The effect of common inversion polymorphisms \textit{In}(2L)t and \textit{In}(3R)Mo on patterns of transcriptional variation in \textit{Drosophila melanogaster}.

Erik Lavington\textsuperscript{1}, Andrew D. Kern\textsuperscript{1,2}

1. Department of Genetics, Rutgers University, Piscataway, New Jersey

2. Human Genetics Institute of New Jersey, Rutgers University, Piscataway, New Jersey
Abstract

Chromosomal inversions are an ubiquitous feature of genetic variation. Theoretical models describe several mechanisms by which inversions can drive adaptation and be maintained as polymorphisms. While inversions have been shown previously to be under selection, or contain genetic variation under selection, the specific phenotypic consequences of inversions leading to their maintenance remain unclear. Here we use genomic sequence and expression data from the Drosophila Genetic Reference Panel to explore the effects of two cosmopolitan inversions, \textit{In(2L)t} and \textit{In(3R)Mo}, on patterns of transcriptional variation. We demonstrate that each inversion has a significant effect on transcript abundance for hundreds of genes across the genome. Inversion affected loci (IAL) appear both within inversions as well as on unlinked chromosomes. Importantly, IAL do not appear to be influenced by the previously reported genome-wide expression correlation structure. We found that five genes involved with sterol uptake, four of which are Niemann-Pick Type 2 orthologs, are upregulated in flies with \textit{In(3R)Mo} but do not have SNPs in LD with the inversion. We speculate that this upregulation is driven by genetic variation in \textit{mod(mdg4)} that is in LD with \textit{In(3R)Mo}. We find that there is little evidence for regional or position effect of inversions on gene expression at the chromosomal level but do find evidence for the distal breakpoint of \textit{In(3R)Mo} interrupting one gene and possibly disassociating the two flanking genes from regulatory elements.

Introduction

Chromosomal inversions, in which a portion of linear DNA sequence is flipped in its orientation, are a common member of the menagerie of DNA polymorphisms, and have been found in diverse organismal populations such as humans, plants, and fruit flies(Krimbas and Powell 1992; Kidd \textit{et al.} 2010; Lowry and Willis 2010). In many cases, large chromosomal inversions have profound impacts on phenotype and disease(Feuk 2010). For instance recurrent inversions are responsible for an estimated 43% of
hemophilia A cases (Lakich et al. 1993). Inversions can also have beneficial effects. A 900kb inversion on human chromosome 17 (q21.31) has been shown to be associated with higher female fecundity in the Icelandic population (Stefansson et al. 2005). In populations of the malaria vector An. gambiae a large chromosomal inversion on chromosome 2L (2La) is associated with desiccation resistance and thus segregates at high frequencies in arid environments (Fouet et al. 2012). These examples are the very tip of the iceberg—inversion polymorphisms have been implicated in numerous phenotypic differences among a host of organisms, however little is known about the mechanisms by which inversions confer their phenotypic effects.

Perhaps the single best studied inversions are those from Drosophila, in part made famous by the pioneering work of Dobzhansky (Dobzhansky and Sturtevant 1938). Dobzhansky focused much attention on spatial and temporal variation in frequency of large inversions of D. pseudoobscura and showed in broad strokes that clear fitness differences were responsible for the regular patterns of frequency change observed. These findings in turn spurred a large body of population genetics theory to explain the establishment and selective persistence of inversions in natural populations (Levene and Dobzhansky 1958; Fraser et al. 1966; Anderson et al. 1967; Tobari and Kojima 1967). As postulated by Sturtevant (1921), crossover suppression induced in inversion heterozygotes can mean that a single adaptive allele within an inversion may suffice for the selective invasion of that rearrangement (Haldane 1937). Such lowered levels of recombination and attendant increases in linkage disequilibrium (LD) could thus present the opportunity for subsequent coadaptation of multiple genes near inversion breakpoints (Sturtevant and Mather 1938; Dobzhansky 1947). Conversely locally adapted alleles that predate the rearrangement on the same chromosome might aid the establishment of an inversion simply because of the reduction in recombination rates between such loci (Kirkpatrick and Barton 2006).

Further, inversions might have direct fitness effects, for instance by deletion or changes in gene
expression near the inversion breakpoints (Kirkpatrick and Kern 2012). At present we have precious little
information as to the variants responsible for differential fitness effects associated with inversions.

In *Drosophila melanogaster* paracentric inversions spanning several megabases are common and have
been found in populations across the globe (Stalker 1976, 1980; Knibb *et al.* 1981; Sezgin *et al.* 2004;
Anderson *et al.* 2005; Umina *et al.* 2005). Much of this segregating inversion polymorphism is associated
with latitudinal clines in *D. melanogaster*, an historically tropical species adapting along tropical-to-
temperate climatic gradients in Australia and North America (Knibb *et al.* 1981; Weeks *et al.* 2002; de
Jong and Bochdanovits 2003; Sezgin *et al.* 2004; Reinhardt *et al.* 2014; Schrider *et al.* 2016). Clinically
varying phenotypes that are associated with inversions include heat resistance, cold tolerance, and body
size (Weeks *et al.* 2002; Anderson *et al.* 2003; de Jong and Bochdanovits 2003). Clinal variation of
inversion frequency in *D. melanogaster* has been shown via population genetic approaches to be due to
selection independent of demography (Reinhardt *et al.* 2014; Kapun *et al.* 2016), though migration has
been suggested to generate these patterns along with local adaptation (Bergland *et al.* 2016). Indeed,
inversions have been observed to have a major effect on several phenotypes that vary between
temperate and tropical populations across several *Drosophila* species and these clines have been stable
since their discovery roughly 80 years ago (see Hoffmann *et al.* 2004 for review). Unfortunately, while
the associations are known, the molecular mechanisms at work determining differential phenotypes as a
result of inversion status are still unknown.

Recent population genomic projects in *D. melanogaster*, such as the Drosophila Genetic Reference Panel
(DGRP) and Drosophila Population Genomics Project (DPGP), have opened the opportunity to study
inversions systematically as these resources have captured segregating inversions from North America
and Africa (Pool *et al.* 2012; Mackay *et al.* 2012; Langley *et al.* 2012; Houle and Márquez 2015). Corbett-
Detig *et al.* (2012) bioinformatically mapped previously unknown breakpoints of several inversions, a
task that was tedious for even single inversions prior to whole genome sequencing (Wesley and Eanes 1994; Andolfatto et al. 1999; Matzkin et al. 2005). For instance Corbett-Detig et al. (2012) discovered the breakpoints associated with numerous inversions and demonstrated the expected increase in LD near inversion breakpoints and elevated differentiation between inverted and standard arrangement chromosomes at the nucleotide level. In parallel with the exponential increase in population genomic resources, large-scale phenotypic association studies of these same genotypes have been accumulating (Mackay et al. 2012; Vonesch et al. 2016; Telonis-Scott et al. 2016). These include numerous phenotypes previously associated with inversion polymorphism such as body size (Weeks et al. 2002) and desiccation resistance (Hoffmann et al. 2005).

A logical place to look for inversion effects that may influence suites of phenotypes would be transcript level variation. Previous findings strongly suggest that inversions could be important drivers of adaptation with gene expression variation as a potential molecular mechanism (Chambers 1991; López-Maury et al. 2008; Fraser 2013). Indeed inversions could affect patterns of transcript variation in a number of ways: 1) genes at or near inversion breakpoints may become disabled or separated from their regulatory apparatus, thus inversions may have direct effects on transcription, 2) increased LD in inversions due to crossover suppression may tie up gene expression Quantitative Trait Loci (eQTL), and thus alternative alleles of the inversion may be associated with differential expression of genes within the inversion (i.e. indirect, cis-eQTL associated with the inversion), 3) eQTL in LD with the inversion might themselves regulate genes outside of the inversion (i.e. indirect, trans-eQTL associated with the inversion), or 4) the large-scale nature of Drosophila inversions may create global changes in the organization of chromatin or nuclear localization of the chromosomes such that genes are differentially regulated between inversion and standard karyotypes. Thus inversions may have a direct effect on global patterns of transcription, and act as trans eQTL themselves. Indeed earlier studies of transcriptional variation in D. melanogaster have hinted at the influence of inversions on genome wide
patterns of gene expression (Ayroles et al. 2009; Massouras et al. 2012; Huang et al. 2015). Here we
address the effect of two cosmopolitan inversions, \textit{In}(2L)t and \textit{In}(3R)Mo, on patterns of transcription by
using whole-genome sequence, gene expression, and inversion call data from the DGRP.

**Methods/Materials**

**Materials:** Processed expression data previously reported in (Ayroles et al. 2009) was downloaded from
ArrayExpress (Kolesnikov et al. 2015). We accepted inversion state calls for each of the DGRP lines
where cytological and bioinformatic inversion calls for \textit{In}(2L)t and \textit{In}(3R)Mo agree and removed lines
from analyses where there was any disagreement (Corbett-Detig et al. 2012a; Huang et al. 2014; Houle
and Márquez 2015). Using the same databases, we removed Individuals from lines likely heterozygous
for \textit{In}(2L)t or \textit{In}(3R)Mo. Expression analyses were performed with 34 lines (136 individuals), with two
lines homozygous for \textit{In}(2L)t and seven lines homozygous for \textit{In}(3R)Mo. LD was calculated with 181 and
197 lines with 19 and 17 inversion lines for \textit{In}(2L)t and \textit{In}(3R)Mo, respectively. To maintain consistency
with Affymetrix library files, dm3/BDGP release 5 genomic coordinates and annotations corresponding
to BDGP version 5.49 were used in conjunction with the Affymetrix Drosophila 2 Release 35 library file
update.

**Methods:**

Statistical analyses were performed in R (R Core Team 2013) using the functions \texttt{lm}, \texttt{anova},
\texttt{quantile}, \texttt{qvalue}, \texttt{phyper}, and \texttt{ggplot2}.

**3' UTR Array Analysis:**

*Correlation structure:* Pairwise gene expression correlation coefficients were calculated by linear
regression on all unique pairwise combinations of probe sets, excepting self-comparisons. Correlation
coefficients reported here as adjusted $r^2$ from the R function (`lm`). Gene expression modules used here were reported in Ayroles et al (2009). Null distributions of uniquely occupied clusters for each inversion were generated by permuting the occupied cluster for each gene 100,000 times and calculating the total unique clusters occupied by inversion affected loci (IAL) for each permutation.

**Inversion effect on expression**: Linear regressions of sex, inversion, and Line effects with expression as the response were performed for each probe set:

$$Y_{ijkl} = \mu + A_i + B_j + C_k + D_{kl} + \varepsilon_{ijkl}$$

for individual expression value, $Y_{ijkl}$, in response to $i^{th}$ Sex, $A_i$, $j^{th}$ In(2L)t state, $B_j$, $k^{th}$ In(3R)Mo state, $C_k$, and the $l^{th}$ Line in the $j^{th}$ In(2L)t state, $D_{kl}$, with $\varepsilon_{ijkl}$ as the error term in Lines. Model testing was performed using (`add1`) and (`drop1`) in R to add interaction terms or remove main effect terms from the above model, respectively. Interaction terms were added one at a time to the main effects and tested for each probe set. An AIC is reported for each model with a lower absolute value being preferred when comparing two models. The effect of adding an interaction term between Sex and In(2L)t or Sex and In(3R)Mo varied by probeset, but the above model was the best fit for 10,082 of the probe sets. Models with and without an interaction term between inversions performed the same for all loci thus we chose the less complex model. Similarly, the above model was the best fit for 12,994 probe sets when compared to dropping any of the main effect terms. We calculated $p$-values of the observed $F$ values as percentiles of the $F$ distributions generated by 10,000 permutations sampling each inversion independently without replacement. Multiple testing correction was performed through Bonferroni corrections to ANOVA derived $p$-values and False Discovery Rate (FDR) by calculating $q$-values using the R package (`qvalue`) with FDR=0.05 (Storey and Tibshirani 2003) on the permutation-derived $p$-values. Proportion of variance explained by each effect was calculated as $\eta^2$. Magnitude and direction of inversion effect was calculated as Cohen’s d (Cohen 1988).
**Functional Annotation Enrichment:** Functional annotation profiling was performed using the g:Profiler online portal of g:GOSt using default settings (Reimand et al. 2016) (version r1622_e84_eg31). Ambiguous 3'UTR probe sets were resolved manually if possible, or ignored if they overlapped transcripts for more than one gene.

**SNP-Inversion LD:** Linkage disequilibrium (LD) was calculated as $r^2 = D'/p_S(1-p_S)p_I(1-p_I)$ for each diallelic SNP S and inversion I, with major allele frequencies $p_S$ and $p_I$, using a custom bash script. For each SNP, significance was calculated as a chi-squared ($\chi^2$) transformation with 1 degree of freedom of $r^2$ as $\chi^2 = Nr^2$, where N is the sample size. Significant LD was defined as a SNP with sample size of at least 60 and $\chi^2 > \text{critical value (p=0.05, d.f.}=1)$ with Bonferroni correction for all SNPs on that chromosome arm (n=967774, $\chi^2 > 29.65329$ and 947970, $\chi^2 > 29.61321$ for chr2L and chr3R, respectively). Genes were considered in significant LD with inversion state if at least one significant SNP was found within the annotated gene region (FlyBase v5.49).

**IAL physical clustering:** To see if IAL were physically clustered within the genome, we examined physical clustering by measuring the coefficient of variance (CV) of distances between genes by chromosome arm. Location and length of each gene was used from FlyBase v5.49. Distance between neighboring genes was calculated as the distance between ends of gene annotated regions of neighboring genes. Distance from the most distal or proximal genes to the distal or proximal endpoint, respectively, was not included. CV was calculated as the standard deviation divided by the mean of the distribution of intergenic distances for each chromosomal arm (Sokal and Rohlf 1995). We defined the intergenic distance between overlapping gene regions as zero. Null distributions of intergenic distances for each chromosome and each inversion were generated by 100,000 random samples, without replacement, of the same number of genes from a chromosome arm as the number of IAL for that chromosome arm and
inversion, then calculating the distances between those genes. Confidence intervals were calculated as 2.5%-97.5% quantiles from the corresponding random sample distribution.

**Transcription factor-target gene interactions:** Transcription factors (TFs) and target genes (TGs) were defined by Drosophila Interaction Database modMine (Contrino et al. 2012) (v2015_12). Genes in this analysis were those that are present in Affymetrix Drosophila2 genome array annotation (Release 35), DroID TF-TG database v2015_12, and FlyBase v5.49 annotation. Over/under-representation of genes with significant inversion effect on expression as targets of transcription factors with SNPs in LD with inversion state was calculated as the probability of the observation given the hypergeometric distribution.

Custom scripts, data, and analysis results can be found online at [https://github.com/kern-lab/lavingtonKern](https://github.com/kern-lab/lavingtonKern), including file descriptions in the AnalysisFiles.readme document.

**Results**

To examine what influence, if any, common inversion polymorphisms have on patterns of transcription in the Drosophila genome we combined publically available genome sequences(Mackay et al. 2012), their associated karyotype calls (Corbett-Detig et al. 2012a; Huang et al. 2014; Houle and Márquez 2015), and previously published microarray based expression data(Ayroles et al. 2009). We validated the use of a model with only main effects of Sex, Line, and two cosmopolitan inversions, In(2L)t and In(3R)Mo using the R functions (add1) and (drop1). This main effects model performed better than any model having any of the main effect terms removed or with the addition of any interaction term (see Methods). After correction for multiple testing (see Methods), we found 229 and 498 total probesets with significant inversion effects for In(2L)t and In(3R)Mo, respectively, hereafter referred to as inversion affected loci (IAL) (Figure 1). These IAL occur both within the inversions themselves (40 In(2L)t, 111
In(3R)Mo), outside the inversions but near the breakpoints (3 In(2L)t, 38 In(3R)Mo within 1Mb of the breakpoint), as well as scattered throughout the genome (134 In(2L)t, 181 In(3R)Mo; see Table 1). This is a large number of loci with transcript abundance variation correlating with inversion state; however, we note that the inversion effect contribution to variance is relatively small for the vast majority of loci (Supplemental data).

One explanation for the large number of IAL found across the genome is that few loci are directly affected by the inversion and the remaining loci are affected indirectly by an expression variation correlation structure, previously described by Ayroles et al (2009). We addressed this correlation structure by the numbers of unique expression modules occupied by, and the distribution of correlation coefficients of, IAL as compared to all genes. If a significant portion of the IAL we observe are due to expression variation correlation, then we would expect that IAL occupy fewer expression modules than the same number of genes drawn at random. We would also expect the mean correlation of IAL between IAL to be higher than the genome wide average. We observe IAL occupy more modules than expected at random for both In(2L)t (71 obs; 38-56 95% c.i.) and In(3R)Mo (108 obs ; 65-87 95% c.i.). We also observed lower mean correlation between IAL than either between IAL and non-IAL, or the genome-wide mean for both inversions (Table 2).

We then examined the inversion effect on gene expression variation for four distinct categories of effect: 1.) cis- or 2.) trans-inversion effects of SNPs in LD with the inversion, 3.) direct effects of the inversion by interrupting genes, and 4.) regional effects of chromosomal rearrangement.

**Cis-Inversion effect of SNPs in LD with the inversion**

Our model explicitly tests the effect of chromosomal arrangement on expression variation and here we focus on SNP variation as the main driver of the inversion effect by taking advantage of LD between the
inversion state and SNP variants. LD with inversions is highest at the breakpoints and decays in both
directions from each breakpoint as expected (Figure 2) (Wesley and Eanes 1994; Andolfatto et al. 2001;
Langley et al. 2012; Corbett-Detig et al. 2012b). To determine cis-effects of the inversions we tested for
over-representation of IAL among loci in LD with the breakpoints. This ignores the location of the loci
and focuses on the correlation of SNP alleles with the inversion state. As expected, few loci in LD with
the inversion were located on a different chromosomal arm (3 with In(2L)t, and 2 with In(3R)Mo) and
IAL in LD are located only on the same chromosomal arm (Table 3).

**Trans-inversion effect of SNPs in LD with the inversion**

Our expectation of a cis-inversion effect is dependent on SNP variation in LD with the inversion. By the
same rationale, a trans-inversion effect may be detected as an IAL without SNP variation in LD with the
inversion, as we observe with a majority of IAL for each inversion (181 of 192 for In(2L)t and 323 of 425
for In(3R)Mo (see Table 3). Assuming SNP variation is the basis of expression variation, one trivial
explanation of a trans-inversion effect is SNP variation in transcription factors in LD with the inversion
acting on downstream targets. The TFs in this case need not be IAL as SNPs in protein coding regions of
TFs can give rise to expression variation in downstream targets. However, we found no over- or
underrepresentation of IAL that are targets of TFs with SNPs in LD with either inversion (Table 4). A
possible example of trans-inversion effect was found by functional analysis of IAL and discussed below.

**Direct effect of inversions by disrupting genes at breakpoints**

Nucleotide sequence variation at each breakpoint of an inversion can truncate transcribed gene regions
and disassociate transcribed regions from transcription factor binding sites and other regulatory
elements. Truncation most likely leads to down-regulation of genes and rearrangement of regulatory
elements and can lead to up- or downregulation of genes at either breakpoint. The closest IAL to any of
the four breakpoints are near the distal breakpoint of \( \text{In}(3R)\text{Mo} \). \text{CG1951} and \( 8\text{GalNAcTB} \) are within 2.5kb inside and outside of the inversion region, respectively, and \text{Ssl2} is interrupted by the breakpoint. All three are downregulated in the presence of the inversion. The next closest IAL is 24 kb away. The closest IAL to the proximal breakpoint of \( \text{In}(3R)\text{Mo} \) are 32kb away. It is also important to note that loci immediately surrounding the proximal breakpoint are not transcriptionally affected by the rearrangement, so the presumed disassociated regulatory elements from the distal end are not altering expression of loci at the proximal end. The closest IAL to the proximal and distal breakpoints of \( \text{In}(2L)t \) are 34kb and 37kb away, respectively, and thus probably too far to have been affected by direct effects of the breakpoint.

**Regional effect of chromosomal rearrangement**

We tested for regional effects of chromosomal rearrangement by looking for over- or underrepresentation of genes in LD with the inversion as IAL. We did observe more IAL than expected in LD with each inversion (Table 5) and a trend of transcriptional downregulation across the \( \text{In}(2L)t \) region (Figure 3 A & D), but note that the effect size is relatively small. Near the breakpoints, patterns of inversion effect direction are generally small and not likely significantly different from zero (Figure 3 B,C,E & F). The pattern of downregulation immediately surrounding the distal breakpoint of \( \text{In}(3R)\text{Mo} \) is moderate and appears to be driven by three loci: \text{CG1951}, \text{Ssl2}, and \( 8\text{GalNAcTB} \). This is likely a direct effect of the inversion on \text{Ssl2}, as the breakpoint interrupts this gene proximal to \text{CG1951} (Corbett-Detig \textit{et al.} 2012a), and possibly disassociation of regulatory elements from coding sequence with respect to \text{CG1951} and \( 8\text{GalNAcTB} \).

We also examined whether IAL tend to cluster together by physical location along chromosomes, which could arise from more localized regional effects. We measured physical clustering as the coefficient of variation (CV) of distances between genes, for the global CV, and between IAL by chromosome arm. For
each arm, IAL for both inversions were less clustered than the distribution of all genes, although

\[
\text{In}(3R)Mo \text{ IAL are more clustered on chromosome 3R than we would expect for the same number of}
\]

randomly drawn genes, but are still less clustered than the genomic background generally (Table 6).

Functional analysis

Coadapted alleles segregating with an inversion should also be in LD with the inversion breakpoints. We

used gProfiler g:GOSt functional profiling to detect overrepresentation of functional groups in sets of

IAL. The sets of IAL that we analyzed were those IAL in LD with the inversion (Table 2) or targets of

another gene with variation segregating with the inversion (Table 3). Functional analysis of IAL for each

inversion yielded significant groups only when considering all \text{In}(3R)Mo IAL or only IAL where inversion

state explains at least 15\% of expression variance (Supplemental data). Sterol transport is significant in

both cases (\(p=0.022\) for all, \(p=0.000116\) for \(\geq 15\%\) variance) and catalytic activity term (GO:0003824) is

significant when considering all IAL(\(p=0.000146\)). We found no significant functional groups when

considering any similar grouping of \text{In}(2L)t IAL.

The sterol transport group found to be enriched among \text{In}(3R)Mo IAL includes four of the eight

Niemann-Pick type 2 orthologs (\text{Npc2b}, \text{Npc2c}, \text{Npc2f}, \text{Npc2g}) (Huang et al. 2007), along with \text{Apoltp}. All

five genes are upregulated with respect to the inverted arrangement, suggesting an increase of sterol

uptake in \text{In}(3R)Mo bearing flies. In the context of the expression samples, most of these genes are

preferentially expressed in the adult gut (modENCODE, Contrino et al. 2012). While the four NPC2s are

on chromosome 3R (Figure 4), \text{Apoltp} is on chromosome 2L, and none of these genes contain a SNP in

the gene region that is in significant LD with \text{In}(3R)Mo, or each other, and only \text{Npc2f} is within the

inversion region (Figure 5). It is possible that the significant upregulation of the sterol transport group is

a downstream effect of \text{mod(mdg4)}, which is an IAL near the proximal breakpoint (Figure 4) and in LD

with \text{In}(3R)Mo (Figure 5). \text{mod(mdg4)} is a chromatin protein associated with \text{Npc2b} as well as other
chromatin proteins and TFs associated with all eight Npc2 orthologs (Supplemental data). We address
this exciting finding in further detail in the Discussion.

Discussion

Despite decades of research on polymorphic inversions in Drosophila melanogaster, and despite an
overwhelming consensus that inversions are maintained due to selection, we have little understanding
of the targets of selection in these inversions that leads to their maintenance (cf. Kirkpatrick and Kern
2012). Here we examine the role of transcriptional variation induced by cosmopolitan inversions to
explore what effect, if any, inversions might have on gene expression that itself might be selectively
favored. Besides gross rearrangement effects, we examined gene disruption at inversion breakpoints,
IAL in LD with the inversion, and IAL not in LD with the inversion. We assumed that multiple functionally
linked IAL to be potentially coadapted so long as they contained SNPs in LD with the inverted
arrangement. We also assumed that trans-inversion effects, IAL not in LD with the inversion, could be
the result of these loci interacting with TFs in LD with the inversion or epistatic interactions with loci in
LD. We note that while sample size of In(2L)t bearing individuals in the expression analysis is small (8 of
136), we were still able to detect a relatively large number of significant loci across the genome. We also
note that we could not address over- or under-dominance in this study as we examined only lines
known to be homozygous for either arrangement of In(2L)t or In(3R)Mo.

Of particular concern was what role, if any, expression correlation had on our findings. To explore this
we considered pairwise correlation coefficients and previously described expression modules (Ayroles et
al. 2009). One could imagine that the true false discovery rate would be much higher than anticipated
due to expression correlation of multiple loci from the same expression module. We found that IAL for
both inversions occupied more expression modules than we would expect when drawing genes at
random. Even when considering just the pairwise expression correlation, we found that IAL for both
inversions were slightly less correlated on average than the genome wide mean. These observations suggest that the inversion effects arose independent of the underlying correlation structure.

We found many loci of significant effect, yet most have a modest contribution to expression variation. Importantly, the only evidence of direct structural influence on patterns of expression was at a single breakpoint that appears to be due to the inversion mutation interrupting gene regions but not moving genes to a different region of the chromosome. That is, there was no appreciable pattern of up- or down-regulation of genes along the chromosome with respect to the location of inversion breakpoints. Perhaps unsurprisingly, we did find an overabundance of expression perturbation within the inversion regions themselves. This would suggest the effect is possibly due to genetic, rather than the structural, variation in LD with the inversion. That is, transcription variation associated with the inversion is due to genetic variation at the gene level and not the rearrangement of loci.

Inversion polymorphism can be maintained by reduced recombination between locally coadapted alleles within inversions (Dobzhansky and others 1970), both with (Nei et al. 1967; Pepper 2003) or without epistasis (Kirkpatrick and Barton 2006). Clinal inversion variation in D. melanogaster within In(3R)Payne in Australian populations and In(3R)Payne, In(2R)NS, and In(2L)t in American populations has been shown to be due to selection rather than demographic history and indicative of coadaptation (Kennington et al. 2006; Kolaczkowski et al. 2011; Kapun et al. 2016). We reasoned that one might be able to detect epistatic coadaptation through functional analysis of IAL. Coadaptation of non-interacting loci would likely result in either no significant functional groups or multiple, unrelated significant groups. We found functional annotation enrichment only when we considered all IALs with respect to In(3R)Mo. This would suggest that either one locus or a few coadapted, but non-interacting, loci segregate with these inversions to maintain polymorphic inversions. However, it is still possible that coadapted loci could be overlooked in our analysis; certainly the existing annotation is incomplete. Moreover, our
statistical power to detect IAL suffers due to the constraints of the number of inversions captured in the DGRP dataset. Nevertheless our observation that there are multiple IAL within the inversion with argues strongly for the role that inversions play as modifiers of recombination which may hold adaptive haplotypes together.

Our strongest functional finding, that sterol uptake associated with \textit{In(3R)Mo}, appears to be driven by genetic variation in a single locus as a \textit{trans}-inversion effect. Four of the five genes in this cluster are located on chromosome 3R, however none of these genes have a SNP in significant LD with \textit{In(3R)Mo}, or with each other (Supplemental data), and only one is found within the inversion region (\textit{Npc2f}). This would rule out effective coadaptation of these genes and suggest that the location of these genes on the same chromosome as the inversion is coincidence. Assuming the upstream effector of the sterol transport is a transcription factor, it could contain a SNP that alters either its protein-coding sequence or expression. Either scenario would require that the genetic variant responsible for the upregulation of the four Npc2s in question would have to be in LD with \textit{In(3R)Mo}. Only one TF, \textit{mod(mdg4)}, annotated in DroIDb as interacting with any (\textit{Npc2b}) of the five sterol transport IAL contains a SNP in LD with \textit{In(3R)Mo}. Furthermore, \textit{mod(mdg4)} is itself an IAL and also located near the proximal breakpoint of \textit{In(3R)Mo}. We speculate that inversion associated expression variation detected in this functional group is under control of \textit{mod(mdg4)}.

Increased sterol uptake fits nicely with the positive correlation of \textit{In(3R)Mo} frequency with latitude (Kapun \textit{et al.} 2014). \textit{Npc2} genes control sterol homeostasis via uptake of dietary sterols in \textit{D.melanogaster}(Huang \textit{et al}. 2007; Carvalho \textit{et al}. 2010; Niwa \textit{et al}. 2011). Two species of \textit{Drosophila} differentially express \textit{Npc2}'s in response to cold acclimation, though the patterns differ between the two (Parker \textit{et al}. 2015). Increased dietary cholesterol increases cold tolerance in \textit{Drosophila melanogaster} (Shreve \textit{et al}. 2007), however, \textit{D.melanogaster} takes up phytosterols more efficiently than
cholesterol (Cooke and Sang 1970). Furthermore, *D. melanogaster* is a sterol auxotroph (Carvalho *et al.* 2010; Niwa *et al.* 2011) that utilizes dietary sterols preferentially to biosynthesizing different sterols from dietary sterols (Carvalho *et al.* 2012). This would suggest that In(3R)Mo carrying *D. melanogaster* could be cold-acclimated due to increased uptake of dietary sterols, rather than the upregulation of cholesterol production.

It is difficult to interpret results for In(2L)t without a clear functional annotation group associated with the inverted state. Assuming In(2L)t is under selection (Kapun *et al.* 2016), the simplest interpretation is that In(2L)t polymorphism is maintained by only a small number of loci in LD with the inversion. It is possible that one or more loci in LD with In(2L)t contain protein coding variation under selection and no appreciable transcript abundance variation with respect to chromosomal arrangement. We note that there are only 14 IAL in LD with In(2L)t (Supplemental data). While a lack of a significant functional group may be dissatisfying, this does provide a manageable candidate list for validation of single targets.

**Conclusion**

We found that two different cosmopolitan inversions in *D. melanogaster* have a moderate effect on the expression of hundreds of genes across the genome. The genetic variation responsible for the observed transcriptional variation is only in small part due to the inversion event itself, with the majority of the variation being the result of either allelic variation in LD with the inversions, trans-effects of regulators that also are in LD with the inversions, or as of yet uncharacterized, indirect effects of the inversions. Our results mirror those of a recent report on transcriptional variation caused by inversions in *Drosophila pseudoobscura* (Fuller *et al.* 2016). Fuller et al. demonstrated quite convincingly that the well-studied polymorphic inversions of *D. pseudoobscura* are modulating levels of transcription at multiple life history stages and even found hints of trans-effects. However, due to experimental design constraints they could not examine inversion effects on unlinked chromosomes. Our findings extend this
pattern to *Drosophila melanogaster* and show that inversions are affecting loci genome-wide. While we have begun to parse the possible causes of IAL, our focus on available data limits the statistical power of our study. Thus it will be important to conduct carefully designed experiments on inversion polymorphism in *D. melanogaster* to elucidate the true extent of influence of inversions on genome-wide patterns of transcription.

**Acknowledgements**

We thank Dan Schrider and Steve Schaeffer for feedback on the manuscript. A.D.K. and E.L. were supported in part by NIH award no. R01GM078204.
Figure 1. Manhattan plot of log transformed, Bonferroni corrected p-values for an \( \text{In}(2L)t \) (top) and \( \text{In}(3R)Mo \) (bottom) inversion effect on transcription for each probe set across the genome. The blue bar above the points indicates the genomic location of the inversion. The red horizontal line represents the genome-wide significance threshold of \( p=5\times10^{-8} \).
Figure 2. Linkage disequilibrium between SNPs and inversion state. Disequilibrium is measured as $r^2$. Inversion breakpoints are depicted as red vertical lines. SNPs that are in significant LD with the inversion are shown in light blue ($p<0.05$, Bonferroni correction for all SNPs detected by chromosome arm).
Figure 3. Cohen’s d of inversion effect for probe sets by location. Cohen’s d for In(2L)t effect across chromosome 2L (A) and for In(3R)Mo effect across chromosome 3R (D). Loess curves (polynomial line) and 95% confidence intervals (shaded area) for 200kb surrounding each breakpoint: In(2L)t (B) proximal and (C) distal, In(3R)Mo proximal (E) and distal (F). Inversion breakpoints are depicted as vertical lines. Loess curves and confidence intervals generated by (geom_smooth) function in R package (ggplot2). Positive values of Cohen’s d represent increased transcript levels associated with the inverted chromosome state and negative values represent decreased transcript levels.
Figure 4. Location of *Npc2s* and *mod(mdg4)* on chromosome 3R. Long vertical lines represent breakpoints of *In(3R)Mo*. Only *mod(mdg4)* contains a SNP in LD with *In(3R)Mo*.
Figure 5. Correlation of SNPs with \textit{In(3R)Mo} by gene. All SNPs within gene regions and within 5kb up- and downstream of each gene: \textit{Npc2b}, \textit{Npc2c}, \textit{Npc2f}, \textit{Npc2g}, and \textit{mod(mdg4)}. Thin horizontal bars above each plot represent the gene region for that gene.
|                  | Within Inversion | Outside Inversion, on same Chr | < 1Mb outside Inversion | Chr 2L | Chr 2R | Chr 3L | Chr 3R | Chr X | Chr 4 |
|------------------|------------------|--------------------------------|-------------------------|--------|--------|--------|--------|-------|-------|
| In(2L)t          | 40               | 18                             | 3                       | 61     | 42     | 36     | 35     | 17    | 4     |
| In(3R)Mo         | 111              | 133                            | 38                      | 44     | 58     | 49     | 244    | 30    | 0     |

Table 1. Location of IAL. Counts of IAL between breakpoints, on the same chromosome as the inversion but outside of the breakpoints, and those on the same chromosome as the inversion but outside, and within 1Mb, of the breakpoints, as well as total IAL on each chromosome arm for each inversion.
### Table 2. Pairwise correlation of gene expression.

Mean $r^2$ was calculated from all pairwise correlation for the set description. IAL x IAL is the set of all pairwise correlation coefficients of expression between IAL for that inversion. IAL x other is the set of correlation coefficients of expression between IAL and all other (non-IAL).

|       | Mean $r^2$ |
|-------|------------|
| **In(2L)t** |            |
| IAL x IAL | 0.0965     |
| IAL x other | 0.0962    |
| **In(3R)Mo** |            |
| IAL x IAL | 0.1025     |
| IAL x other | 0.1042    |
| Genome-wide | 0.1181    |
| In(2L)t | chr2L | chr2R | chr3L | chr3R | chrX |
|---------|-------|-------|-------|-------|------|
| All loci | 370   | 1     | 0     | 1     | 1    |
| IAL     | 11    | 0     | 0     | 0     | 0    |

| In(3R)Mo | chr2L | chr2R | chr3L | chr3R | chrX |
|----------|-------|-------|-------|-------|------|
| All loci | 0     | 2     | 0     | 663   | 0    |
| IAL      | 0     | 0     | 0     | 110   | 0    |

Table 3. Location of loci with SNPs in LD with inversions. Loci are identified as BDGP v5.49 and Affymetrix library v.35 annotated gene regions containing SNPs in significant LD with the inversion (see Methods).
| Inversion | Total Genes | DroID TFs with Sig SNPs | Total DroID TF targets | Targets of Sig SNP DroID TFs also IAL | E[X] | p(E[X]>x) |
|-----------|-------------|-------------------------|-----------------------|-------------------------------------|------|----------|
| In(2L)t   | 10623       | 3                       | 3802                  | 181                                 | 53   | 65       | 0.96260   |
| In(3R)Mo  | 8           | 2373                    | 385                   | 99                                  | 86   | 0.04801  |

Table 4. IAL as targets of Transcription Factors with SNPs in LD with inversion. Expected counts E[X] and p-values calculated by hypergeometric distribution with uncorrected p-value cutoff of p=0.025 for the two-tailed test. Genes included in these analyses are annotated in BDGP v5.49, Affymetrix library v.35, and DroIDb v2015_12.
| Inversion | Total Genes | Total IAL | IAL in LD with Inversion | Genes in LD with Inversion | E[X] | p(E[X] > x) |
|-----------|-------------|-----------|--------------------------|---------------------------|------|-------------|
| In(2L)t   | 11969       | 192       | 11                       | 372                       | 6    | 0.01683     |
| In(3R)Mo  | 425         | 102       | 744                      | 26                        | 6.08 x 10^{-25} |

Table 5. IAL in LD with inversion. Expected counts (E[X]) and p-values are calculated by hypergeometric distribution with uncorrected p-value cutoff of p=0.025 for the two-tailed test. Genes included in analyses are annotated in both BDGP v5.49 and Affymetrix library version 35.
|     | In(2L)t | Genome | In(3R)Mo |
|-----|---------|--------|----------|
|     | 95% conf. int. | Obs | Obs | 95% conf. int. |
| chr2L | 0.882-1.366 | 1.338 | 2.612 | 1.417 | 0.834-1.305 |
| chr2R | 0.834-1.341 | 1.066 | 3.11 | 1.176 | 0.853-1.287 |
| chr3L | 0.806-1.352 | 0.979 | 2.812 | 1.129 | 0.829-1.278 |
| chr3R | 0.786-1.341 | 1.015 | 2.681 | 1.676 | 1.07-1.312 |
| chrX  | 0.724-1.399 | 1.152 | 2.814 | 0.944 | 0.734-1.295 |

Table 6. Gene location dispersion as described by the Coefficient of Variation (CV). Observed genome CV’s calculated from loci in both Flybase ver.5.49 and Affymetrix Drosophila2 r35 annotation databases by chromosome arm. See Methods for details.
Supplemental Tables

| Inversion | Total Genes | Total IAL | IAL Inside Inversion | Genes Inside Inversion | E[X]     | p(E[X]>x)     |
|-----------|-------------|-----------|----------------------|------------------------|----------|---------------|
| In(2L)t   | 11969       | 192       | 41                   | 1279                   | 20       | 4.3045 × 10^-6|
| In(3R)Mo  | 425         | 112       | 871                  | 31                     | 7.6 × 10^-36|

Supplemental Table 1. IAL inside inversion region. Expected counts E[X] and p-values are calculated by hypergeometric distribution with uncorrected p-value cutoff of p=0.025 for the two-tailed test. Genes included in analyses are annotated in both BDGP v5.49 and Affymetrix library version 35.

| Inversion | Total Genes | Total IAL | IAL as Fst outliers | Total Fst outliers | E[X]     | p(E[X]>x)     |
|-----------|-------------|-----------|---------------------|-------------------|----------|---------------|
| In(2L)t   | 11969       | 192       | 35                  | 1992              | 32       | 0.24089       |
| In(3R)Mo  | 425         | 79        | 71                  |                   | 0.12336  |

Supplemental Table 2. IAL as Fst outliers in clinal populations. Populations from Reinhardt et al (2014). Number of genes reflect the overlap between the unique Affymetrix Drosophila 2 annotated genes and FlyBase r5.49 annotations. Outliers are from the p=0.05 tail of the empirical distribution of 1kb windows (see Reinhardt et al (2014)).
Supplemental Methods

**Pairwise LD between SNPs:** was calculated with a custom Linux script as $r^2$ and $\chi^2$ calculated as $\chi^2 = Nr^2$, as described in Methods. We used the DGRP and excluded lines where either $In(2L)t$ or $In(3R)Mo$ was suspected to be segregating. We included SNPs that fell within the specified gene regions as annotated in FlyBase v5.49. SNPs with no base calls across all tested lines were excluded. Significance level was corrected for 843,051 tests (all unique, LD calculations for 1299 SNPs).
Literature Cited

Anderson W. W., Dobzhansky T., Kastritsis C. D., 1967 Selection and inversion polymorphism in experimental populations of Drosophila pseudoobscura initiated with the chromosomal constitutions of natural populations. Evolution (N. Y). 21: 664–671.

Anderson A. R., Hoffmann A. A., McKechnie S. W., Umina P. A., Weeks A. R., 2005 The latitudinal cline in the In(3R)Payne inversion polymorphism has shifted in the last 20 years in Australian Drosophila melanogaster populations. Mol. Ecol. 14: 851–8.

Andolfatto P., Wall J. D., Kreitman M., 1999 Unusual Haplotype Structure at the Proximal Breakpoint of In(2L)t in a Natural Population of Drosophila melanogaster. Genetics 153: 1297–1311.

Andolfatto P., Depaulis F., Navarro A., 2001 Inversion polymorphisms and nucleotide variability in Drosophila. Genet. Res. (Camb). 77: 1–8.

Aulard S., David J. R., Lemeunier F., 2002 Chromosomal inversion polymorphism in Afrotropical populations of Drosophila melanogaster. Genet. Res. (Camb). 79: 49–63.

Ayroles J. F., Carbone M. A., Stone E. A., Jordan K. W., Lyman R. F., Magwire M. M., Rollmann S. M., Duncan L. H., Lawrence F., Anholt R. R. H., Mackay T. F. C., 2009 Systems genetics of complex traits in Drosophila melanogaster. Nat. Genet. 41: 299–307.

Bergland A. O., 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: 1157–74.

Carvalho M., Schwudke D., Sampaio J. L., Palm W., Riezman I., Dey G., Gupta G. D., Mayor S., Riezman H., Shevchenko A., Kurzchalia T. V., Eaton S., 2010 Survival strategies of a sterol auxotroph. Development 137.

Carvalho M., Sampaio J. L., Palm W., Brankatschk M., Eaton S., Shevchenko A., Adachi-Yamada T., Gotoh T., Sugimura I., Tateno M., Nishida Y., Onuki T., Date H., Aldaz S., Escudero L., Freeman M., Baker K., Thummel C., Beadle G., Tatum E., Clancy C., Berndt A., Bernard A., Clouet P., Bhattacharyya A., Connor W., Birse R., Choi J., Reardon K., Rodriguez J., Graham S., Diop S., Ocorr K., Bodmer R., Oldham S., Blanksby S., Mitchell T., Borradaile N., Schaffer J., Burr G., Carvalho M., Schwudke D., Sampaio J., Palm W., Riezman I., Dey G., Gupta G., Mayor S., Riezman H., Shevchenko A., Kurzchalia T., Eaton S., Chan R., Uchil P., Jin J., Shui G., Ott D., Mothes W., Wenk M., Clayton R., Dennis E., Diagne A., Fauvel J., Record M., Chap H., Douste-Blazy L., Duan L., Wang H., Wang D., Eckel R., Grundy S., Zimmert P., Ejsing C., Meiring B., Bahr U., Duchoslav E., Karas M., Simons K., Shevchenko A., Ejsing C., Sampaio J., Surendranath V., Duchoslav E., Ekrko K., Klemm R., Simons K., Shevchenko A., Farooqui A., Horrocks L., Farquhar J., Ahrens E., Fast P., Gerl M., Sampaio J., Urbano S., Kaldova L., Verbavatz J., Binnington B., Lindemann D., Lingwood C., Shevchenko A., Schroeder C., Simons K., Gronke S., Milder S., Fellt L., Tennagels N., Petry S., Muller G., Jackle H., Kuhnein R., Gross R., Han X., Guo Y., Walther T., Rao M., Stuurman N., Goshima G., Terayama K., Wong J., Vale R., Walter P., Farese R., Hama H., Hamad M., Cooper B., Fisher N., Montooth K., Karty J., Han X., Yang K., Gross R., Harkewicz R., Dennis E., Herr D., Fyrst H., Phan V., Heinecke K., Georges R., Harris G., Saba J., Herzog R., Schwudke D., Schuhmann K., Sampaio J., Bornstein S., Schroeder M., Shevchenko A., Hobson R., Holman R., Johnson S., Hatch T.
Huang X., Suyama K., Buchanan J., Zhu A., Scott M., Iqbal J., Hussain M., Jiang C., Baehrecke E., Thummel C., Kalvodova L., Sampaio J., Cordo S., Ejsing C., Shevchenko A., Simons K., Koivusalo M., Haimi P., Heikinheimo L., Kostiainen R., Somerharju P., Kunde A., Matthews K., Rawson R., Leopold P., Perrimon N., Liebisch G., Binder M., Schifferer R., Langmann T., Schulz B., Schmitz G., Little S., Lynch M., Manku M., Nicolaou A., Lofgren H., Pascher I., Maldonado E., Alderson N., Monje P., Wood P., Hama H., Maldonado E., Casanave E., Aveldano M., Merkey A., Wong C., Hoshizaki D., Gibbs A., Mirth C., Riddiford L., Nelson W., Yeaman C., Nijhout H., Grunert L., Nijhout H., Williams C., Niwa R., Niwa Y., Palm W., Sampaio J., Branckatschk M., Carvalho M., Mahmoud A., Shevchenko A., Eaton S., Panakova D., Sprong H., Marois E., Thiele C., Eaton S., Pavlidis P., Ramaswami M., Tanouye M., Porter F., Reisse S., Rothardt G., Volkl A., Beier K., Riccardi G., Giacco R., Rivellese A., Rusten T., Lindmo K., Juhasz G., Sass M., Seglen P., Brech A., Stenmark H., Salem N., Litman B., Kim H., Gawrisch K., Sampaio J., Gerl M., Klose C., Ejsing C., Beug H., Simons K., Shevchenko A., Sandhoff R., Brugger B., Jeckel D., Lehmann W., Wieland F., Schuhmann K., Almeida R., Baumert M., Herzog R., Bornstein S., Shevchenko A., Schwudke D., Hennich J., Suredranath V., Grimard V., Moehring T., Burton L., Kurzhalia T., Shevchenko A., Schwudke D., Schuhmann K., Herzog R., Bornstein S., Shevchenko A., Simons K., Souza C., Schwabe T., Ploierz B., Leitner E., Guan X., Wenk M., Riezman I., Riezman H., Tanasescu M., Cho E., Manson J., Hu F., Tarr P., Arling E., Bojanic D., Edwards P., Baldan A., Tschepe J., Hammerschmied C., Muhl-Versen M., Athenstaedt K., Daum G., Kretzschmar D., Uauy R., Dangour A., Uchida Y., Hama H., Alderson N., Douangpanya S., Wang Y., Murkina D., Elias P., Holleran W., Meier G., van Wende A., Abel E., Wenk M., Yehuda S., Rabinovitz S., Mostofsky D., Yeo Y., Philbrick D., Holub B., Yetukuri L., Ekkos K., Vidal-Puig A., Oresic M., Yu L., Bergmann K., von Latzjohn D., Hobbs H., Cohen J., Zeisel S., Zhang H., Abraham N., Khan L., Hall D., Fleming J., Göbel V., 2012 Effects of diet and development on the Drosophila lipidome. Mol. Syst. Biol. 8: 600.

Chambers G. K., 1991 Gene expression, adaptation and evolution in higher organisms. Evidence from studies of Drosophila alcohol dehydrogenases. Comp. Biochem. Physiol. Part B Comp. Biochem. 99: 723–730.

Cohen J., 1988 Statistical Power Analysis for the Behavioral Sciences.

Contrino S., Smith R. N., Butano D., Carr A., Hu F., Lyne R., Rutherford K., Kalderimis A., Sullivan J., Carbon S., Kephart E. T., Lloyd P., Stinson E. O., Washington N. L., Perry M. D., Ruzanov P., Zha L., Lewis S. E., Stein L. D., Micklem G., 2012 modMine: flexible access to modENCODE data. Nucleic Acids Res. 40: D1082-8.

Cooke J., Sang J. H., 1970 Utilization of sterols by larvae of Drosophila melanogaster. J. Insect Physiol. 16: 801–812.

Corbett-Detig R. B., Cardeno C., Langley C. H., 2012a Sequence-based detection and breakpoint assembly of polymorphic inversions. Genetics 192: 131–7.

Corbett-Detig R. B., Hartl D. L., Sturtevant A., Noor M., Grams K., Bertucci L., Reiland J., Rieseberg L., Kirkpatrick M., Barton N., Sandler L., Hiraizumi Y., Sandler I., Jaenike J., Presgraves D., Gerard P., Cherukuri A., Lyttle T., Hoffmann A., Rieseberg L., Stalker H., Mettler L., Voelker R., Mukai T., Stalker H., Knibb W., Oakenshot J., Gibson J., Inoue Y., Watanabe T., Watanabe T., Watanabe T., Andolfatto P., Wall J., Kreitman M., Matzkin L., Merritt T., Zhu C., Eanes W., Wesley C., Eanes W., Hasson E., Eanes W., Corbett-Detig R., Cardeno C., Langley C., Aulard S., David J., Lemeunier F.,
Pool J., Corbett-Detig R., Sugino R., Stevens K., Cardeno C., Mackay T., Richards S., Stone E.,
Barbadilla A., Ayroles J., Li H., Stephan W., Thornton K., Andolfatto P., Glinka S., Ometto L.,
Mousset S., Stephan W., Lorenzo D. De, Li H., Durbin R., Li H., Handsaker B., Wysoker A., Fennell T.,
Ruan J., Smith J. M., Haigh J., Langley C., Stevens K., Cardeno C., Lee Y., Schrider D., Wallace A.,
Detweiler D., Schaeffer S., Tajima F., Schaeffer S., Pritchard J., Seielstad M., Perez-Lezaun A.,
Feldman M., Beaumont M., Zhang W., Balding D., Nielsen R., Williamson S., Kim Y., Hubisz M., Clark A.,
Andolfatto P., Depaulis F., Navarro A., Aulard S., Monti L., Chaminade N., Lemeunier F., Ranz J.,
Maurin D., Chan Y., Grotthuss M., Hillier L., Pool J., Aquadro C., Navarro A., Betran E., Barbadilla A.,
Ruiz A., Patterson N., Price A., Reich D., Kulathinal R., Stevison L., Noor M., Payne F., Navarro A.,
Barbadilla A., Ruiz A., Baudry E., Viginier B., Veuille M., Smith J. M., Haigh J., Charlesworth B.,
Richards S., Liu Y., Bettencourt B., Hradecky P., Letovsky S., Reed F., Reeves R., Aquadro C., Novitski E.,
Hanks G., Hanks G., Adams M., Celniker S., Holt R., Evans C., Gocayne J., Thompson J., Higgins D.,
Gibson T., Hu T., Eisen M., Thornton K., Andolfatto P., Hudson R., Hudson R., Slatkin M., Maddison W., 2012b Population Genomics of Inversion Polymorphisms in Drosophila melanogaster (HS Malik, Ed.). PLoS Genet. 8: e1003056.

Dobzhansky T., Sturtevant a. H., 1938 Inversions in the chromosomes of Drosophila pseudoobscura. Genetics 23: 28–64.

Dobzhansky T., 1947 Genetics of Natural Populations. Xiv. a Response of Certain Gene Arrangements in the Third Chromosome of DROSOPHILA PSEUDOOSBSCURA to Natural Selection. Genetics 32: 142–60.

Dobzhansky T., others, 1970 Genetics of the evolutionary process. Columbia University Press New York.

Feuk L., 2010 Inversion variants in the human genome: role in disease and genome architecture. Genome Med. 2: 11.

Fouet C., Gray E., Besansky N. J., Costantini C., Reig O., 2012 Adaptation to Aridity in the Malaria Mosquito Anopheles gambiae: Chromosomal Inversion Polymorphism and Body Size Influence Resistance to Desiccation (J Pinto, Ed.). PLoS One 7: e34841.

Fraser A., Burnell D., Miller D., 1966 Simulation of genetic systems X. Inversion polymorphism. J. Theor. Biol. 13: 1–14.

Fraser H. B., 2013 Gene expression drives local adaptation in humans. Genome Res. 23: 1089–96.

Fuller Z. L., Haynes G. D., Richards S., Schaeffer S. W., 2016 Genomics of Natural Populations: How Differentially Expressed Genes Shape the Evolution of Chromosomal Inversions in Drosophila pseudoobscura. Genetics 204.

Haldane J. B. S., 1957 The conditions for coadaptation in polymorphism for inversions. J. Genet. 55: 218–225.

Hoffmann A. A., Sgrò C. M., Weeks A. R., 2004 Chromosomal inversion polymorphisms and adaptation. Trends Ecol. Evol. 19: 482–8.

Hoffmann A. A., Shirriffs J., Scott M., 2005 Relative importance of plastic vs genetic factors in adaptive differentiation: Geographical variation for stress resistance in Drosophila melanogaster from
eastern Australia. Funct. Ecol. 19: 222–227.

Houle D., Márquez E. J., 2015 Linkage Disequilibrium and Inversion-Typing of the Drosophila melanogaster Genome Reference Panel. G3 (Bethesda). 5: 1695–701.

Huang X., Warren J. T., Buchanan J., Gilbert L. I., Scott M. P., 2007 Drosophila Niemann-Pick type C-2 genes control sterol homeostasis and steroid biosynthesis: a model of human neurodegenerative disease. Development 134: 3733–42.

Huang W., Massouras A., Inoue Y., Peiffer J., Ràmia M., Tarone A. M., Turlapati L., Zichner T., Zhu D., Lyman R. F., Magwire M. M., Blankenburg K., Carbone M. A., Chang K., Ellis L. L., Fernandez S., Han Y., Highnam G., Hjelmen C. E., Jack J. R., Javaid M., Jayaseelan J., Kalra D., Lee S., Lewis L., Munidasa M., Ongeri F., Patel S., Perales L., Perez A., Pu L., Rollmann S. M., Ruth R., Saada N., Warner C., Williams A., Wu Y.-Q., Yamamoto A., Zhang Y., Zhu Y., Anholt R. R. H., Korbel J. O., Mittelman D., Muzny D. M., Gibbs R. A., Barbadilla A., Johnston J. S., Stone E. A., Richards S., Deplancke B., Mackay T. F. C., 2014 Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. Genome Res. 24: 1193–208.

Huang W., Carbone M. A., Magwire M. M., Peiffer J. A., Lyman R. F., Stone E. A., Anholt R. R. H., Mackay T. F. C., 2015 Genetic basis of transcriptome diversity in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 112: E6010-6019.

Jong G. de, Bochdanovits Z., 2003 Latitudinal clines in Drosophila melanogaster: Body size, alleloyme frequencies, inversion frequencies, and the insulin-signalling pathway. J. Genet. 82: 207–223.

Kapun M., Schalkwyk H. van, McAllister B., Flatt T., Schlötterer C., 2014 Inference of chromosomal inversion dynamics from Pool-Seq data in natural and laboratory populations of Drosophila melanogaster. Mol. Ecol. 23: 1813–1827.

Kapun M., Fabian D. K., Goudet J., Flatt T., 2016 Genomic Evidence for Adaptive Inversion Clines in Drosophila melanogaster. Mol. Biol. Evol.: msw016-.

Kennington W. J., Partridge L., Hoffmann A. A., 2006 Patterns of diversity and linkage disequilibrium within the cosmopolitan inversion In(3R)Payne in Drosophila melanogaster are indicative of coadaptation. Genetics 172: 1655–63.

Kidd J. M., Graves T., Newman T. L., Fulton R., Hayden H. S., Malig M., Kallicki J., Kaul R., Wilson R. K., Eichler E. E., 2010 A Human Genome Structural Variation Sequencing Resource Reveals Insights into Mutational Mechanisms. Cell 143: 837–847.

Kirkpatrick M., Barton N., 2006 Chromosome inversions, local adaptation and speciation. Genetics 173: 419–34.

Kirkpatrick M., Kern A., 2012 Where’s the money? Inversions, genes, and the hunt for genomic targets of selection. Genetics 190: 1153–5.

Knibb W. R., Oakeshott J. G., Gibson J. B., 1981 Chromosome Inversion Polymorphism in Drosophila melanogaster. I. Latitudinal Clines and Associations between Inversions in Australian Populations. Genetics 98: 833–847.
Kolaczkowski B., Kern A. D., Holloway A. K., Begun D. J., 2011 Genomic differentiation between temperate and tropical Australian populations of Drosophila melanogaster. Genetics 187: 245–60.

Kolesnikov N., Hastings E., Keays M., Melnichuk O., Tang Y. A., Williams E., Dylag M., Kurbatova N., Brandizi M., Burdett T., Megy K., Pilicheva E., Rustici G., Tikhonov A., Parkinson H., Petryszak R., Sarkans U., Brazma A., 2015 ArrayExpress update--simplifying data submissions. Nucleic Acids Res. 43: D1113-6.

Krimbas C. B., Powell J. R., 1992 Drosophila Inversion Polymorphism. CRC Press.

Lakich D., Kazazian H. H., Antonarakis S. E., Gitschier J., 1993 Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. Nat. Genet. 5: 236–241.

Langley C. H., Stevens K., Cardeno C., Lee Y. C. G., Schrider D. R., Pool J. E., Langley S. A., Suarez C., Corbett-Detig R. B., Kolaczkowski B., Fang S., Nista P. M., Holloway A. K., Kern A. D., Dewey C. N., Song Y. S., Hahn M. W., Begun D. J., 2012 Genomic variation in natural populations of Drosophila melanogaster. Genetics 192: 533–98.

Levene H., Dobzhansky T., 1958 New evidence of heterosis in naturally occurring inversion heterozygotes in Drosophila pseudoobscura. Heredity (Edinb). 12: 37–49.

López-Maury L., Marguerat S., Bähler J., 2008 Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat. Rev. Genet. 9: 583–93.

Lowry D. B., Willis J. H., 2010 A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS Biol. 8: e1000500.

Mackay T. F. C., Richards S., Stone E. A., Barbadilla A., Ayroles J. F., Zhu D., Casillas S., Han Y., Magwire M. M., Cridland J. M., Richardson M. F., Anholt R. R. H., Barrón M., Bess C., Blankenburg K. P., Carbone M. A., Castellano D., Chaboub L., Duncan L., Harris Z., Javaid M., Jayaseelan J. C., Jhangiani S. N., Jordan K. W., Lara F., Lawrence F., Lee S. L., Librado P., Linheiro R. S., Lyman R. F., Mackey A. J., Munidasa M., Muzny D. M., Nazareth L., Newsham I., Perales L., Pu L.-L., Qu C., Râmia M., Reid J. G., Rollmann S. M., Rosas J., Saada N., Turlapati L., Worley K. C., Wu Y.-Q., Yamamoto A., Zhu Y., Bergman C. M., Thornton K. R., Mittelman D., Gibbs R. A., 2012 The Drosophila melanogaster Genetic Reference Panel. Nature 482: 173–8.

Massouras A., Waszak S. M., Albarca-Aguilera M., Hens K., Holcombe W., Ayroles J. F., Dermitzakis E. T., Stone E. A., Jensen J. D., Mackay T. F. C., Deplancke B., 2012 Genomic variation and its impact on gene expression in Drosophila melanogaster. PLoS Genet. 8: e1003055.

Matzkin L. M., Merritt T. J. S., Zhu C.-T., Eanes W. F., 2005 The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion In(3R)Payne in Drosophila melanogaster. Genetics 170: 1143–52.

Nei M., Kojima K. I., Schaffer H. E., 1967 Frequency changes of new inversions in populations under mutation-selection equilibria. Genetics 57: 741–50.

Niwa R., Niwa Y. S., Niwa R., Niwa Y. S., 2011 The Fruit Fly Drosophila melanogaster as a Model System to Study Cholesterol Metabolism and Homeostasis. Cholesterol 2011: 1–6.
Parker D. J., Vesala L., Ritchie M. G., Laiho A., Hoikkala A., Kankare M., 2015 How consistent are the transcriptome changes associated with cold acclimation in two species of the Drosophila virilis group? Heredity (Edinb). 115: 13–21.

Pepper J. W., 2003 The evolution of evolvability in genetic linkage patterns. Biosystems 69: 115–126.

Pool J. E., Corbett-Detig R. B., Sugino R. P., Stevens K. A., Cardeno C. M., Crepeau M. W., Duchen P., Emerson J. J., Saelao P., Begun D. J., Langley C. H., 2012 Population Genomics of sub-saharan Drosophila melanogaster: African diversity and non-African admixture. PLoS Genet. 8: e1003080.

R Core Team, 2013 R: A Language and Environment for Statistical Computing.

Reimand J., Arak T., Adler P., Kolberg L., Reisberg S., Peterson H., Vilo J., 2016 g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res.: gkw199-.

Reinhardt J. A., Kolaczkowski B., Jones C. D., Begun D. J., Kern A. D., Akey J. M., Zhang G., Zhang K., Jin L., Shriver M. D., Beaumont M. A., Nichols R. A., Begun D. J., Aquadro C. F., Cavalli-Sforza L. L., Edwards A. W., Christin P. A., Weinreich D. M., Besnard G., Coop G., Witonsky D., Rienzo A. Di, Pritchard J. K., Daborn P. J., Yen J. L., Bogwitz M. R., Goff G. Le, Feil E., David J. R., Capy P., Duchen P., Živković D., Hutter S., Stephans W., Laurent S., Fabian D. K., Kapun M., Nolte V., Kofler R., Schmidt P. S., Gompel N., Prud’homme B., Hoffmann A. A., Weeks A. R., Hoffmann A. A., Sgrò C. M., Weeks A. R., Hohenlohe P. A., Bassham S., Etter P. D., Stiffler N., Johnson E. A., Huang D. W., Sherman B. T., Zheng X., Yang J., Imamichi T., Jones F. C., Grabherr M. G., Chan Y. F., Russell P., Mauceli E., Keller A., Knibb W. R., Kirkpatrick M., Barton N., Kirkpatrick M., Kern A., Kolaczkowski B., Kern A. D., Holloway A. K., Begun D. J., Kopp A., Kyriacou C. P., Peixoto A. A., Sandrelli F., Costa R., Tauber E., Lachaise D., Cariou M. L., David J. R., Lemeunier F., Tsacas L., Langley C. H., Stevens K., Cardeno C., Lee Y. C. G., Schrider D. R., Li H., Durbin R., Losos J. B., Mackay T. F., Richards S., Stone E. A., Barbadilla A., Ayroles J. F., McKechnie S. W., Blacket M. J., Song S. V., Rako L., Carroll X., Nei M., Pavlidis P., Jensen J. D., Stephans W., Stamatakis A., Pickrell J. K., Coop G., Novembre J., Kudaravalli S., Li J. Z., Sakton T. B., Kulathinal R. J., Bergman C. M., Quinlan A. R., Dopman E. B., Saitou N., Nei M., Schmidt J. M., Good R. T., Appleton B., Sherrard J., Raymang C. G., Singh R. S., Long A. D., Stalker H. D., Stephans W., Li H., Teshima K. M., Coop G., Przeworski M., Turner T. L., Levine M. T., Eckert M. L., Begun D. J., Voelker R. A., Cockerham C. C., Johnson F. M., Schaffer H. E., Mukai T., Voight B. F., Kudaravalli S., Wen X., Pritchard J. K., 2014 Parallel geographic variation in Drosophila melanogaster. Genetics 197: 361–73.

Schrider D. R., Hahn M. W., Begun D. J., 2016 Parallel Evolution of Copy-Number Variation across Continents in Drosophila melanogaster. Mol. Biol. Evol. 33: 1308–1316.

Sezgin E., Duvernell D. D., Matzkin L. M., Duan Y., Zhu C.-T., Ferrelli B. C., Eanes W. F., 2004 Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. Genetics 168: 923–31.

Shreve S. M., Yi S. X., Lee R. E., 2007 Increased dietary cholesterol enhances cold tolerance in Drosophila melanogaster. 28: 33–37.

Sokal R. R., Rohlf F. J., 1995 Biometry.

Stefansson H., Helgason A., Thorleifsson G., Steinhorsdottir V., Masson G., Barnard J., Baker A.,
Jonasdottir A., Ingason A., Gudnadottir V. G., Desnica N., Hicks A., Gyflason A., Gudbjartsson D. F., Jonsdottir G. M., Sainz J., Agnarsson K., Birgisdottir B., Ghosh S., Olafsdottir A., Cazier J.-B., Kristjansson K., Frigge M. L., Thorgerisson T. E., Gulcher J. R., Kong A., Stefansson K., 2005 A common inversion under selection in Europeans. Nat. Genet. 37: 129–37.

Storey J. D., Tibshirani R., 2003 Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. 100: 9440–5.

Sturtevant A. H., Mather K., 1938 The Interrelations of Inversions, Heterosis and Recombination. Am. Nat. 72: 447–452.

Teroni Scott M., Sgrò C. M., Hoffmann A. A., Griffin P. C., 2016 Cross-Study Comparison Reveals Common Genomic, Network, and Functional Signatures of Desiccation Resistance in Drosophila melanogaster. Mol. Biol. Evol. 33: 1053–1067.

Tobari Y. N., Kojima K. I., 1967 Selective modes associated with inversion karyotypes in Drosophila ananassae. I. Frequency-dependant selection. Genetics 57: 179–88.

Umina P. A., Weeks A. R., Kearney M. R., McKechnie S. W., Hoffmann A. A., 2005 A rapid shift in a classic clinal pattern in Drosophila reflecting climate change. Science 308: 691–3.

Vonesch S. C., Lamparter D., Mackay T. F. C., Bergmann S., Hafen E., Markow T., Jensen L., Lee S., Wee C., Hoffmann A., 2016 Genome-Wide Analysis Reveals Novel Regulators of Growth in Drosophila melanogaster (GS Barsh, Ed.). PLOS Genet. 12: e1005616.

Weeks A. R., McKechnie S. W., Hoffmann A. A., 2002 Dissecting adaptive clinal variation: markers, inversions and size/stress associations in Drosophila melanogaster from a central field population. Ecol. Lett. 5: 756–763.

Wesley C. S., Eanes W. F., 1994 Isolation and analysis of the breakpoint sequences of chromosome inversion In(3L)Payne in Drosophila melanogaster. Proc. Natl. Acad. Sci. 91: 3132–3136.