection (Fig. 5). Interestingly, for *spin*, *Aldh-III* and *eca*, the highest F_{ST} sites are located in noncoding regions (intron region for *spin*, downstream of the gene for *Aldh*-III and upstream of the gene for *eca*). While for *AGO2*, the highest F_{ST} site was located in the protein coding sequence.

Further, we identified seven genes showing consistent cis-effects across two population pairs (cis-effect favored expression of the same cold or warm parental alleles). Similarly, these shared cis- effect genes might show high genetic differentiation specific for cold populations in the two focal pairs. Using the "Population Branch Excess" statistic (PBE) results from Pool et al. 2017, we found that one gene named *Tollo* contained SNPs showing high cold-population specific differentiation (PBE quantile < 0.05) in both Eth and SAf pairs. *Tollo* is known to be have several important functions: innate immune response, glucose and protein metabolism regulation, and peripheral nervous system development (Seppo et al. 2003; Yagi et al. 2010; Akhouayri et al. 2011; Ballard et al. 2014).

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

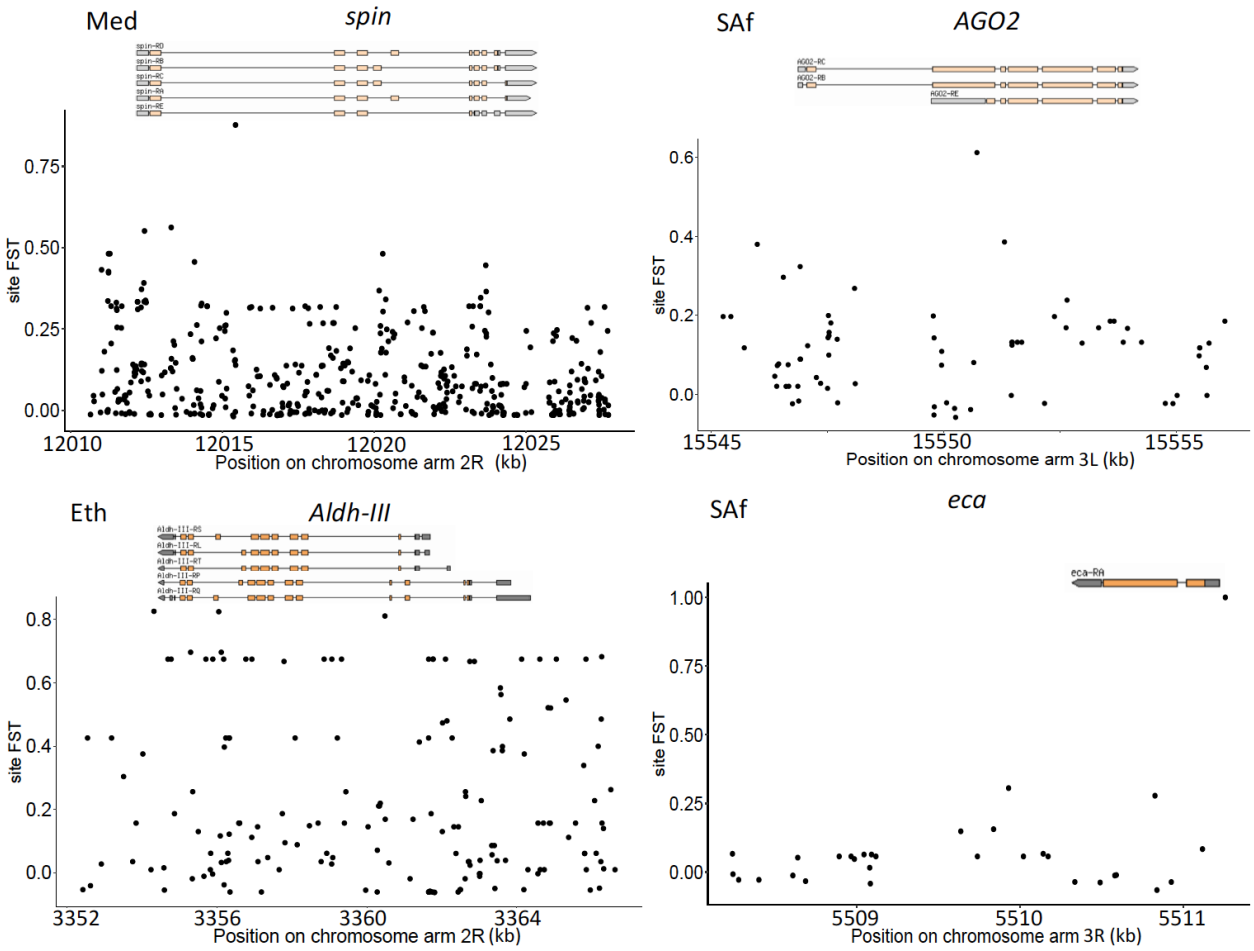

Fig 5. SNP F_{ST} along the gene with flanking region of 2kb for *spin* in Med pair, *Aldh-III* in Eth pair and *AGO2* and *eca* in SAf pair. The top diagram depicts the coding (orange) and non-coding (gray) exon, captured from GBrowse 2 of *D. melanogaster* (R5.57) from FlyBase (St. Pierre et al. 2014).

pair	expression type	Total tests	Cis only	Trans only	Both co-dir	Both anti-dir	Neither
Med	P_{ST} outliers	282	41 (15%)	107 (38%)	12 (4%)	34 (12%)	88 (31%)
	Non-outliers	4887	431 (9%)	2015 (41%)	288 (6%)	917 (19%)	1236 (25%)
Eth	P_{ST} outliers	251	14 (6%)	75 (30%)	20 (8%)	104 (41%)	38 (15%)
	Non-outliers	4851	967 (20%)	985 (20%)	147 (3%)	1905 (39%)	847 (17%)
SAf	P_{ST} outliers	224	29 (13%)	54 (24%)	1 (0.4%)	49 (22%)	91 (41%)

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

	Non-outliers	4721	431 (9%)	780 (17%)	60 (1%)	1289 (27%)	2161 (46%)

Table 2. Numbers of gene expression abundance traits showing different regulatory effects for P_{ST} outliers and non-outliers. The percentage in parentheses indicates the fraction of genes in each category relative to total genes in the tests.

pair	Intron type	Total tests	Cis only	Trans only	Both co-dir	Both anti-dir	Neither
Med	P_{ST} outliers	21	2 (10%)	3 (14%)	0	3 (14%)	13 (62%)
	Non-outliers	375	78 (21%)	45 (12%)	2 (1%)	81 (22%)	169 (45%)
Eth	P_{ST} outliers	37	5 (14%)	9 (24%)	1 (3%)	10 (27%)	12 (32%)
	Non-outliers	456	170 (37%)	46 (10%)	5 (1%)	115 (25%)	120 (26%)
SAf	P_{ST} outliers	18	3 (17%)	1 (6%)	0	10 (56%)	4 (22%)
	Non-outliers	460	62 (13%)	41 (9%)	1 (0.2%)	168 (37%)	188 (41%)

Table 3. Numbers of introns showing different regulatory effects for P_{ST} outliers and non-outliers. The percentage in parentheses indicates the fraction of introns in each category relative to the total introns in the tests.

Cis- and trans-acting contributions to differential intron usage

For all intron usage, we found the magnitude of trans-effects on average to be higher than that of cis-effects (mean absolute cis effects and trans effects are: Med pair, 0.15 vs. 0.18, p = 6.2e-06; Eth pair, 0.32 vs. 0.34, p < 3.3e-5; SAf pair, 0.20 vs. 0.22, p = 0.00027. 'Mann-Whitney' paired test.). Although there are few outlier introns tested for cis- and trans-regulatory effects (Table 3) because of the limited diagnostic SNPs located within the intron regions, we found the numbers of significant trans only introns were higher than that of significant cis only introns summing across three population pairs. While for non-outlier introns, the significant trans only introns are fewer than the significant cis only introns (the numbers of cis only vs. trans only introns are ten vs. 13 for outliers; 311 vs. 132 for non-outliers; $X^2 = 6.1$; df = 1; P = 0.014). Thus, trans-regulated splicing changes appear to be relatively more common for putatively adaptive than for putatively neutral population differences, although the pattern varies geographically (Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

For the outlier introns showing cis-effects (including only cis and both cis and trans), the maximum F_{ST} (F_{ST_SNPmax}) around their splice sites tends to be higher than that for non-outliers (average F_{ST_SNPmax} for cis outlier vs. non-outliers: Med: 0.184 vs. 0.153; Eth: 0.152 vs. 0.134; SAf: 0.081 vs. 0.055). Because there are few cis outlier introns with SNPs located around splice sites, all three comparisons are non-significant based on Wilcoxon signed-rank test. Across the three comparisons, four genes contained cis-regulated introns with high F_{ST_SNPmax} around splice sites (top 15% quantile of F_{ST_SNPmax}). One identified in Med pair is *Usp10*, which is known to regulate Notch Signaling during development (Zhang et al. 2012). One gene identified in SAf pair is *Sdc*, which has been shown to have neuromuscular functions (Johnson et al. 2006; Chanana et al. 2009). The other two genes (*DOR* and *Jabba*) were related to lipid metabolism (Francis 2010; McMillan et al. 2018). Since the two lipid related genes were identified in highland pairs (*DOR* in SAf and *Jabba* in Eth), putative changes in lipid metabolism might facilitate adaptation to high altitude environments.

Discussion

Parallel evolution has often been studied at the population genetic and trait levels, but it has less frequently been analyzed at the transcriptome level (Stern 2013; Juneja et al. 2016). In this study, we used three recent instances of adaptation to colder climates in *Drosophila melanogaster* to study the evolution of gene expression and alternative splicing. The signal of parallel evolution in expression abundance varied among developmental stages, with a higher degree of parallelism for larva and adult stages than pupa. Further, we studied cis- and trans-regulatory evolution in the context of this ecological adaptation. For gene expression abundance, we found geographically variable patterns of cis- versus trans-effects for highly differentiated expression outliers relative to the other genes. Specifically, P_{ST} outliers show enrichment of cis-effect relative to background genes in Med pair while outliers show enrichment of trans-effect in Eth pair. For splicing, we also found P_{ST} outliers enriched for trans-effects in Eth pair. This pattern of trans-effects contributing to differential expression in the Ethiopian pair raises the possibility of large-scale gene regulatory network changes in this phenotypically distinctive highland population, which might result from a few genetic changes or from many.

Although there are significant patterns of parallel evolution in expression abundance between population pairs, the majority of outlier genes/intron usages are not shared between pairs. The low level of detected parallelism could reflect a high false negative rate, for example due to limited spatiotemporal expression of relevant differences (perhaps contributing to the greater parallelism

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

339 detected in larvae, which have somewhat less tissue diversity). Alternatively, it might reflect the
340 different selection agents in the different natural habitats as well as the demographic histories for
341 these populations. The cold FR population colonized a higher latitude environment than the related
342 warm population EG, whereas the other two cold populations colonized higher altitude environments
343 where the selection agents may include air pressure, desiccation and ultraviolet radiation (Pool et al.
344 2017). Also, the Med pair has experienced the trans-Saharan bottleneck (Pool et al. 2012;
345 Sprengelmeyer et al. 2019) and the standing genetic variation may be altered, potentially resulting in
346 a distinct evolutionary path for FR compared to other two cold populations. Although EF and SD
347 have both adapted to higher altitudes (EF at 3,070 meters above sea level, SD at 2,000), SD is
348 seasonally cold (like FR) whereas EF is perpetually cool. Notably, the EF population exhibits distinct
349 phenotypic evolution such as darker pigmentation (Bastide et al. 2014), larger body size (Pitchers et
350 al. 2013; Lack et al. 2016), and reduced reproductive rate (Lack et al. 2016). Therefore, the
351 underlying transcriptomic evolution for EF may partly reflect its unique phenotypic evolution.
352 Indeed, the Eth pair shows the least parallelism at gene level with the other pairs (Table 1; Fig. 2),
353 although it shared some parallel functional categories with the Med pair (Table S1).
354
355 Compared to the expression abundance, the pattern of parallelism is much weaker for intron usage
356 (Fig. 2, Fig S1), which may partly stem from lower power to detect intron usage change (only a small
357 proportion of reads are informative for exon junctions). However, we still found the Med pair and
358 SAf pair show more parallel changes than the combinations with the Eth pair, which is consistent
359 with results for expression abundance. Given the increasing evidence for alternative splicing
360 contributing to environmental response and adaptation (e.g., Singh et al. 2017; Signor and Nuzhdin
361 2018; Smith et al. 2018), we need to study both expression abundance and splicing to fully
362 understand the evolution at the transcriptome level. The development of sequencing approaches with
363 long reads that cover the entire transcripts will enable us to quantify isoforms frequency directly and
364 broaden the scope of alternative splicing variation that can readily be quantified. Since splicing
365 changes during development and among tissues (Brown et al. 2014; Gibilisco et al. 2016), a detailed
366 sampling throughout development of different tissues will also be necessary to understand the role of
367 splicing on ecological adaptation.
368
369 We found trans-effects are generally larger than the cis-effects across the transcriptome, which is
370 consistent with some previous studies (e.g., McManus et al. 2010; Coolon et al. 2014; Albert et al.
371 2018; Hart et al. 2018) but not with others (e.g., Lemmon et al. 2014; Mack et al. 2016; Verta and

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

372 Jones 2018). The transcriptome-wide stronger trans-effects can be caused by random regulatory
373 changes biased toward trans-regulation because of the larger trans-mutational target size (Landry et
374 al. 2007). To focus on the evolved changes related to adaptation, we compared the ratios of genes
375 with trans-effects to those with cis-effects between P_{ST} outliers and non-outliers and saw patterns
376 varied among population pairs (Fig. 3). Cis only genes are enriched in the outliers of Med pair while
377 trans only genes are enriched in the outliers of Eth pair, suggesting different adaptive regulatory
378 mechanisms responding to ecological shifts. These results suggest that both cis- and trans-acting
379 expression changes may be viable mechanisms of adaptive evolution. For intron usage, we found
380 more differences showing cis-effects than trans-effects across the transcriptome (Table 3), consistent
381 with splicing differences between *Drosophila* species studied by McManus et al. 2014. These results
382 may be unsurprising since alternative splicing in *Drosophila* is mostly regulated by nearby sequences
383 (Venables et al. 2011; Kurmangaliyev et al. 2015). However, particularly for the Ethiopian pair, we
384 observed a relative excess of trans-regulation among P_{ST} outliers, which is consistent with expression
385 abundance results for this same population pair (Fig. 3). Therefore, the genetic basis of gene
386 regulatory evolution may depend on the mechanism (e.g. transcription vs. splicing), the evolutionary
387 scale, and population-specific evolutionary events.
388
389 When we considered genes/introns showing both cis- and trans- effects, we observed that the two
390 types of effects were generally in opposite directions (anti-directional. Table 3). This is consistent with
391 the idea that gene expression is under stabilizing selection in general and gene regulatory networks
392 evolve negative feedback to buffer effects of regulatory changes (Denby et al. 2012; Coolon et al.
393 2014; Bader et al. 2015; Fear et al. 2016). With regard to our P_{ST} outliers, it is possible that cis-acting
394 changes might have evolved to compensate for unfavorable pleiotropic impacts of adaptive trans-
395 regulatory evolution. However, negative correlations between cis- and trans-effects can also be an
396 artifact coming from the measurement error on F1 expression data. Because the F1 data was used to
397 estimate ASE and compared it to 0.5 (cis-effect null) and to parental expression proportion (trans-
398 effect null), measurement error will introduce artifactual negative correlation between cis- and trans-
399 acting changes. Therefore, whether the opposing effects between cis- and trans-acting changes are
400 biologically meaningful will require further study. As Fraser (2019) and Zhang and Emerson (2019)
401 proposed, using independent F1 replicates or other approaches such as eQTL mapping to infer cis-
402 and trans-effects separately is necessary to affirm evidence of compensatory evolution.
403

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We expect that the adaptive expression divergence caused by cis-regulatory changes should leave a signal in the nearby genomic region. Therefore, we used F_{ST} statistics to quantify genetic differentiation for the region around the focal genes. Window F_{ST} is sensitive to classic hard sweeps, and relatively useful for incomplete sweeps and moderately soft sweeps, but it is less useful for soft sweeps with higher initial frequencies of the beneficial allele (Lange and Pool 2016), for which SNP F_{ST} may be more sensitive. Here, we only found enrichment of window F_{ST} outliers in cis-effect genes for the Med pair. Interestingly, a previous genomic study on these populations found a stronger signal of parallel change for SNP F_{ST} than for window F_{ST} genome-wide (Pool et al. 2017). In light of the lack of elevated SNP F_{ST} among our cis-regulatory P_{ST} outliers, the previously-observed population genetic parallelism may primarily reflect changes other than the cis-regulatory events identified from our whole-organism RNAseq data.

Methods and Materials

Ecologically and phenotypically differentiated populations

The three *Drosophila melanogaster* cold-warm population pairs used in this study, France-Egypt (Med), Ethiopia (Eth) and South Africa (SAf), were described in previous publications (Pool et al. 2012; Lack et al. 2015; Pool 2017). The three cold derived populations have evolved increased cold tolerance in parallel. A previous study has shown that female adults from the cold populations were more likely to recover after 96 hours at 4 °C than the respective warm populations (Pool et al. 2017). Here to confirm increased cold tolerance for the cold populations for egg-to-adult survival, we selected three strains from each of the FR, EF and SD populations as well as from the ancestral warm population ZI as control.

Developmental success was assayed at 15 °C as the cold environment and 25 °C as the warm control environment. 40 mated female flies were allowed to lay eggs in a half pint glass milk bottle with a standard medium at room temperature for 15 hours. Each strain had ~8 bottles. After the flies were removed and the number of eggs were counted, about half of the bottles were incubated at warm environment and the other half were incubated at cold environment. The numbers of adult flies emerged from each bottle were counted after 14 days and 42 days from warm and cold environments respectively. Bottles with more adults than recorded eggs were scored as 100% survival. Developmental success for each strain was measured as the average emergence proportion among bottles, which is the number of emerged adults divided by the number of eggs. To determine

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

436 significance, unpaired t-tests between each cold population and the ZI population were performed for
437 both temperature conditions.
438
439 *RNA sample collections and sequencing*
440 Within each population of the three warm-cold pairs (six populations in total), we selected 16 strains
441 and assigned them into eight crosses. Before the crossing, all the strains had been inbred for eight
442 generations. The criterion for choosing parental strains for a cross was based on minimal genomic
443 regions of overlapping heterozygosity. Among the strains chosen within each population, we used
444 similar criteria to select four strains to perform crosses between the warm and the respective cold
445 populations. Two of the four strains were used as the maternal lines and the other two were used as
446 paternal lines in the between-population crosses. One cross between SD and SP populations was lost.
447 We also collected adult female samples from the parental inbred lines used in the crosses.
448
449 All the flies were reared at 15°C, which approximated the derived cold condition. 20 virgin females
450 and 20 males were collected from maternal and paternal lines respectively for each cross and allowed
451 to mate and lay eggs for a week in half pint bottles. Each bottle contained standard *Drosophila* medium
452 (containing molasses, cornmeal, yeast, agar, and antimicrobial agents). For the within-population
453 crosses, samples at three developmental stages were collected: larva, pupa and female adult. Third-
454 instar larvae were collected on the surface of the medium. For pupa, new yellow pupae were collected
455 within one day of pupation. For adult, female flies were collected 4-5 days after eclosion. For samples
456 from between-population crosses and parental lines, only female adults were collected. All the
457 samples were shock-frozen in liquid nitrogen immediately after collection.
458
459 Approximate 50 larvae or 50 pupae or 30 female adults were used for RNA extraction for each
460 sample. Total mRNA was extracted using the Magnetic mRNA Isolation Kit (New England Biolabs,
461 Ipswich, MA) and RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany). Strand-specific
462 libraries were prepared using the NEBNext mRNA Library Prep Reagent Set for Illumina. Libraries
463 were sized selected for approximately 150 bp inserts using AMPureXP beads (Beckman Coulter, CA,
464 USA). The libraries were quantified using Bioanalyzer and manually multiplexed for sequencing. All
465 libraries were sequenced on a HiSeq2500 (V4) with 2×75bp paired-end in two flow cells.
466
467 *Quantifying gene expression and exon usage frequency*

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

468 The paired-end sequence reads for the within-population cross samples were mapped to the
469 transcribed regions annotated in *D. melanogaster* (release 6, BDGP6.84) using STAR with parameters
470 from ENCODE3's STAR-RSEM pipeline (Li and Dewey 2011; Dobin et al. 2013). For gene
471 expression, the numbers of reads mapped to each gene were quantified using RSEM (Li and Dewey
472 2011). Reads mapped to the rRNA were excluded in the analysis. The expression abundance for
473 each gene was standardized by the numbers of reads mapped to the total transcriptome of the
474 sample.
475
476 To quantify exon usage, we used Leafcutter (Li et al. 2018) to estimate the excision frequencies of
477 alternative introns. This phenotype summarizes different major splicing events, including skipped
478 exons, 5' and 3' alternative splice-site usage, intron retention. Leafcutter took the alignment files
479 generated by STAR as input to quantify the usage of each intron. Then Leafcutter formed clusters
480 that contain all overlapping introns that shared a donor or accept splice site. The default parameters
481 were used: ≥ 50 reads supporting each intron cluster and ≤ 500kb for introns length. The exon usage
482 frequency is the number of intron excision events divided by the total events per cluster. It is worth
483 noting that Leafcutter only detects exon-exon junction usage and it is unable to quantify 5' and 3'
484 end usage and intron retention (Alasoo et al. 2018), which were not examined here.
485
486 *Identifying outliers in gene expression and intron usage differentiation*
487 To identify candidate genes under differential evolution between the warm and cold populations in
488 each pair, we first controlled for the potential transcriptome skew caused by very highly expressed
489 genes. For each expressed gene, we calculated the average expression of the cold samples ($AvgExp_{cold}$)
490 and that of the warm samples ($AvgExp_{warm}$). Then we obtained the median of the ratio of
491 $AvgExp_{cold}/AvgExp_{warm}$ across all expressed genes for the population pair. Gene expression for the warm
492 samples was normalized by multiplying this median before subsequent analysis. This correction was
493 designed to avoid a scenario in which either the cold population or the warm population had
494 important expression changes in one or more highly expressed genes that caused the relative
495 expression of all other genes to shift, even if their absolute expression level did not.
496
497 We used P_{ST} statistics to quantify gene expression divergence between cold and warm populations in
498 each population pair:

499
$$P_{ST} = \frac{V_{between}}{V_{between} + 2V_{within}}$$

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where $V_{between}$ is between-populations variance for expression abundance, V_{within} is the average variance for expression abundance within populations. Although both within- and between-population components of variance can be confounded by the environmental variance, P_{ST} is still a useful statistic to quantify phenotypic differentiation (Merila 1997; Brommer 2011; Leinonen et al. 2013). Here, environmental variance should be reduced by the common laboratory environment. To reduce sampling variance before calculating P_{ST}, for each gene, we required the total mapped reads across all 48 within-population samples to exceed 200 for a given developmental stage. Then for each population/stage, we excluded the crosses/samples with the highest and lowest gene expression for each gene (to avoid high P_{ST} values being driven by single anomalous values), resulting in six samples per population/stage. The P_{ST} quantile based on data excluding extreme samples is concordant with the P_{ST} quantile calculated using all the crosses for most cases (Fig. S3).

We chose the above P_{ST}-based approach instead of simply testing for differential expression in part because our within-population samples reflect real variation as opposed to technical replicates. Also, many alternative methods make assumptions about the data (e.g., negative binomial distribution for transcript counts) which are difficult to apply to splicing, even if they hold for expression. P_{ST} and the population genetic index F_{ST} are under the same theoretical framework, and are often directly compared to search for evidence of adaptive trait differentiation. However, environmental and measurement variance will downwardly bias P_{ST}, making targets of local adaptation less likely to reach a threshold defined by genome-wide high F_{ST} outliers. Hence, in this study we simply focus on the highest quantiles of P_{ST} for a given trait/population comparison, as detailed below.

As with gene expression, we used P_{ST} to estimate the intron usage differentiation between cold and warm populations, with $V_{between}$ as the between-populations variance for a given intron's usage frequency, V_{within} as the average within populations variance for intron usage frequency. Before calculating the P_{ST}, for each exon-exon junction, we summed the intron excision events (n_i) and the alternative events (n_j) of the cluster across all samples in a developmental stage. The minimum for both types of event had to be at least 5 ($n \in [n_i, n_j] \geq 5$) for the exon-exon junction to be included in subsequent analysis. Then for each exon-exon junction, we excluded the sample with highest and lowest intron usage in a population/stage and calculated P_{ST}.

Examining co-directional change for outliers shared between population pairs

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

For gene expression differentiation, we used the upper 5% quantile of P_{ST} as outlier cutoff to identify candidate genes potentially under geographically differential selection. To study the degree of parallel evolution in gene expression, we identified outlier genes shared between two population pairs and showing consistent changes in the cold populations relative to the warm ones (co-directional). Whether the number of shared outliers with co-directional change was significantly greater than expected by chance from the total shared genes between population pairs was determined by a one-tailed binomial test. The statistics here and those below assume the expression changes are independent among genes/introns, which is not always the case (genes can interact with each other via regulatory networks).

The second approach used to examine parallelism of gene expression evolution was to focus on the outlier genes for a specific population pair (*outlier pair*) and examine whether the expression changes in other pair (*directional pair*) follow the same directions. If cold adaptation causes similar evolution in gene expression, those genes in the *directional pair* should have changes in the same directions as the *outlier pair*. Each of the pairwise population combinations had two comparisons; a population pair was assigned as the *outlier pair* in one comparison and as the *directional pair* in the other comparison. To generate a control set of genes for the null expectation of co-directional change proportion, we identified genes in the bottom 50% quantile for P_{ST} in both the o*utlier pair* and the responding *directional pair*. We tested whether the proportion of co-directional change is higher in the outliers than that in the control using the Chi-squared Test.

To identify exon usage outliers, a cutoff of the upper 5% P_{ST} is used. If multiple exon junctions had P_{ST} pass the top 5% cutoff, only the exon junction with the highest P_{ST} would be kept as an outlier to control for nonindependence. Because the numbers of shared exon usage outliers in both population pairs are small (<10), we only performed the second type of analysis studying the proportion of co-directional changes between *outlier pair* and *directional pair* for the top 5% exon usage events. We identified exon usage events in the bottom 50% P_{ST} in both population pairs as control.

GO enrichment test for P_{ST} outlier genes

The Gene Ontology enrichment tests were performed using the R package "clusterProfiler" (Yu et al. 2012) based on the fly genome annotation (Carlson 2018). The types of GO terms being tests contained all three sub-Ontologies: Biological Process (BP), Cellular Component (CC) and Molecular Function (MF). Selection of overrepresented GO terms was based on adjusted p-value < 0.1 using

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

565 "BH" method (Benjamini and Hochberg 1995) for each sub-Ontology. For gene expression, the
566 upper 5% P_{ST} outliers for each population pair were tested for GO enrichment. To determinate
567 whether the shared significant GO terms between pairs were more than expected by chance, we
568 randomly sampled the same numbers of genes as the outliers and performed the GO test for both
569 pairs and identified the shared significant GO terms between pairs. We repeated the process 1000
570 times to get a set of numbers for the shared significant GO terms and compared to the actual number
571 of shared significant GO terms to get a permuted p-value.
572
573 To access the functional categories of the differential intron usage, we calculate the quantile of P_{ST} for
574 each exon usage. To rank the differentiation for a gene, we used the highest quantile (the most
575 extreme differentiation) among the exon usages within the gene as the gene quantile (q_{gene}). To
576 account for the multiple testing of the exon usages for a gene, the adjusted total numbers of testing is
577 calculated as $n_{sum} = \sum_{i=1}^{i=j}(n_i - 1)$, where n_i is the number of testing for a cluster and j is the number
578 of clusters for the gene. Then adjust gene quantile is q'_{gene} = 1- (1- q_{gene}) x n_{sum}. The upper 5% q'_{gene} was
579 used to identify the most differentiated genes for intron usage and they were tested for GO
580 enrichment as described above.
581
582 *Cis- and trans-effects of regulatory divergence*
583 To study the contributions of cis- and trans-regulatory effects on expression and exon usage
584 divergent, we focused our analysis on the upper 5% P_{ST} outliers for gene expression/exon usage. For
585 each gene/exon junction in each population pair, we selected a representative cross showing the
586 greatest difference between parental strains for this analysis. In addition, this difference needed to be
587 larger than the average difference between the cold and warm populations for its pair.
588
589 To study allele-specific expression/exon junction, we obtained the genomic sequences of the two
590 parental strains aligned separately to the FlyBase *D. melanogaster* 5.77 assembly (Lack et al. 2015;
591 2016). The SNP calling from the reference genome was done by samtools (Li et al. 2009). To avoid
592 mapping bias for the RNAseq reads (Degner et al. 2009; Stevenson et al. 2013), we updated the
593 reference based on the SNPs for the two parental stains by masking the SNPs as "N". The F1 female
594 adult RNA-seq reads were mapped to the updated reference using STAR with options: --chimFilter
595 None --outFilterMultimapNmax 1 (Dobin et al. 2013). Because of the high level of heterozygosity
596 within our inbred lines (Lack et al. 2015), we used a parental ancestry proportion statistic (*f*) to study

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the allele-specific expression instead of focusing on fixed difference between parental strains. The parental proportion in gene expression level/exon usage in the F1 RNA-seq sample was estimated as

$$f = (p_{F1} - p_w)/(p_c - p_w)$$

where p_{F1} is the allele frequency in the RNA reads for the F1 sample, p_c and p_w are the allele frequency in the genomic reads for the cold- and warm-adapted parental lines respectively. SNPs were filtered with read counts ≥ 10 in the F1 RNA-seq sample and the parental samples as well as parental frequency difference $|p_c - p_w| \geq 0.25$. The parental proportion for each candidate gene was the average f for all sites located in the gene ($\bar{f}$).

We tested two null hypotheses corresponding to cis only and trans only regulatory differences. Under the null hypothesis that cis-regulatory effects are absent, the $\bar{f}$ is expected to be near 0.5 because the cold parental strain contributes half of the alleles to F1 offspring, and alleles from different parents express similarly in these F1s (Cowles et al. 2002; McManus 2010; Meiklejohn et al. 2014). Under the null hypothesis that trans-regulatory effects are absent, $\bar{f}$ is expected to approximate the ratio of the cold parental strain expression to the total expression of both parental strains (Wittkopp et al. 2004): $r_{F0} = E_c/(E_c + E_w)$. However, sampling effects can cause $\bar{f}$ to deviate from the null expectations.

We accounted for different types of uncertainty on estimating f. The first is the uncertainty on estimating parental strain frequencies p_c and p_w from the genomic data. For each SNP used in the calculation, we resampled 60 alleles based on the estimated allele frequency, representing the 30 individuals used for genome sequencing (Lack et al. 2015). Then we sampled reads by drawing with replacement among the resampled 60 alleles until we reached the observed read depth of the site to calculate the p_c' and p_w'. To account for the measurement uncertainty in F1 expression, we sampled with replacement for the F1 reads mapped to each gene (based on p_c' and p_w') until we reached the numbers of reads mapped to the gene. Then we recalculated the p_{F1}' for each SNPs and together with p_c' and p_w' to calculate the $\bar{f}$' for each gene. We repeated the above process 1000 times to get a distribution of $\bar{f}$'. A 95% confidence interval of the distribution not overlapping with 0.5 suggested the existence of a cis-effect.

However, there is another type of sampling effect if the regulatory variants are not fixed different between parental strains. For example, one strain may be heterozygous for a causative regulatory variant, which might be located outside the exons and hence absent from the RNAseq data. The null

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

hypothesis for inferring a cis-effect is that only trans-effects are present and the $\bar{f}$ is 0.5. Sampling of trans-regulatory polymorphism does not affect the null expectation since trans-effect influences both target alleles similarly. However, the sampling of cis-regulatory polymorphism affects the null expectation for trans-effect because the F1 expression proportion $\bar{f}$ can deviate from the parental expression ratio r_{F0}, potentially causing false positive inferences of trans-effect. Although there is no information about the frequency and effect size for the cis-regulatory mutations, we chose simple assumptions about them to make a relatively conservative approach for inferring trans-effects. We assumed that the frequency of the cis-regulatory allele is 0.5 in the cold-adapted strain (heterozygous, *Aa*) and 0 in the warm adapted strain (homozygous, *aa*). This simplest polymorphism condition maximized the sampling effect within the cold strain. Then we assigned the effect size for the *a* cis-regulatory allele as the expression level of the warm adapted strain (E_w). The effect size for the *A* cis-regulatory allele is $2E_c$-E_w, where the E_c is the expression level of the cold adapted strain. Then we sampled 30 alleles randomly from *Aa* with replacement to create diploid individuals and calculated the average expression for the sampled individuals from the cold strain E_c'. The updated r_{F0}' is calculated as $E_c'/(E_c' + E_w)$. The sampling and calculation were repeated 1000 times. Each time the r_{F0}' is paired with a $\bar{f}$' described above to calculate the difference $D' = \bar{f}' - r_{F0}'$. A 95% confidence interval of D' not overlapping with 0 suggested the existence of a trans-effect.

Based on the tests above, the set of candidate genes were classified into categories (McManus 2010; Schaefke et al. 2013; Chen et al. 2015) including no significant cis- or trans-effect, cis only, and trans only. For genes showing both cis- and trans-effects, we further classified them based on whether these two effects favored expression of the same (co-directional) or different parental allele (anti-directional). For exon usage differentiation, we applied a similar approach to classified the differentiated exons into the five categories, accounting for different sampling effects and measurement errors. Instead of analyzing expression level of the parental strains (E), we analyzed their exon usage frequency for the sets of outlier exon junctions.

For the P_{ST} outlier introns identified as cis only or both cis- and trans-effect, we hypothesized that causative cis-regulatory elements may show elevated allele frequency differentiation between the warm and cold populations. For expression abundance, the majority of cis-regulatory SNPs are located within 2kb upstream of the transcription start site and downstream of the transcription end site (Massouras et al. 2012). Therefore, we used the interval from 2kb upstream to 2kb downstream as

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the focal region of a gene for this analysis. We calculated window F_{ST} and SNP F_{ST} using sequenced genomes from *Drosophila* Genome Nexus (Lack et al. 2015 & 2016). For window F_{ST}, the division of windows within a gene region is based on 250 non-singleton variable sites per window in the ZI population (Pool et al. 2017). The highest F_{ST} for the windows overlapping the focal region was assigned as its F_{ST_winmax}. To determinate the statistical significance of F_{ST_winmax}, we calculated F_{ST_winmax} for all other blocks of the same number of windows along the same chromosome arm where cross-over rates were above 0.5cM/Mb (Comeron et al. 2012), but excluding those within 10 windows of the focal region. The specific non-low recombination regions are: 2.3–21.4 Mb for the X chromosome, 0.5–17.5 Mb for arm 2L, 5.2–20.8 Mb for arm 2R, 0.6–17.7 Mb for arm 3L, and 6.9–26.6 Mb for arm 3R. SNP F_{ST} was calculated for all sites within the focal region and the highest value (F_{ST_SNPmax}) was thus obtained for the focal gene. Analogous to our F_{ST_winmax} permutation, we also calculated F_{ST_SNPmax} for permuted regions with the same number of SNPs as the focal region, along the non-low cross-over rate region on the same chromosome arm. For both F_{ST_winmax} and F_{ST_SNPmax}, we then focused on regions in the upper 10% quantile of permuted values for further analysis. We tested whether the proportion of genes with high F_{ST} is higher in the cis-effect genes than that in control non-outliers using the Fisher's Exact Test because of the low counts.

We also identified genes with significant cis-effects shared in two population pairs and examined whether the effects favored expression of the same cold or warm parental alleles (consistent cis-effect). Also, we tested whether the shared cis-effect genes also show elevated population genetic differentiation in the two pairs. We obtained "Population Branch Excess" statistic (*PBE*) specific for cold populations for SNPs from Pool et al. 2017. We used ±2kb around the gene regions to look for any shared cis-effect genes containing SNPs with high *PBE* statistic for cold population in both pairs (*PBE* quantile < 0.05 in both pairs).

For exon usage, because the cis-regulation is largely contributed by the splice sites (Kurmangaliyev et al. 2015), we calculated the F_{ST} value for the splice sites, which are located within ±15 base pair around the two intron/exon boundaries. The maximum F_{ST} among the splice sites for each intron is chosen as the SNP F_{ST} for the focal intron. We compared the F_{ST_SNPmax} of the cis outlier introns and the non-outlier introns to see whether the cis outlier introns showed elevated F_{ST} based on Wilcoxon signed-rank test. To examine the potential function of splicing differentiation, genes containing high

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SNP F_{ST} (upper 15% quantile of F_{ST_SNPmax}) flanking cis-regulated introns were identified as candidate genes.

Acknowledgements

We thank Colin Dewey for helpful discussions and the UW-Madison Center for High Throughput Computing (CHTC) for cluster usage. This work was funded by NSF DEB grant 1754745 to JEP and by NIH NIGMS grant F32GM106594 to JBL.

Reference

Adrion JR, Hahn MW, Cooper BS. 2015. Revisiting classic clines in *Drosophila melanogaster* in the age of genomics. Trends Genet 8:434–444.

Akhouayri I, Turc C, Royet J, Charroux B. 2011. Toll-8/Tollo negatively regulates antimicrobial response in the *Drosophila* respiratory epithelium. PLoS Pathog 7:e1002319.

Albert FW, Bloom JS, Siegel J, Day L, Kruglyak L. 2017. Genetics of trans-regulatory variation in gene expression. eLife 100:371–44.

Arthaud L, Ben Rokia-Mille S, Raad H, Dombrovsky A, Prevost N, et al. 2011. Trade-off between toxicity and signal detection orchestrated by frequency- and density-dependent genes. PLoS One 6: e19805.

Bader DM, Wilkening S, Lin SG, Tekkedil MM, Dietrich K, et al. 2015. Negative feedback buffers effects of regulatory variants. Mol Syst Biol 11:785.

Ballard SL, Miller DL, Ganetzky B. 2014. Retrograde neurotrophin signaling through Tollo regulates synaptic growth in *Drosophila*. J Cell Biol 204:1157–1172.

Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, et al. 2012. The evolutionary landscape of alternative splicing in vertebrate species. Science 338:1587–93.

Bastide H, Yassin A, Johanning EJ, Pool JE. 2014. Pigmentation in *Drosophila melanogaster* reaches its maximum in Ethiopia and correlates most strongly with ultra-violet radiation in sub-Saharan Africa. BMC Evol Biol 14:222–14.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate—a practical and powerful approach to multiple testing. J Royal Stat Soc B 57:289–300.

Bergland AO, Tobler R, González J, Schmidt P, Petrov D. 2016. Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in *Drosophila* melanogaster. Mol Ecol 25:1157–1174.

Božičević V, Hutter S, Stephan W, Wollstein A. 2016. Population genetic evidence for cold adaptation in European *Drosophila melanogaster* populations. Mol Ecol 25:1175–1191.

Brommer JE. 2011. Whither P_{ST}? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. J Evol Biol 24:1160–1168.

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

728 Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, et al. 2014. Diversity and dynamics
729 of the *Drosophila* transcriptome. Nature. 7515:393–9.

730 Catalan A, Hutter S, Parsch J. 2012. Population and sex differences in *Drosophila melanogaster*
731 brain gene expression. BMC Genomics 13:1–12.

732 Chakraborty M, VanKuren NW, Zhao R, Zhang X, Kalsow S, Emerson JJ. 2017. Hidden genetic
733 variation shapes the structure of functional elements in *Drosophila*. Nat Genet 50:20–25.

734 Chen J, Nolte V, Schlötterer C. 2015. Temperature Stress Mediates Decanalization and Dominance
735 of Gene Expression in *Drosophila melanogaster*. PLoS Genet 11:e1004883.

736 Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in *Drosophila*
737 *melanogaster*. PLoS Genet 8:e1002905.

738 Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and
739 convergence in natural populations. Proc Royal Soc B 279:5039–5047.

740 Coolon JD, McManus CJ, Stevenson KR, Graveley BR, Wittkopp PJ. 2014. Tempo and mode of
741 regulatory evolution in Drosophila. Genome Res 24:797–808.

742 Cowles CR, Hirschhorn JN, Altshuler D, Lander ES. 2002. Detection of regulatory variation in
743 mouse genes. Nat Genet 32:432–437

744 Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkadori E, Gilad Y, Pritchard JK. 2009. Effect of read-
745 mapping biases on detecting allele-specific expression from RNA-sequencing data.
746 Bioinformatics 25:3207–3212.

747 Denby CM, Im JH, Yu RC, Pesce CG, Brem RB. 2012. Negative feedback confers mutational
748 robustness in yeast transcription factor regulation. Proc Natl Acad Sci USA 109:3874–3878.

749 Deshpande G, Calhoun G& Schedl P. 2005. Drosophila Argonaute-2 is required early in
750 embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division,
751 nuclear migration, and germ-cell formation. Gene Dev 19:1680–1685.

752 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR.
753 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21.

754 Duneau D, Sun H, Revah J, San Miguel K, Kunerth HD, Caldas IV, Messer PW, Scott JG, Buchon
755 N. 2018. Signatures of insecticide selection in the genome of *Drosophila melanogaster*. G3
756 (Bethesda) 8:3469–3480.

757 Emerson JJ. 2019. Inferring compensatory evolution of cis- and trans-regulatory variation. Trends
758 Genet 35:1–3.

759 Fabian DK, Kapun M, Nolte V, Kofler R, Schmidt PS, Schlötterer C, Flatt T. 2012. Genome-wide
760 patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North
761 America. Mol Ecol 21:4748–4769.

762 Fear JM, Leon-Novelo LG, Morse AM, Gerken AR, Van Lehmann K, Tower J, Nuzhdin SV,

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

McIntyre LM. 2016. Buffering of genetic regulatory networks in *Drosophila melanogaster*. Genetics 203:1177–1190.

Francis VA, Zorzano A, Teleman AA. 2010. dDOR is an EcR coactivator that forms a feed-forward loop connecting insulin and ecdysone signaling. Curr Biol 20:1799–1808.

Fraser HB. 2019. Improving estimates of compensatory cis–trans regulatory divergence. Trends Genet 35:3–5.

Gamazon ER, Stranger BE. 2014. Genomics of alternative splicing: evolution, development and pathophysiology Hum Genet 133:679-687

Gibilisco L, Zhou Q, Mahajan S, Bachtrog D. 2016. Alternative splicing within and between *Drosophila* species, sexes, tissues, and developmental stages. PLoS Genet 12: e1006464.

Gilad Y, Rifkin SA, Pritchard JK. 2008. Revealing the architecture of gene regulation: the promise of eQTL studies. Trends Genet 24:408–415.

Green L, Battlay P, Fournier-Level A, Good RT, Robin C. 2019. Cis- and trans-acting variants contribute to survivorship in a naïve *Drosophila melanogaster* population exposed to ryanoid insecticides. Proc Natl Acad Sci USA 116:10424–10429.

Hanson D, Hu J, Hendry AP, Barrett RDH. 2017. Heritable gene expression differences between lake and stream stickleback include both parallel and antiparallel components. Heredity 119:339–348.

Hart JC, Ellis NA, Eisen MB, Miller CT. 2018. Convergent evolution of 1493 gene expression in two high-toothed stickleback populations. PLoS Genetics 6:e1007443.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 615: 995–1016.

Hoffmann AA, Hallas RJ, Dean JA, Schiffer M. 2003. Low potential for climatic stress adaptation in a rainforest *Drosophila* species. Science 301:100–102.

Hoffmann AA, Weeks AR. 2007. Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temper-ate-tropical clines in *Drosophila melanogaster* from eastern Australia. Genetica 129:133–147.

Huylmans AK, Parsch J. 2014. Population- and sex-biased gene expression in the excretion organs of *Drosophila melanogaster*. G3 (Bethesda). 4:2307–15.

Johnson KG, Tenney AP, Ghose A, Duckworth AM, Higashi ME, Parfitt K, Marcu O, Heslip TR, Marsh JL, Schwarz TL, et al. 2006. The HSPGs Syndecan and Dallylike bind the receptor phosphatase LAR and exert distinct effects on synaptic development. Neuron 49:517–531.

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55–61.

Juneja P, Quinn A, Jiggins FM. 2016. Latitudinal clines in gene expression and cis-regulatory

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

798 element variation in Drosophila melanogaster. BMC Genomics:1–11.

799 Kang J, Kim J, Choi K-W. 2011. Novel cytochrome P450, cyp6a17, is required for temperature
800 preference behavior in *Drosophila*. PLoS ONE 6:e29800.

801 Kitano J, Ishikawa A, Kusakabe M. 2018. Parallel transcriptome evolution in stream threespine
802 sticklebacks. Devel Growth Differ 61:104–113.

803 Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011. Genomic differentiation between
804 temperate and tropical australian populations of *Drosophila melanogaster*. Genetics 187:245–
805 260.

806 Kurmangaliyev YZ, Favorov,AV, Osman NM, Lehmann K-V, Campo D, Salomon MP, et al. 2015.
807 Natural variation of gene models in Drosophila melanogaster. BMC Genomics 16:198

808 Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. 2016. A thousand fly genomes: An
809 expanded Drosophila genome nexus. Mol Biol Evol 33:3308–3313.

810 Lack JB, Monette MJ, Johanning EJ, Sprengelmeyer QD, Pool JE. 2016. Decanalization of wing
811 development accompanied the evolution of large wings in high-altitude *Drosophila*. Proc Natl
812 Acad Sci USA 113:1014–1019.

813 Lai Z, Kane NC, Zou Y, Rieseberg LH. 2008. Natural variation in gene expression between wild and
814 weedy populations of *Helianthus annuus*. Genetics 179:1881–1890.

815 Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL. 2007. Genetic properties influencing the
816 evolvability of gene expression. Science 317:118–121.

817 Lange JD, Pool JE. 2016. A haplotype method detects diverse scenarios of local adaptation from
818 genomic sequence variation. Mole Ecol 25:3081–3100.

819 Leder EH, McCairns RJS, Leinonen T, et al. 2015. The evolution and adaptive potential of
820 transcriptional variation in sticklebacks—signatures of selection and widespread heritability.
821 Mol Biol and Evol 32:674–689.

822 Leinonen T, McCairns RJS, O'Hara RB, Merilä J. 2013. Q_{ST}–F_{ST} comparisons: evolutionary and
823 ecological insights from genomic heterogeneity. Nat Rev Genet. 14:179–190.

824 Lemmon ZH, Bukowski R, Sun Q, Doebley JF. 2014. The role of cis regulatory evolution in maize
825 domestication. PLoS Genet 10:e1004745–15.

826 Li B, Dewey Colin N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or
827 without a reference genome. BMC Bioinform 12:323

828 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. 2009. The Sequence alignment/map (SAM)
829 format and SAMtools. Bioinformatics 25:2078-2079.

830 Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, Pritchard JK. 2017.
831 Annotation-free quantification of RNA splicing using LeafCutter. Nat Genet 50:151–158.

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

832 Losos JB. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–1840.
833
834 Machado HE, Bergland AO, O'Brien KR, Behrman EL, Schmidt PS, Petrov DA. 2016. Comparative
835 population genomics of latitudinal variation in *Drosophila simulans* and *Drosophila*
836 *melanogaster*. Mol Ecol 25:723–740.

837 Mack KL, Campbell P, Nachman MW. 2016. Gene regulation and speciation in house mice. Genome
838 Res 26:451–461.

839 Massouras A, Waszak SM, Albarca-Aguilera M, Hens K, Holcombe W, Ayroles JF, Dermitzakis ET,
840 Stone EA, Jensen JD, Mackay TFC, et al. 2012. Genomic Variation and Its Impact on Gene
841 Expression in *Drosophila melanogaster*. PLoS Genet 8:e1003055.

842 Mateo L, Ullastres A, González J. 2014. A transposable element insertion confers xenobiotic
843 resistance in *Drosophila*. PLoS Genet 10:e1004560.

844 McGirr JA, Martin CH. 2018. Parallel evolution of gene expression between trophic specialists
845 despite divergent genotypes and morphologies. Evol Lett 2:62–75.

846 McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. 2010. Regulatory
847 divergence in *Drosophila* revealed by mRNA-seq. Genome Res 20:816–825.

848 McManus CJ, Coolon JD, Eipper-Mains J, Wittkopp PJ, Graveley BR. 2014. Evolution of splicing
849 regulatory networks in *Drosophila*. Genome Res 24:786–796.

850 McMillan EA, Longo SM, Smith MD, Broskin S, Lin B, Singh NK, Strochlic TI. 2018. The protein
851 kinase CK2 substrate Jabba modulates lipid metabolism during *Drosophila* oogenesis. J Biol
852 Chem 293:2990–3002.

853 Meiklejohn CD, Coolon JD, Hartl DL, Wittkopp PJ. 2014. The roles of cis- and trans-regulation in
854 the evolution of regulatory incompatibilities and sexually dimorphic gene expression. Genome
855 Res 24:84–95.

856 Merilä J. 1997. Quantitative trait and allozyme divergence in the greenfinch (*Carduelis chloris*,
857 Aves: Fringillidae). Biol J Linn Soc 61:243–266.

858 Nakano Y, Fujitani K, Kurihara J, Ragan J, Usui-Aoki K, Shimoda L, Lukacsovich T, Suzuki K,
859 Sezaki M, Sano Y, et al. 2001. Mutations in the novel membrane protein Spinster interfere with
860 programmed cell death and cause neural degeneration in *Drosophila melanogaster*. Mol Cell
861 Biol 21:3775–3788.

862 Nandamuri SP, Conte MA, Carleton KL. 2018. Multiple trans QTL and one cis-regulatory deletion
863 are associated with the differential expression of cone opsins in African cichlids. BMC
864 Genomics. 19:945.

865 Nayak A, Berry B, Tassetto M, Kunitomi M, Acevedo A, Deng C, Krutchinsky A, Gross J,
866 Antoniewski C, Andino R. 2010. Cricket paralysis virus antagonizes Argonaute 2 to modulate
867 antiviral defense in *Drosophila*. Nat Struct Mol Biol. 17:547–554.

868 Osada N, Miyagi R, Takahashi A. 2017. Cis- and trans-regulatory effects on gene expression in a

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

869 natural population of *Drosophila melanogaster*. Genetics 206:2139–2148.

870 Pitchers W, Pool JE, Dworkin I. 2012. Altitudinal clinal variation in wing size and shape in african
871 *Drosophila melanogaster*: one cline or many? Evolution 67:438–452.

872 Pool JE, Braun DT, Lack JB. 2016. Parallel evolution of cold tolerance within *Drosophila*
873 *melanogaster*. Mol Biol Evol 34:349–360.

874 Pool JE, Corbett-Detig RB, Sugino RP, Stevens KA, Cardeno CM, Crepeau MW, Duchen P,
875 Emerson JJ, Saelao P, Begun DJ, et al. 2012. Population genomics of sub-Saharan *Drosophila*
876 *melanogaster*: African diversity and non-African admixture. PLoS Genet 8:e1003080–24.

877 Pool JE. 2015. The Mosaic Ancestry of the Drosophila Genetic Reference Panel and the *D.*
878 *melanogaster* reference genome reveals a network of epistatic fitness interactions. Mol Biol Evol
879 32:3236–3251.

880 Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel geographic variation
881 in *Drosophila melanogaster*. Genetics 197:361–373.

882 Rong et al. 2011. Spinster is required for autophagic lysosome reformation and mTOR reactivation
883 following starvation. Proc Natl Acad Sci USA 108:7826–7831.

884 Rosenblum EB, Parent CE, Brandt EE. 2014. The molecular basis of phenotypic convergence. Annu
885 Rev Ecol Evol Syst 45:203–226.

886 Sackton TB, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds.
887 Science 364:74–78.

888 Schaefke B, Emerson JJ, Wang T-Y, Lu M-YJ, Hsieh L-C, Li W-H. 2013. Inheritance of gene
889 expression level and selective constraints on trans- and cis-regulatory changes in yeast. Mol Biol
890 Evol 30:2121–2133.

891 Schluter D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, New York.

892 Schmidt PS, Paaby AB. 2008. Reproductive diapause and life-history clines in North American
893 populations of *Drosophila melanogaster*. Evolution 62:1204–1215.

894 Seppo A. 2003. Induction of neuron-specific glycosylation by Tollo/Toll-8, a *Drosophila* Toll-like
895 receptor expressed in non-neural cells. Development 130:1439–1448.

896 Signor S, Nuzhdin S. 2018. Dynamic changes in gene expression and alternative splicing mediate the
897 response to acute alcohol exposure in *Drosophila melanogaster*. Heredity 121:342-360.

898 Singh P, Börger C, More H, Sturmbauer C. 2017. The role of alternative splicing and differential
899 gene expression in cichlid adaptive radiation. Genome Biol Evol 9:2764–2781.

900 Smith CCR, Tittes S, Mendieta JP, Collier-Zans E, Rowe HC, Rieseberg LH, Kane NC. 2018.
901 Genetics of alternative splicing evolution during sunflower domestication. Proc Natl Acad Sci
902 USA 115:6768-6773.

bioRxiv preprint doi: https://doi.org/10.1101/795716; this version posted October 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

903 Sprengelmeyer QD, Mansourian S, Lange JD, Matute DR, Cooper BS, Jirle EV, Stensmyr MC, Pool
904 JE. 2018. Discovery of *Drosophila melanogaster* from Wild African Environments and Genomic
905 Insights into Species History. biorxiv 45:1153–13.

906 St Pierre SE, Ponting L, Stefancsik R, McQuilton P, the FlyBase Consortium. 2013. FlyBase 102—
907 advanced approaches to interrogating FlyBase. Nucleic Acids Res 42:D780–D788.

908 Stern DL. 2013. The genetic causes of convergent evolution. Nat Rev Genet 14:751–764.

909 Stevenson KR, Coolon JD, Wittkopp PJ. 2013. Sources of bias in measures of allele-specific
910 expression derived from RNA-seq data aligned to a single reference genome. BMC Genomics
911 14:536.

912 Venables JP, Tazi J, Juge F. 2011. Regulated functional alternative splicing in *Drosophila*. Nucleic
913 Acids Res 40:1–10.

914 Verta JP, Jones FC. 2019. Predominance of cis-regulatory changes in parallel expression divergence
915 of sticklebacks. eLife. 8:e43785.

916 Wittkopp PJ, Kalay G. 2011. Cis-regulatory elements: molecular mechanisms and evolutionary
917 processes underlying divergence. Nat Rev Genet 13:59–69.

918 Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation.
919 Nature 430:85–88.

920 Wittkopp PJ, Haerum BK, Clark AG. 2008. Regulatory changes underlying expression differences
921 within and between *Drosophila* species. Nat Genet 403:346–350.

922 Yagi Y, Nishida Y, Ip YT. 2010. Functional analysis of Toll-related genes in *Drosophila*. Dev
923 Growth Differ 52:771–783.

924 Yu G, Wang L, Han Y, He Q. 2012. clusterProfiler: an R package for comparing biological themes
925 among gene clusters. OMICS 16:284-287.

926 Zhao L, Begun DJ. 2017. Genomics of parallel adaptation at two timescales in *Drosophila*. PLoS
927 Genet 13:e1007016.

928 Zhao L, Wit J, Svetec N, Begun DJ. 2015. Parallel gene expression differences between low and high
929 latitude populations of *Drosophila melanogaster* and *D. simulans*. PLoS Genet 11:e1005184.

930 Zhang J, Liu M, Su Y, Du J, Zhu AJ. 2012. A targeted in vivo RNAi screen reveals deubiquitinases
931 as new regulators of Notch signaling. G3 (Bethesda) 2:1563–1575.

932