Genomics of Recombination Rate Variation in Temperature-Evolved *Drosophila melanogaster* Populations

Ari Winbush and Nadia D. Singh*

Department of Biology, Institute of Ecology and Evolution, University of Oregon, Eugene, Oregon, USA

*Corresponding author: E-mail: nsingh@uoregon.edu.

Accepted: 24 November 2020

Abstract

Meiotic recombination is a critical process that ensures proper segregation of chromosome homologs through DNA double-strand break repair mechanisms. Rates of recombination are highly variable among various taxa, within species, and within genomes with far-reaching evolutionary and genomic consequences. The genetic basis of recombination rate variation is therefore crucial in the study of evolutionary biology but remains poorly understood. In this study, we took advantage of a set of experimental temperature-evolved populations of *Drosophila melanogaster* with heritable differences in recombination rates depending on the temperature regime in which they evolved. We performed whole-genome sequencing and identified several chromosomal regions that appear to be divergent depending on temperature regime. In addition, we identify a set of single-nucleotide polymorphisms and associated genes with significant differences in allele frequency when the different temperature populations are compared. Further refinement of these gene candidates emphasizing those expressed in the ovary and associated with DNA binding reveals numerous potential candidate genes such as *Hr38*, *EcR*, and *mamo* responsible for observed differences in recombination rates in these experimental evolution lines thus providing insight into the genetic basis of recombination rate variation.

Key words: recombination rate, *Drosophila*, temperature, genome sequencing, experimental evolution, divergence.

Significance

The genetic basis of recombination rate variation in *Drosophila melanogaster* remains an area of active study with few genes identified thus far. By performing whole-genome sequencing in a set of temperature-evolved Drosophila populations that have evolved heritable differences in recombination rate, this study identifies a large set of gene candidates responsible for recombination rate variation in these populations with broader implications in this field. Furthermore, these strains have been previously used in several other studies and our data are therefore also relevant in the broader context of these strains.

Introduction

Meiotic recombination is the transfer of genetic material between homologous chromosomes during prophase I of meiosis and is necessary for proper segregation of homologous chromosomes during the first meiotic division arising as a consequence of DNA double-strand break repair. This process is distinct from noncrossover-associated gene conversions which are normally associated with the nonreciprocal exchange of genetic information from one chromosome homolog to another (reviewed in Baudat et al. [2013]; Hughes et al. [2018]). Evolutionary and genomic consequences of recombination are many and include the shuffling of alleles to create unique haplotypes and loss of linkage disequilibrium (Hill and Robertson 1966; Kliman and Hey 1993). Although recombination can be beneficial by bringing together advantageous alleles, recombination can also be deleterious by breaking apart certain allele combinations. Notably, an absence of recombination can lead to aneuploidy (Hill and

Robertson 1966; Barton 1995; Charlesworth and Barton 1996; reviewed in Barton and Charlesworth 1998; Otto 2009; Ritz et al. 2017), illustrating its vital role for chromosome segregation.

Recombination rate shows variation across genomes, between and within populations and among species (reviewed in Stapley et al. 2017; reviewed in Ritz et al. 2017). Variation in recombination rate within populations and between sexes has been extensively demonstrated in vertebrate systems including mice, humans, cattle, and sheep, as well as in plants such as Arabidopsis (Dumont et al. 2009; Kong et al. 2014; Ma et al. 2015; Johnston et al. 2016; Lambing et al. 2017; reviewed in Wang and Copenhaver 2018). Research into the genetic basis of recombination rate in vertebrate systems has led to the discovery of several prominent genomic loci with roles in mediating recombination rate within genomes and between individuals. This includes *PRDM9* which encodes a histone methyltransferase and appears to play a critical role in specifying recombination hotspots in the human and mouse genome, potentially by directing DSB formation via targeted recruitment of the SPO11–Rad50 protein complex (Baudat et al. 2010; Berg et al. 2011; reviewed in Baudat et al. [2013]). Allelic variation at this locus is associated with fine-scale variation in recombination hotspots among human populations (Hinch et al. 2011), and evidence for positive selection for *PRDM9* variants at the level of zinc-finger domains across several taxa have been demonstrated (Oliver et al. 2009; Schwartz et al. 2014). *PRDM9* has also been shown to be associated with individual variation in recombination rate at the genome level (Ma et al. 2015).

Additional loci associated with variation in recombination rate within populations and between sexes in vertebrates are *RNF212* which encodes a RING finger protein that is homologous to crossover factor Zip3 in yeast, the meiosis-specific cohesion proteins *REC8* and *RAD21L*, the E3 ubiquitin ligase *CCNB1IP1*, and the protein kinase encoding gene *NEK9* (Chowdhury et al. 2009; Fledel-Alon et al. 2011; Sandor et al. 2012; Kong et al. 2014; Ma et al. 2015; reviewed in Baudat et al. 2013; Stapley et al. 2017). Likewise, different orientations of a 900-kb inversion on chromosome 17q21.31 in humans, for example, show patterns of divergence in recombination rate between different populations. One distinct orientation, under positive selection in European lineages, is associated with increased female recombination rate and fecundity (Stefansson et al. 2005).

Compared with vertebrates, substantially less is known about genomic loci responsible for variation in recombination rate in *Drosophila melanogaster*. Empirical data, however, suggest the existence of such loci. Recombination rates show distinct patterns of heritability and variation between different fly populations and species, individuals, and within genomes (Brooks and Marks 1986; True et al. 1996; Ortiz-Barrientos et al. 2006; Singh et al. 2009; Chan et al. 2012; Manzano-Winkler et al. 2013; Smukowski Heil et al. 2015; Brand et al. 2018). In *D. melanogaster*, experimental evidence supports the existence of recombination modifiers across the three major chromosomes, evidenced by successful directional selection on recombination rate (Chinnici 1971; Kidwell 1972), and recombination rate has been shown to be subject to variation in genetic background (Hunter et al. 2016).

Previous work has identified a small number of candidate loci associated with recombination rate variation. The meiotic gene *mei-1* was previously associated with localized changes in recombination on the *X*-chromosome (Valentin 2009). Several loci were implicated in recombination rate variation among a set of inbred *D. melanogaster* lines over a limited set of chromosomes including: *lola*, *CG10864*, *Eip75B*, and *Ptp61F* (Hunter et al. 2016). Transcriptome analysis of ovarian germaria cells undergoing early meiosis likewise reveals several candidates associated with proteolysis undergoing high levels of transcription that are both associated with, and the target of DSB events (Adrian and Comeron 2013). Furthermore, variation within the *mei-217/mei-218* locus has been linked to recombination rate variation as demonstrated by differences in crossover events per tetrad between two closely related *Drosophila* species (Brand et al. 2018).

In order to identify additional candidate loci that potentially mediate recombination rate variation, we took advantage of several experimentally evolved *D. melanogaster* lines that were maintained for multiple generations in either a warm, cold, or fluctuating temperature regime (Yeaman et al. 2010; Cooper et al. 2012). These strains have diverged in recombination rate based on the corresponding temperature regime in which they previously evolved (Kohl and Singh 2018). When recombination rates were measured in these populations using a classic two-step backcrossing protocol utilizing visible recessive markers, the warm-regime populations exhibited the highest recombination rate as measured by quantifying the proportion of recombinant offspring. The fluctuating-regime populations exhibited an intermediate recombination rate, whereas the cold-regime populations had the lowest recombination rates. These differences were observed regardless of rearing temperature, although recombination rates were significantly higher at warmer temperatures in all three populations (Kohl and Singh 2018).

We therefore used whole-genome sequencing on these lines to identify candidate loci underlying divergence in recombination rate. Such loci would be expected to show differences in single-nucleotide polymorphism (SNP) allele frequency between populations based on the temperature regime in which they evolved. We note that these populations are also likely to have experienced selection on other traits associated with adaptation to thermal regime. Indeed, divergence in several additional traits has been documented in these populations across temperature regimes (Cooper et al. 2012; Condon et al. 2014, 2015; Adrian et al. 2016; Le Vinh Thuy et al. 2016; Alton et al. 2017). Given these findings, it is

Table 1
Replicate Populations from Each Temperature Regime and the Number of Isogenic Lines within Each Population

Cold-Regime	No. of Lines	Warm-Regime	No. of Lines	Temp-Regime	No. of Lines
C1	10	H1	9	T1	8
C2	9	H2	9	T2	11
C3	9	H3	10	T3	9
C4	10	H4	10	T4	10
C5	10	H5	10	T5	10

Note.—Individual embryos from each line were utilized in our studies ensuring homogenous representation of each replicate population from their corresponding temperature regime.

likely that many loci we identify in this manner include potential candidates that underlie evolutionary responses to selection on traits other than recombination. However, narrowing our dataset to genes with particular functions can partially overcome this limitation. Therefore, to identify the candidates most likely related to divergence in recombination rate, we focused almost exclusively on genes with annotated expression in the ovaries. We then further narrowed our results to genes associated with DNA binding and those that functioned as zinc finger transcription factors as direct interaction with DNA is likely a factor in regulating recombination rate variation. Our results highlight a set of candidate genes underlying population divergence in recombination rate.

These candidates included loci previously identified as mediating population level variation in recombination rate (Hunter et al. 2016). In addition to identifying novel candidates for the genetic basis of recombination rate variation, our data suggest the loci associated with variation in recombination rate within populations also contribute to variation in recombination rate between populations.

Materials and Methods

Experimentally Evolved Populations Used in This Study

Experimental populations utilized in this study have been previously generated and described (Yeaman et al. 2010). Mated *D. melanogaster* females were captured near Cawston, British Columbia, Canada and used to establish a set of reciprocally crossed isofemale lines which were allowed to expand for six generations, resulting in a population of at least 64,000 individuals. Three groups of five replicate populations were established from this initial population and each group was maintained in one of three environments that differed by temperature regime. A "Cold" group was maintained continuously at 16 °C, a "Warm" group at 25 °C, and finally, a "Temp" group was maintained in a temporally fluctuating environment by switching between both the Warm and Cold temperature regimes every four weeks (Yeaman et al. 2010). After three years, 225 isofemale lines representing 15 replicate strains from each of the 15 populations were established (Cooper et al. 2012). We maintained these lines under standard laboratory conditions (20.5 °C, 12 h:12 h light/dark cycle) for approximately 30 months prior to experimentation.

Experimental Crosses

Following previous studies (Langley et al. 2011), we generated haploid embryos for whole-genome sequencing. We took advantage of a male-sterile mutation, *ms(3)K81*, in which the majority of the progeny produced from males homozygous for this mutation arrest during the blastoderm stage, with only approximately 1% surviving to the first instar before also terminating development (Fuyama 1984). However, embryos that survive are haploid and their genome content is entirely maternally derived. These embryos are thus suitable and advantageous for whole-genome sequencing given that no haplotype phasing is required.

Females from each of the isofemale lines were crossed to *ms(3)K81* males and the resulting gynogenetically derived haploid embryos were allowed to develop for 12–24 h. Dechorionated embryos were examined, and those that exhibited normal initial development consistent with healthy 24-h embryos, were individually collected from each population. We collected one embryo each from 144 of the 225 isofemale lines representing each of the 15 replicate populations for a total of 48 embryos for each of the Cold, Warm, and Temp populations (table 1).

DNA Library Preparation and Sequencing

Single embryo DNA extraction was as described previously (Dean et al. 2002; Langley et al. 2011). Briefly, embryonic DNA was extracted and amplified using the REPLI-g Midi kit (Qiagen, Valencia CA) according to the manufacturer's instructions. Embryos were homogenized in D2 buffer and, following a 10-min incubation period on ice, stop solution was added plus the amplification master mix. Amplification was performed for 16 h at 30 °C followed by heat inactivation at 65 °C for 3 min. DNA fragmentation was performed using NEBNext dsDNA Fragmentase (New England Biolabs, Ipswich, MA). All intermediate quantification and cleanup steps utilized the Quant-iT PicoGreen dsDNA Assay kit (ThermoFisher), and AMPure XP kit (Beckman Coulter, Indianapolis, IN), respectively. End repair, dA-tailing, and

adapter ligation were performed using standard in-house protocols. PCR enrichment steps used NEXTFLEX DNA barcoded adapters and primers (Applied Genomics, Waltham, MA) coupled with Phusion HS FLEX DNA polymerase (New England Biolabs, Ispwich, MA).

Barcoded DNA libraries were sequenced on three lanes using the Illumina HiSeq2000 (48 samples per lane) producing 100-bp paired-end reads. On an average, 1 μg of each sample was loaded on their respective lanes and run with standard Illumina protocols.

Mapping and Variant Calling

Raw FASTQ files were preprocessed using custom Perl scripts to allow for identification of read pairs. These reads were then mapped to the *D. melanogaster* genome (Flybase, r6.22) using the MosaikAligner suite of tools (v.2.2.3; Lee et al. 2014). A jump database was created from the *D. melanogaster* genome using a hash size of 13 and FASTQ files were processed using MosaikBuild with the following parameters: -st Illumina -mfl 250. Additionally, all samples were given a unique read and group id. Resulting files were aligned using MosaikAligner with the following parameters: -hs 13 -mm 4 -a all -m unique -mhp 100 -act 20. Resulting BAM files containing only uniquely aligned reads were further processed using SAMtools (v.1.9) to produce coordinate sorted files. Duplicate reads were removed from each BAM file with the Picard Toolkit (v.2.18.14; Broad Institute, http://broadinstitute.github.io/picard/), using MarkDuplicates and setting "remove_duplicates" to "true." On an average, 72% of read pairs were uniquely mapped. Five samples from the cold-regime population and two samples from the warm-regime population exhibited poor alignment percentages and were not utilized in downstream variant calling.

Variant calling was performed using both FreeBayes software (v.1.2.0; Garrison and Marth 2012) which uses a Bayesian statistical framework that considers multiallelic loci and nonuniform copy number across samples (Marth et al. 1999; Garrison and Marth 2012), and the Joint Genotyper for Inbred Lines (JGIL) software (v.1.6; Stone 2012), which uses a maximum likelihood algorithm to obtain estimates that maximize the joint probability of the read data across all populations with given parameters. Processed BAM files, and the *Drosophila* genome reference (Flybase, r6.22) served as input to FreeBayes, which was run with the "-ploidy" value set to "1." The resulting variant call file (VCF) served as input to downstream postprocessing with the FreeBayes vcflib and bcftools (v.1.9; Li 2011) software. Briefly, VCF files were filtered to exclusively include SNPs with a phred-encoded quality score >15. Additionally, any SNP within 10 bp of an indel were also removed, and any sites with more than one minor allele (likely stemming from sequencing errors) were filtered. VCF files were then split by chromosome arm into separate tab-delimited files listing SNP read counts for each sample. Custom Perl scripts were used to produce tall-array format text files containing read counts and coordinates for every SNP for each sample. Additional Perl scripts were then used produce a SNP "observation file" that groups all samples based on the temperature regime in which they evolved ("Warm," "Cold," or "Temp") and lists the total read count for both the reference and minor allele in each regime for each SNP. For JGIL, processed BAM files likewise served as input. JGIL was run with the inbreeding generation count set to 50, the base quality threshold at 15, and output format to tall-array. The resulting tall-array files were likewise converted into observation files using Perl scripts.

Population Comparisons

Samples grouped together according to the temperature regime in which they evolved (Warm, Cold, and Temp; see Materials and Methods) are treated as a single population in this study. For simplicity, samples grouped in this manner are referred to as "populations" (Cold population, Warm population, Temp population) in the proceeding text. In order to glean population-specific data, we assessed total coverage across the three populations. Observation files containing total SNP observations across the three populations were further filtered using standard Linux shell tools to only include SNPs with at least 100 reads across each of the three temperature regimes. We also removed SNPs with total observation counts exceeding 750 as this likely represents read coverage bias based on the overall distribution of SNP coverage for each population (supplementary fig. S1, Supplementary Material online). Results from both JGIL and FreeBayes showed strong overlap in observed SNPs (see Results). We therefore limited our analysis to overlapping SNPs from both methods.

Because we are combining all the reads for a given temperature regime, we used Fisher's exact test to identify regions with significant differences in allele frequency. We recognize that the Cochran–Mantel–Haenszel (CMH) test (Kofler et al. 2011) has arguably become an industry standard since, in contrast to Fisher's exact test, it considers independent biological replicates in its contingency tables (Vlachos et al. 2019). However, this test cannot be applied to our data given its underlying structure. Our observation files were converted into a format suitable for Fisher's test using Popoolation2 (Kofler et al. 2011), and utilizing allele counts from both JGIL and Freebayes for each overlapping SNP. Resulting *P* values are reported in $-\log_{10}$ format. We performed tests centered on three pairwise population comparisons: Cold versus Warm, Cold versus Temp, and Warm versus Temp.

In addition, we supplemented this analysis using a sliding window approach using methods previously described in Burke et al. (2010). Using the windowscanr (v.0.1; Tavares 2019) and boot (v.1.3-24; Canty and Ripley 2020) packages implemented in R, we calculated in each window the mean and quantile score that 5% of *P* values exceeded using a

window size of 100 kb and step size of 10 kb. As windows with low SNP density could cause excess variation in the metric, especially in areas with low coverage (i.e., centromeres), we filtered out windows with SNP counts less than the value that 95% of all counts exceeded across all windows of a chromosome arm. This resulted in 2,349, 2,528, 2,810, 3,207, and 2,354 windows, respectively, for chromosomes 2L, 2R, 3L, 3R, and X. A significance threshold was calculated for each comparison and chromosome by taking the standard deviation of all quantiles that 5% of *P* values exceeded for 100 bootstrap replicate samples in each window and computing a metric that represented the quantile of this value at the 75th percentile. The product of this metric and the quantile representing qnorm(0.999) in R was added to the median value of quantile scores across all windows to compute a threshold which represents a 99.9% upper bound (Burke et al. 2010).

Heterozygosity within each population was calculated as *2pq* with allele frequencies representing coverage for each allele divided by total coverage for the SNP in question. The population-specific metrics, nucleotide diversity, and Tajima's *D* were calculated using the PopGenome package in R (v.2.7.5; Pfeifer et al. 2014) in the following manner: a VCF file containing the final set of SNPs following filtering (see Results) was split into individual VCF files, each one representing one of the five chromosome arms. VCF files were loaded into PopGenome which then grouped all samples into their corresponding populations. Nucleotide diversity was calculated over sliding windows (100 kb; 10 kb step size) using the built-in *diversity.stats* function. In the case of missing sites, PopGenome ignores those positions and calculates nucleotide diversity for the valid nucleotides. All nucleotide diversity values were divided by the number of sites per window (100,000) to obtain per-site nucleotide diversity within a population. Tajima's *D* values were likewise calculated using the built-in *neutrality.stats* function over sliding windows.

Identification of Candidate Loci

The quantile value score representing the 95th percentile of all *P* values was computed for all SNPs for each comparison and SNPs with values exceeding this threshold were retained. Compared with the metric used in our sliding window analysis, this threshold is less stringent and takes into consideration all SNPs across all chromosomes. Retained SNPs that fell within candidate genes were identified by creating a database of gene identifiers and coordinates using FlyMine (v.46.0; July 2018; Lyne et al. 2007). Custom Perl scripts were then used to match SNPs to their corresponding genes. Genes with multiple SNPs were only counted once to create a list of candidate genes for each comparison. Strong overlap of SNPs with *P* values exceeding this threshold was observed when *P* values based on allele frequencies generated by Freebayes or JGIL were compared. We therefore limited our analysis to these overlapping SNPs when identifying associated candidate loci. Candidate genes were further filtered to only include those with annotated expression in the ovaries for downstream analysis. Additionally, as part of our downstream analysis, we also filtered ovary annotated genes to only include those with known roles in DNA binding.

Variant Annotation

Identified variants were annotated using SnpEff (v. 4.3; Cingolani et al. 2012) using annotation from release *D. melanogaster* 5.86. We considered variants within exons, introns, UTRs within a single gene. Regions more than 500 bp upstream or downstream were considered intergenic.

Gene Ontology Enrichment Analysis and Overlap Plots

Gene ontology (GO) enrichment analysis was performed using GOWINDA (v. 1.12; Kofler and Schlotterer 2012). Compared with standard analysis of GO, GOWINDA corrects for the number of SNPs associated with a gene of interest. Variation in SNP counts associated with genes affects the probabilities of individual genes being sampled with genes having higher SNP counts clustering together. We used gene annotations based on *D. melanogaster* release 6.22 (Flybase) and set the minimum number of genes for each GO term to ten resulting in 3,430 potential ontologies for each comparison. The number of simulations was set to 100,000 and reported *P* values are FDR-corrected.

To identify genes in our data sets associated with DNA binding, a template query was made to Flymine to create a list of all *D. melanogaster* genes associated with the GO term "DNA binding." Genes in our data sets were then cross-checked against this list to create a modified data set containing only genes associated with this term. These data sets were further filtered to exclusively include zinc finger domain-containing candidates by cross checking against a list of zinc finger transcription factors generated through a query to Flybase (Flybase ID: FBgg0000729).

Upset plots generated with the UpsetR package in R were used to visualize both Freebayes and JGIL SNP overlaps and overlapping genes in our population comparisons (v.1.4.0; Conway et al. 2017).

Results and Discussion

Assessment of Sequencing and SNP Data

The focus of this study is on the identification of genomic loci underlying divergence in recombination rate among the experimentally evolved populations. Given our experimental design, it was important to establish a set of high confidence variants for downstream analysis. We therefore started out by assessing overall sample coverage of reads prior to variant calling and downstream analysis as part of standard quality

control. Likewise, we chose to use two methods of variant calling (JGIL and Freebayes) as we have the highest confidence in SNPs common to both methods since each utilizes a distinct statistical framework. Finally, it was important to establish a set of SNPs for which coverage was sufficient and uniform across our three temperature regimes.

We assessed read coverage among the samples to verify genomic coverage among the three populations. Among the samples from the Cold population, approximately 331 million reads were uniquely aligned corresponding to an average of 71% of total reads per sample, or 65% of total reads per sample when properly aligned read pairs were counted. For the Warm and Temp population samples, approximately 362 million and 382 million reads were properly aligned, respectively, corresponding to an average of 72% of total reads per sample in each of those regimes, and 68% of total reads when counting read pairs per sample in both regimes. Average sequencing depth across all samples was 5.2× per embryo corresponding to a population coverage of approximately 80× (supplementary table S1, Supplementary Material online).

We next examined the distributions of read coverage across the five major chromosome arms for each population. Comparisons of coverage patterns between each population, in which total reads from all samples were plotted following removal of duplicate reads, revealed a similar profile along the major chromosome arms (supplementary fig. S2*A*–*C*, Supplementary Material online). Among the three populations, a slight bias in read density was observed toward centromeric regions of chromosomes two and three, especially on the right arms. Despite this, coverage across the chromosomes was relatively uniform across all five arms. Likewise, the overall read density distribution for each chromosome was similar across the three populations, with a slightly narrower distribution among the Temp population indicating better coverage among sites (supplementary fig. S3, Supplementary Material online). In summary, the read coverage and chromosome distribution profiles of the three populations showed no major variation making the data suitable for downstream analysis.

Having established sufficient read coverage in all our included samples, we next sought to identify SNPs both common and consistent in coverage among the three populations utilizing Freebayes and JGIL (see Materials and Methods). Given the differing statistical frameworks utilized by these two programs, we have the highest confidence in SNPs common to both programs. Limiting our analysis to SNPs in the five major chromosome arms and chromosome 4, a total of 2,689,118 raw variants were identified using JGIL and 7,410,885 with FreeBayes. After filtering the FreeBayes data set to remove low-quality SNPs that failed to meet our criteria (see Materials and Methods), a total of 1,688,314 SNPs remained in the FreeBayes data set with strong overlap between it and the JGIL data set (fig. 1*A*). Average coverage for all SNPs ranged from 200× to 350× across the three populations for both filtered data sets and the average number of samples with coverage for each SNP was 114. In total, 850,298 SNPs were shared between the FreeBayes and JGIL across all three regimes which corresponded to one variant per 157 bases (fig. 1*B*). These numbers are similar to other studies of natural variation in North American of *D. melanogaster* (Singh et al. 2007; Sackton et al. 2009; Huang et al. 2014). Interestingly, 221,500 (26%) of our variants overlapped with those identified in an analysis of a large set of inbred lines comprising the *D. melanogaster* Genetic Research Panel (freeze2, Huang et al. 2014). This suggests that a subset of variants in our populations is common and widespread across other *D. melanogaster* populations. Approximately 52% of SNPs mapped to genes with 42% mapping to introns. In summary, we have the highest confidence in these 850,298 SNPs as they were identified using two different software methods, exhibit uniform and consistent coverage among our populations, and partially overlap previous data from other populations.

Identification of Divergent Alleles

We tested for divergence in SNP allele frequencies using Fisher's exact test within the following population comparisons: Cold versus Warm, Cold versus Temp, Warm versus Temp. Although the lack of sequencing data from the founders precludes an analysis of SNPs that may have emerged during the experimental evolution phase, our analysis is still expected to identify divergent genomic regions and candidate loci that contribute to differences in recombination rate between the three populations. For each comparison, we set a genome-wide threshold for statistical significance using the $-\log_{10}(P \text{ value})$ corresponding to the 95th percentile of all such values (fig. 1*C* and supplementary fig. S4*A*–*E*, Supplementary Material online). A large overlap was observed in SNPs with genome-wide significant differences in allele frequency, in each comparison, when either the JGIL or FreeBayes allele counts were utilized (fig. 2*A*–*C*) enabling us to limit our analysis to common SNPs in each comparison with high confidence. This corresponded to 38,070 (4.4%), 37,385 (4.3%), and 37,321 (4.3%) SNPs in the respective Cold versus Warm, Cold versus Temp, and Warm versus Temp comparisons that exceeded this threshold.

Identification of Divergent Regions

To identify divergent genomic regions among the three populations, we also carried out a sliding window analysis of our Fisher's test results (see Materials and Methods). Compared with individual tests, this method allows for the identification of divergent regions due to linkage. Our analysis identified multiple regions of divergence relative to genomic background in all three comparisons (fig. 3). When examining the Cold versus Warm comparison, large areas of divergence

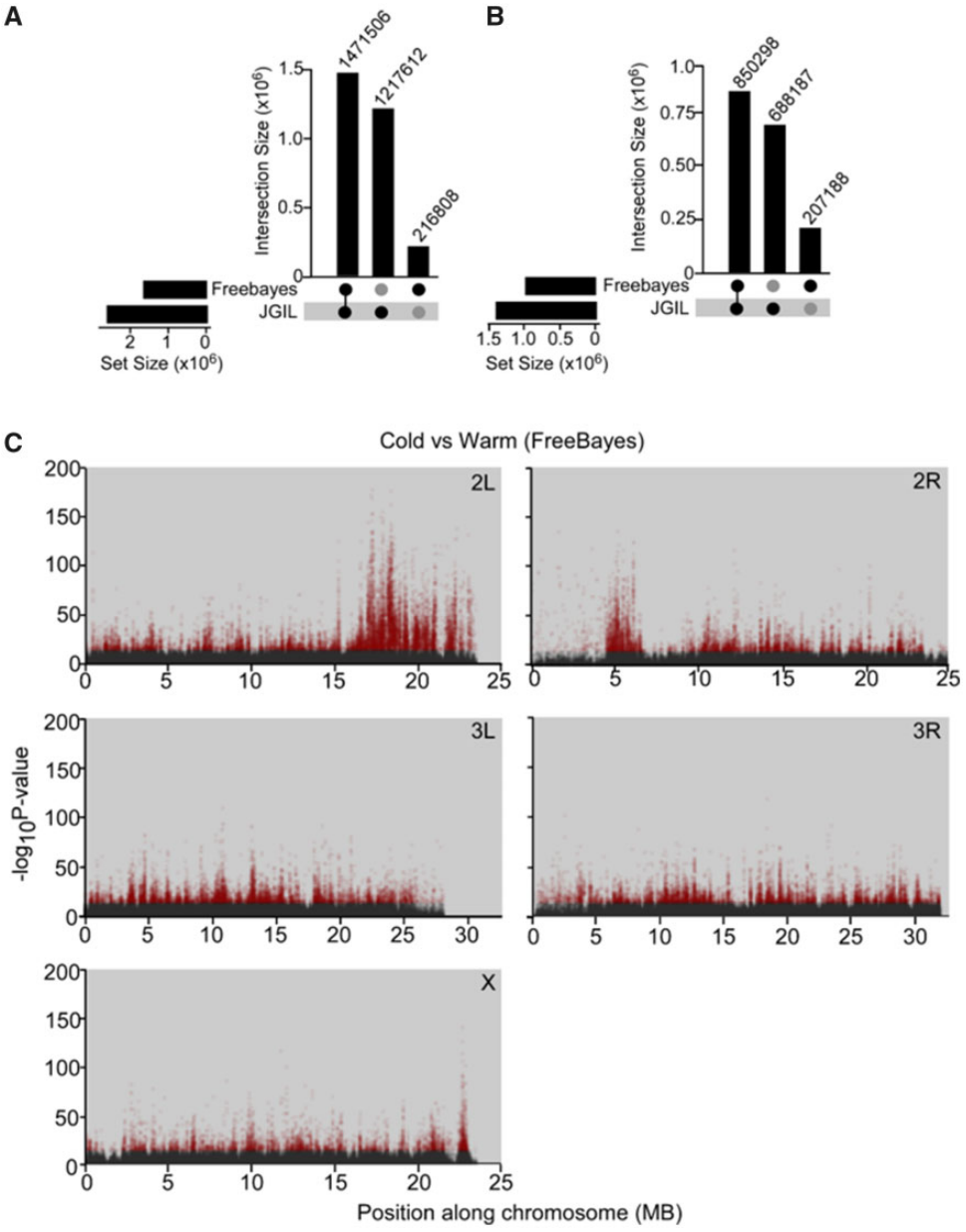

Fig. 1..—UpSet plots depicting overlap of SNPs detected by JGIL and Freebayes after (*A*) filtering to remove low quality SNPs, and (*B*) following a final filtration to remove SNPs failing to meet coverage criteria in all three populations. Numbers SNPs shared between methods are specified in the upper bar plot and methods used (JGIL or Freebayes) are indicated by dark circles below the bar chart. (*C*) *P* values in $-\log_{10}$ format (Fisher's exact test) when comparing allele frequencies between the Cold and Warm populations using allele frequencies generated by FreeBayes plotted along the five major chromosomes. Red points depict values exceeding the quantile corresponding to the 95th percentile of *P* values across all chromosomes. Black points depict values below this threshold.

with peaks exceeding genomic background threshold were observed toward the centromeric regions of chromosomes two and three with a more homogenous distribution of divergent regions in the X chromosome. We compared this result with that obtained when differences in allele frequencies between two random subsets each representing half of all samples within the three populations, Cold, Warm, and Temp, were evaluated against the respective Cold versus Warm, Cold versus Temp, and Warm versus Temp groups (fig. 3, blue lines). As expected, far fewer peaks exceeding threshold were observed when the two random subsets of samples within a population were compared (fig. 3, blue lines). Statistical analyses comparing the proportion of windows exceeding threshold between these comparisons likewise revealed a significant difference for each chromosome arm (supplementary table S2, Supplementary Material online).

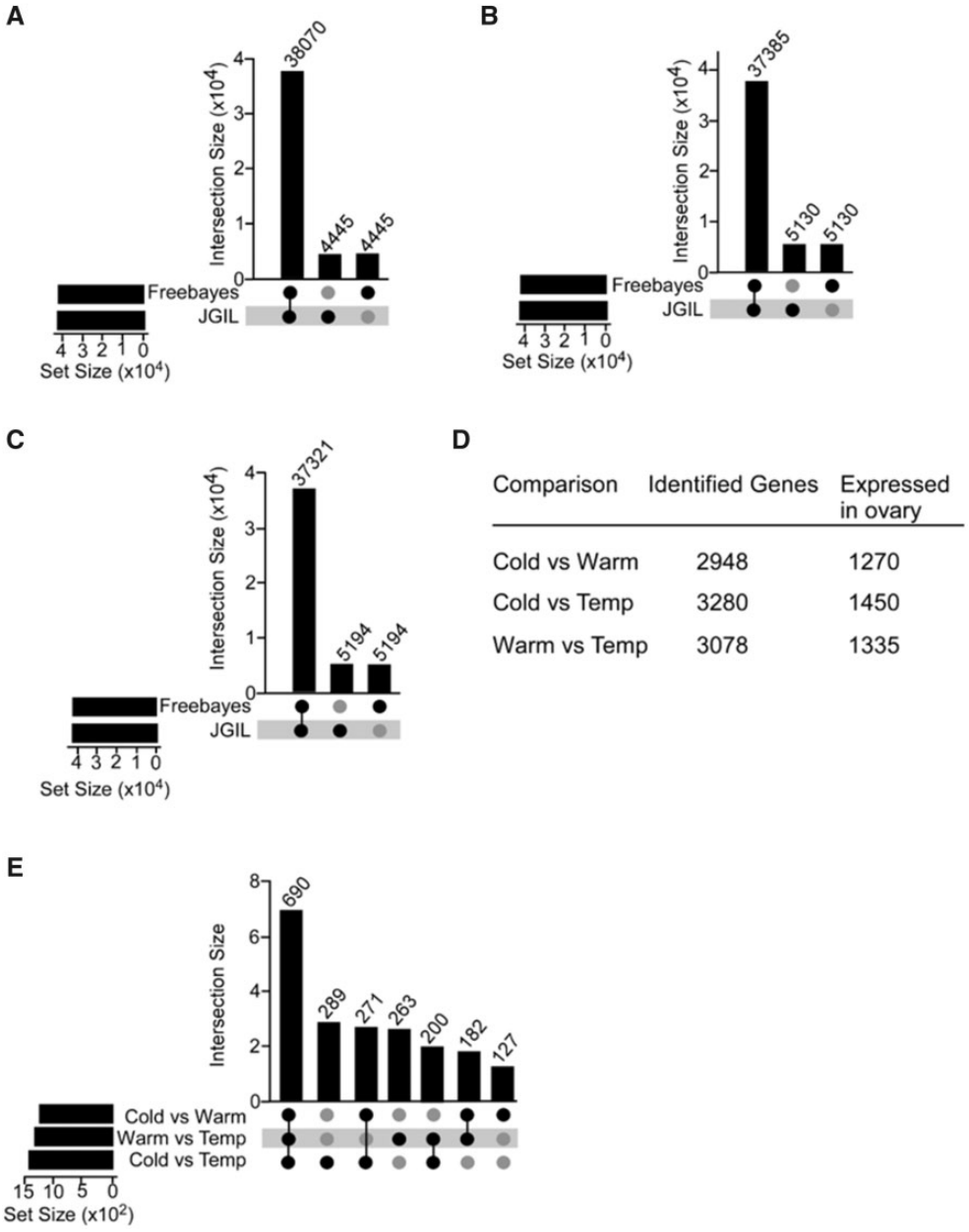

Fig. 2.—SNPs and associated gene counts exceeding the significance threshold representing the quantile corresponding to the 95th percentile of $-\log_{10}P$ values in the three population comparisons using allele frequencies generated by FreeBayes and JGIL. UpSet plots depicting overlap of SNPs exceeding this threshold in the (*A*) Cold versus Warm, (*B*) Cold versus Temp, and (*C*) Warm versus Temp population comparisons. Numbers of shared SNPs between comparisons are specified in the upper bar chart and comparisons are indicated by dark circles below the bar chart. (*D*) Table depicting counts for genes associated with these SNPs for each comparison overall and for those with annotated expression in the ovary. (*E*) UpSet plot depicting overlap of genes associated with SNPs exceeding the significance threshold (see Results) for the three population comparisons (Cold vs. Warm, Cold vs. Temp, Warm vs. Temp). Numbers of shared genes between comparisons are specified above the upper bar chart and comparisons are indicated by dark circles below the bar chart.

Interestingly, a relatively high number of peaks exceeding genomic background was observed in the random Warm population subset comparison (Warm vs. Temp comparison, blue line), indicating higher levels of genetic variation within this population of experimentally evolved lines and perhaps divergence among replicate lines. Nevertheless, divergence at the level of the temperature regime is still much more prominent qualitatively, especially toward the centromeric regions. Previous data showed the largest difference in recombination rate when comparing the Warm and Cold

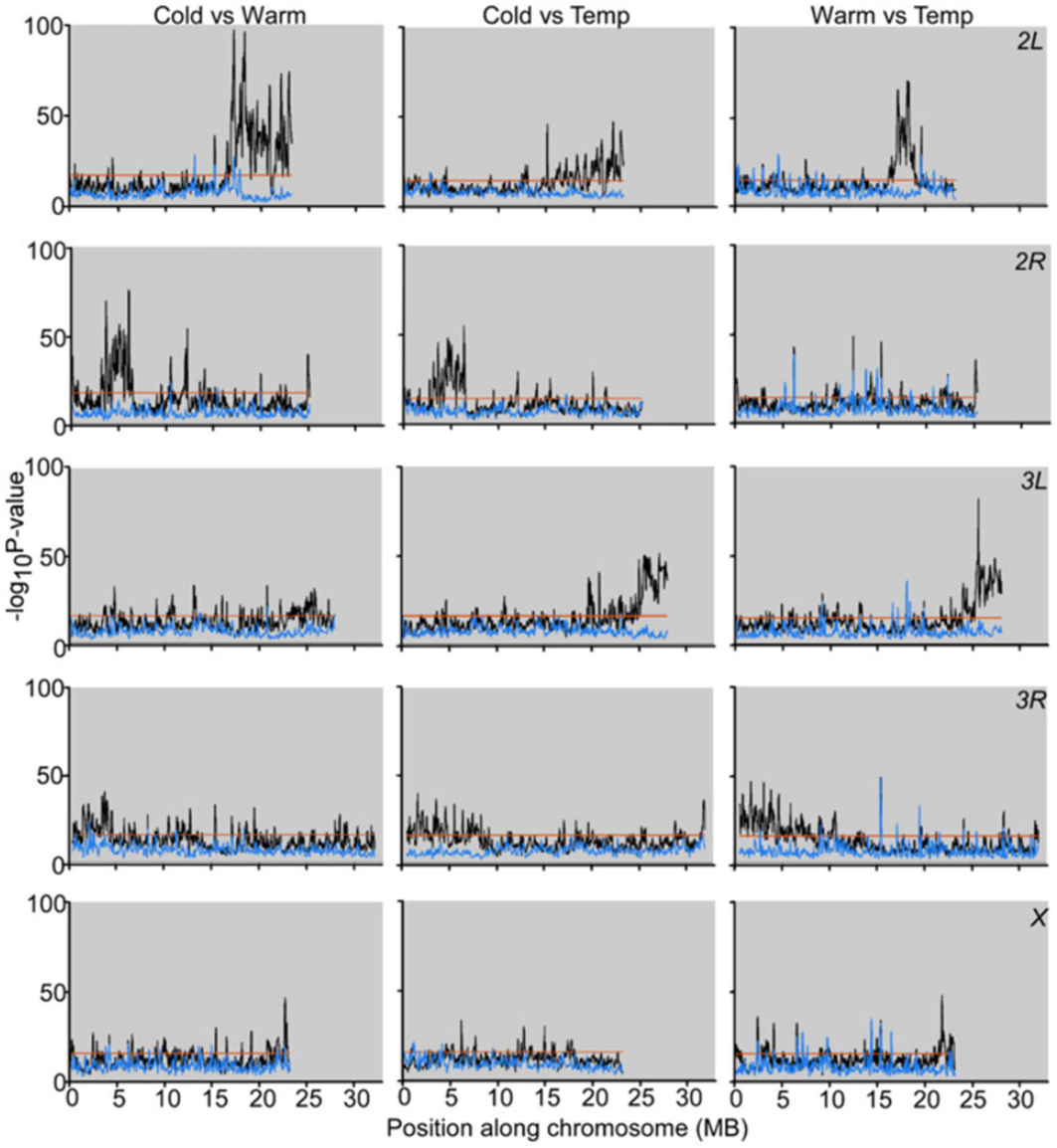

Fig. 3.—Sliding window plots (100-kb window; 10-kb step size) depicting allele frequency differences between the Cold and Warm populations (left column), Cold and Temp populations (middle column), and Warm and Temp populations (right column), for the five major chromosome arms (rows). Black line depicts $-\log_{10}P$ value based on Fisher's exact test corresponding to the 95th percentile for each window along the chromosome. Red lines depict genomic background threshold at 0.1% probability (see Materials and Methods). Blue lines depict the same scores when a random subset representing half of the Cold (first column), Temp (second columns), or Warm (third column) population allele frequencies are compared with that of the remaining subset.

populations (Kohl and Singh 2018). The most striking areas of divergence were likewise noted in this comparison (compare fig. 3, Cold vs. Warm to Cold vs. Temp and Warm vs. Temp).

We also examined nucleotide diversity (π), heterozygosity, and Tajima's *D* in sliding windows for the three populations. Reductions in these metrics are expected to flank nearby areas undergoing positive selection. Localized reductions in mean heterozygosity were noted among the three populations across all five chromosome arms with areas of distinct differences in heterozygosity noted between the populations (fig. 4*A*). The most striking areas exhibiting reduced heterozygosity were again focused among centromeric regions particularly on the right arms of chromosomes two and three within all three populations. Marked differences in heterozygosity at the level of populations were especially noted in these same areas which seemed to correlate with areas previously noted as having high divergence within the population comparisons (compare figs. 3 and 4*A*).

We note three distinct areas on chromosome 2L where mean heterozygosity approaches zero in the Warm population, one distinct area on chromosome 3L where mean heterozygosity approaches zero in the Temp population, and a

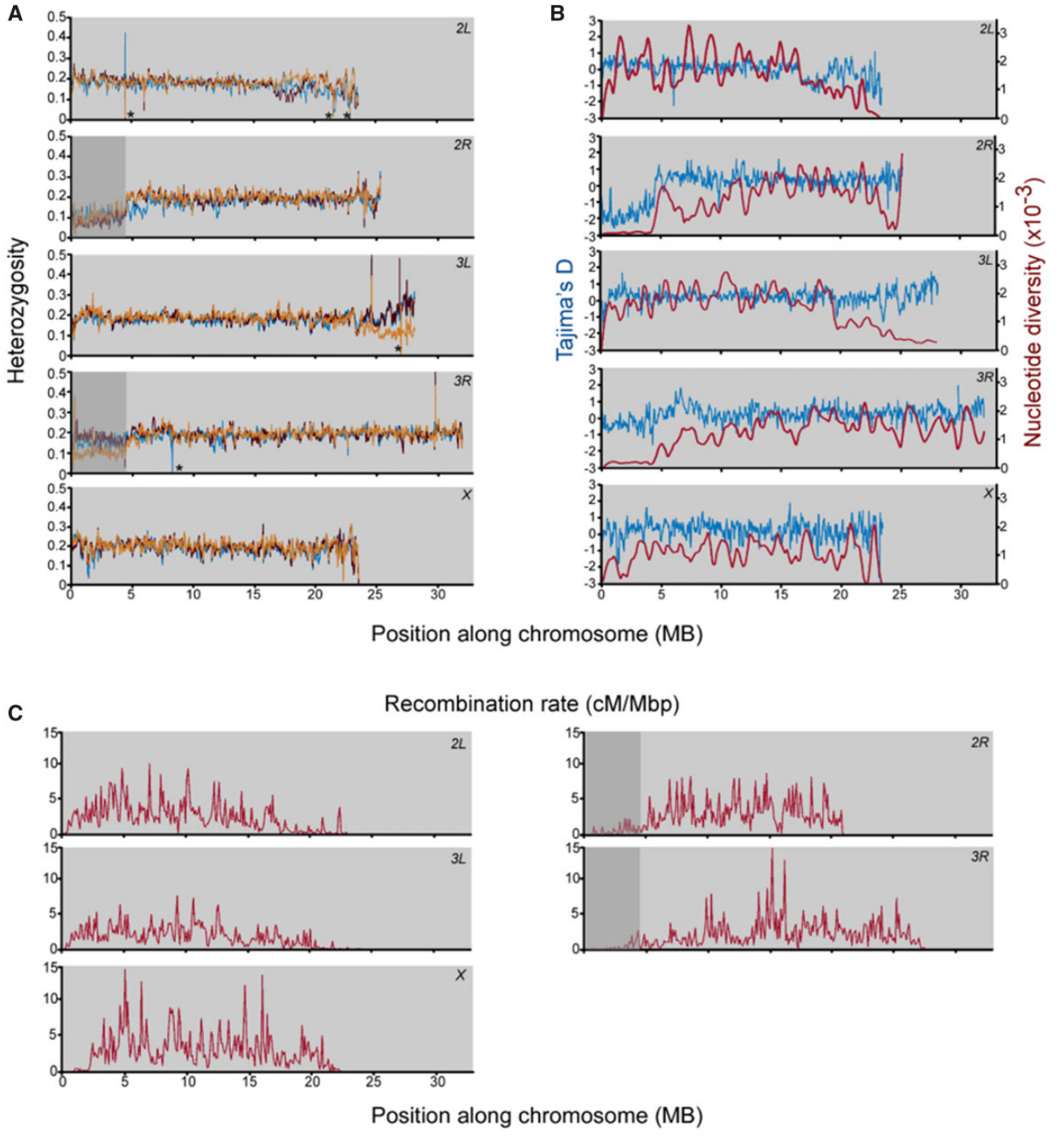

Fig. 4.—(*A*) Heterozygosity within the three populations across the genome. Sliding window plots (100-kb window; 10-kb step size) depicting mean heterozygosity across the five major chromosome arms for the Cold-regime population (blue line), Warm-regime population (dark red line) and Temp-regime population (orange line). Dark gray shading depicts peri-centromeric regions with general reductions in mean heterozygosity in all three populations. Asterisks denote areas where mean heterozygosity approaches zero for a temperature-regime population. (*B*) Tajima's *D* and nucleotide diversity. Sliding window plot (100-kb window size; 10-kb step) of Tajima's *D* (blue) and nucleotide diversity (π) (dark red) values plotted along the five major chromosome arms for the Warm population. Nucleotide diversity plots are smoothened using loess regression. (*C*) Recombination rate estimates across the *Drosophila melanogaster* genome. Recombination rate calculator (RRC) software was utilized to estimate recombination rate in 100-kb intervals (see Comeron et al. [2012]). Recombination rates (cM/Mb) based on available data are depicted across the five major chromosome arms. Gray shading depicts centromeric areas with comparatively low rates of recombination that strongly overlap areas of reduced heterozygosity and Tajima's *D* within the three temperature-regime populations.

final area on chromosome 3R where mean heterozygosity approaches zero in the Cold population suggesting possible fixation in these populations for those regions (fig. 4*A* asterisks). A sharp contrast to the Cold population is noted in the first instance on chromosome 2L. However, no other areas approaching zero are noted.

A similar pattern of reduction in nucleotide diversity and Tajima's *D* was observed in the Warm (fig. 4*B*) and Temp populations (supplementary fig. S5*A*, Supplementary Material online); general reductions in Tajima's *D* values were noted toward centromeric regions of chromosome two and three in these regimes in contrast to the Cold-regime population (supplementary fig. S5*B*, Supplementary Material online). All three populations regardless of temperature regime showed very similar patterns of nucleotide diversity (compare fig. 4*B* and supplementary fig. S5*A* and *B*, Supplementary Material online). As these patterns were most notable in the Warm and Temp populations compared with the Cold and seemed to overlap previously noted areas of allele frequency-divergence within the population comparisons (fig. 3), this data would suggest that these areas in the Temp and especially Warm populations may have experienced a stronger genomic response to the Warm temperature regime.

Considering that recombination is generally suppressed in centromeric regions in *D. melanogaster* (Beadle 1932; reviewed in Lindsley and Sandler 1977), it is tempting to speculate that reductions in these metrics might arise more readily in these regions and contribute to the observed patterns of divergence. To this end, we used the latest release of the *D. melanogaster* recombination rate calculator (v2.3; https://petrov.stanford.edu/cgi-bin/recombination-rates_updateR5.pl; last accessed December 13, 2020; Comeron et al. 2012) to estimate recombination rates across the five major chromosome arms at 100-kb intervals. As expected, estimated recombination rates were especially suppressed on the right arms of chromosomes two and three toward the centromeres (fig. 4*C*). These areas of suppression overlapped our observed areas of reduced heterozygosity, Tajima's *D*, and nucleotide diversity (compare fig. 4*A*–*C*). In addition, an overlap was noted in observed areas of divergence in the pairwise temperature regime comparisons (compare figs. 3 and 4*C*). Lack of sequencing data from the founders, however, limits our ability to make empirical inferences about the evolution of these lines. For instance, it is highly likely that SNPs we identified were present in the founders but at unknown frequencies. In summary, the observed divergence in allele frequencies between the three populations, and local reductions in diversity statistics would suggest divergence of the SNPs initially segregating in the founders based on the temperature regime in which the populations evolved; especially in areas of reduced recombination.

Finally, despite uniform read coverage and SNP observations across the three regime-based populations, a slight bias in read coverage was noted toward centromere regions of the right arms of chromosomes two and three. These areas partially overlap with the observed areas of divergence in the three comparisons (supplementary fig. S2*A*–*C*, Supplementary Material online), as well as the observed reduced heterozygosity and nucleotide diversity noted in the populations. This would suggest that the ability to detect changes in these evolutionary parameters in these regions is substantially greater compared with other areas of the genome thereby skewing our results.

Identification of Candidate Loci

Having identified genomic regions with significant differences in allele frequency, we next sought to identify candidate loci underlying genetic differences in recombination rate between populations. Based on the experimental evolution study design, recombination rate could potentially vary based on several levels including the lines within each population (see Materials and Methods), among replicates within a given temperature regime and among temperature regimes in which the populations evolved (table 1). Here, we focused exclusively on the temperature regime in which the replicate populations evolved. SNPs were mapped to their corresponding genes using an annotated list of chromosome coordinates and gene ids (see Materials and Methods). Exclusion of SNPs with $-\log_{10}(P \text{ values})$ below the established threshold based on the 95th percentile (see above) in each comparison resulted in 2,948, 3,280, and 3,078 divergent genes in the respective Cold versus Warm, Cold versus Temp, and Warm versus Temp comparisons (fig. 2*D*).

To account for divergence in other phenotypes in addition to recombination rate (see Introduction), we focused almost exclusively on genes with annotated expression in the ovaries (fig. 2*D*). Filtering candidates in this manner resulted in 1,270 candidate genes in the Cold versus Warm comparison, 1,450 candidate genes in the Cold versus Temp comparison, and 1,335 candidate genes in the Warm versus Temp comparison. Substantial overlap was noted among the three comparisons of genes expressed in the ovary (fig. 2*E*). Overall analysis of genes in each comparison is summarized in supplementary table S3, Supplementary Material online. For each comparison, we also analyzed candidate genes unique to that comparison (supplementary table S4, Supplementary Material online), and candidates associated with nonsynonymous SNPs (supplementary table S5, Supplementary Material online).

Population Comparisons of Candidates

Within the 1,270 candidates of the Cold versus Warm comparison, we noted 437 SNPs (1.1% of SNPs that exceeded threshold) that were either putative nonsynonymous mutations, splice variants, or stop codons representing 271 candidate genes (supplementary table S5, Supplementary Material

online). Gene ontology analysis of the initial 1,270 candidates revealed 540 enrichment terms (supplementary table S3, Supplementary Material online). Among the top categories were terms associated with protein metabolism and oogenesis. It should be noted that since we limited our analysis to genes with annotated expression in the ovaries, a slight bias toward categories associated with oogenesis and related terms are to be expected. Analysis of the 271 candidate genes with putative protein modifying SNPs, revealed 100 significantly overrepresented terms with top categories associated with DNA and chromatin metabolism and binding (supplementary table S5, Supplementary Material online). As the greatest difference in recombination rates were noted between the Cold and Warm populations, genes in this category are of primary interest.

We were also interested in candidate genes unique to the Cold versus Warm comparison. Enrichment analysis of the 127 genes in this category revealed 61 significantly overrepresented terms with top terms strongly associated with DNA recombination and mitotic and meiotic cell cycle processes (supplementary table S4, Supplementary Material online). Genes associated with DNA recombination in this category included *spn-B*, the *mei-217/218* complex, *Fancm*, and *SMC5*. We also identified 16 nonsynonymous SNPs associated with 16 loci (supplementary table S4, Supplementary Material online). Most of these candidates are associated with nucleic acid metabolism, however, *Fancm* was also noted in this list. *Fancm* encodes a DNA helicase and plays a critical role in DNA DSB repair while acting to suppress crossover events in the context of mitosis and meiosis (Kuo et al. 2014), making this an intriguing candidate. Another gene in this list is *Tollo* which based on sequence homology may be a member of the Toll-like receptor (TLR) superfamily. TLRs are a critical component of the innate immune response to bacterium (reviewed in Ferrandon et al. 2007). Interestingly, data suggest that infection by pathogenic bacteria (Singh et al. 2015) and the bacteria *Wolbachia pipientis* is associated with an increase in recombination rate (Hunter et al. 2016; Singh 2019); however, it is not known whether the innate immune response to this bacterium in the ovaries is responsible for this effect. Although it is interesting to speculate that changes in innate immunity could lead to variation in recombination fraction, lack of information on bacterial infection in our populations precludes further speculation.

We also examined candidate loci stemming from comparisons of the Cold versus Temp populations and Warm versus Temp populations which resulted in 1,450 and 1,335 candidates, respectively (supplementary table S3, Supplementary Material online). Within these sets of candidates, 435 SNPs (1.1% of SNPs that exceeded threshold) representing 290 genes in the Cold versus Temp comparison, and 406 SNPs (1.0% of SNPs exceeding threshold) representing 263 genes in the Warm versus Temp comparison were found to encode putative nonsynonymous SNPs, splice variants, or stop codon polymorphisms (supplementary table S5, Supplementary Material online). Genes unique to each comparison and genes representing nonsynonymous SNPs in each comparison along with their corresponding GO analyses are summarized in supplementary tables S4 and S5, Supplementary Material online.

Notable in these different comparisons is *rad50* with nonsynonymous SNPs in the Cold versus Warm and Cold versus Temp comparisons. *rad50* encodes an exonuclease with critical roles in DSB repair via homologous recombination (Ciapponi et al. 2004; Gorski et al. 2004). Although expression of this protein is noted in females, it is not known whether this protein plays a role in DSB repair in the context of meiotic recombination.

We also examined SNP allele frequencies among the putative nonsynonymous SNPs identified in our comparisons. For this analysis, SNP allele frequencies were analyzed in each of the collective Cold, Warm, or Temp populations depending on the comparison in question (supplementary tables S4 and S5, Supplementary Material online). However, we also split each population into the five constituent population replicates comprising that population overall (table 1) and examined SNP allele frequencies in each replicate. Similar results were obtained in either case with fixation noted in very few candidate SNPs.

Refinement of Candidates

After refining our list to only include candidate genes with annotated expression in the ovary, we are still left with a large set of loci with wide-ranging functions outside the scope of recombination. For example, *phospholipase D* is noted as a potential candidate for regulating plasticity in cell membrane lipid composition in response to temperature in these populations (Cooper et al. 2012). Our own assay noted SNP allele frequency divergence in the Cold versus Temp and Cold versus Warm comparisons within this gene (supplementary tables S3 and S5, Supplementary Material online). We likewise suspect that many of our loci are likewise candidates in the context of thermal adaptation as average temperatures in the region where collection took place max out at 21 °C (https://climate.weather.gc.ca; last accessed December 13, 2020).

Given these factors, we sought to further refine our lists to only include candidates with annotated functions related to DNA binding or zinc finger transcription factors. Mammalian, *PRDM9*, for example, contains a C_2H_2 zinc finger array which plays a critical role in DNA hotspot-specific binding to facilitate recombination via histone trimethylation (Baudat et al. 2010). Filtering for genes associated with DNA binding resulted in 805, 919, and 837 candidates in the respective Cold versus Warm, Cold versus Temp, and Warm versus Temp comparisons (supplementary table S6, Supplementary Material online). When we further refined these lists to exclude candidates lacking zinc finger domains, we were left with 50 candidates in the Cold versus Warm and Warm versus

Table 2

Chromosome and Gene Symbol of Candidate Loci with Annotated Expression in the Ovary in Each Comparison (columns) That Were Also Identified to Encode Zinc Finger Domain-Containing DNA-Binding Proteins

Cold versus Warm		Cold versus Temp		Warm versus Temp	
Chromosome	**Symbol**	**Chromosome**	**Symbol**	**Chromosome**	**Symbol**
2L	*dbr*	2L	*dbr*	2L	*dbr*
2L	*ush*	2L	*ush*	2L	*ush*
2L	*bowl*	2L	*bowl*	2L	*CG3407*
2L	*CG3407*	2L	*CG3407*	2L	*CG13775*
2L	*CG4496*	2L	*CG13775*	2L	*CG4496*
2L	*prg*	2L	*prg*	2L	*prg*
2L	*ab*	2L	*ab*	2L	*ab*
2L	*crol*	2L	*CG9932*	2L	*crol*
2L	*CG9932*	2L	*wek*	2L	*CG9932*
2L	*noc*	2L	*her*	2L	*wek*
2L	*her*	2L	*BuGZ*	2L	*CG17328*
2L	*BuGZ*	2L	*CG17568*	2L	*her*
2L	*CG10431*	2L	*Hr38*	2L	*BuGZ*
2L	*CG17568*	2L	*Hr39*	2L	*CG10431*
2L	*Hr38*	2L	*CG31612*	2L	*CG17568*
2L	*Hr39*	2L	*tsh*	2L	*Hr38*
2L	*CG31612*	2R	*EcR*	2R	*EcR*
2L	*tsh*	2R	*jing*	2R	*jing*
2R	*EcR*	2R	*rgr*	2R	*rgr*
2R	*jing*	2R	*lola*	2R	*lola*
2R	*lola*	2R	*shn*	2R	*shn*
2R	*shn*	2R	*chn*	2R	*chn*
2R	*chn*	3L	*ERR*	2R	*CG8089*
2R	*CG8089*	3L	*ssp*	2R	*ken*
3L	*CG10147*	3L	*CG17359*	3L	*ssp*
3L	*CG6765*	3L	*Trl*	3L	*CG11560*
3L	*CG17359*	3L	*Eip75B*	3L	*CG7372*
3L	*Eip75B*	3L	*ftz-f1*	3L	*Eip75B*
3L	*ftz-f1*	3L	*CG11456*	3L	*ftz-f1*
3L	*CG11456*	3L	*Aef1*	3L	*Aef1*
3L	*Aef1*	3L	*Hr78*	3L	*jim*
3L	*jim*	3L	*jim*	3R	*MTA1-like*
3R	*rn*	3R	*MTA1-like*	3R	*rn*
3R	*dsx*	3R	*rn*	3R	*dsx*
3R	*CG14710*	3R	*dsx*	3R	*CG14710*
3R	*CG6808*	3R	*hb*	3R	*CG6808*
3R	*l(3)neo38*	3R	*CG6808*	3R	*l(3)neo38*
3R	*srp*	3R	*srp*	3R	*MBD-R2*
3R	*mld*	3R	*sqz*	3R	*mld*
3R	*zfh1*	3R	*mld*	3R	*zfh1*
3R	*ttk*	3R	*ttk*	3R	*ttk*
X	*Hr4*	X	*ovo*	X	*ovo*
X	*CG12236*	X	*CG12236*	X	*CG32767*
X	*CG9650*	X	*mamo*	X	*CG9650*
X	*mamo*	X	*disco-r*	X	*mamo*
X	*disco-r*	X	*zld*	X	*disco-r*
X	*CG1529*	4	*dati*	X	*zld*
4	*dati*	4	*pho*	X	*CG1529*
4	*pho*			4	*pho*

NOTE.—Loci in red are common across all three comparisons.

Temp comparison and 49 candidates in the Cold versus Temp comparison (table 2). Within these candidates, we identified 14, 10, and 17 loci, containing nonsynonymous SNPs in the respective comparisons (table 3). Several genes in this list have not been annotated including *Aef1* and *CG17568*. Examination of SNP allele frequencies in whole populations and individual replicates comprising a population revealed no fixation of any candidates. However, we note several sites for which the SNP allele frequency is zero in one population of a comparison, suggesting a role for that SNP in the contrasting population's difference in recombination rate (supplementary table S7, Supplementary Material online). For example, SNP allele frequencies at sites associated with the gene *Aef1* in the Cold population, but not the Warm and Temp populations, are zero which could imply a role for variants in this gene at those sites in the previously observed increased recombination rate in the Warm and Temp populations.

As a final step, we sought to identify SNPs associated with loci which were divergent across all three comparisons. An overlap of genes is noted among the three comparisons (fig. 2*E*), however, these overlaps are at the level of genes as opposed to SNPs. Overlaps at the SNP level were observed in a total of 78 genes with 58 genes associated with the term DNA binding (supplementary table S8, Supplementary Material online). Of those 58 genes, *ush*, *Hr38*, *EcR*, *lola*, and *mamo* contain zinc finger domains. Nonsynonymous SNPs were associated with *ncm*, *kon*, *CG34417*, and *mamo*. The recurrence of these genes in our data sets and their annotated functions make them ideal candidates in the regulation of recombination rate variation.*ush* encodes a cofactor that associates with the GATA transcription factor Pannier and plays a critical role in several cell-differentiation and developmental processes (Cubadda et al. 1997). In mammals, including humans, mutations in homologous friend of GATA (FOG) proteins are implicated in gonadal dysgenesis and sterility (Tevosian et al. 2002; Brauner et al. 2016).

Both *EcR* and *Hr38* encode nuclear hormone receptors for the insect steroid family ecdysone. Circulating levels of ecdysone in adult female *Drosophila* play a critical role in oogenesis (Buszczak et al. 1999). Along these lines, defects in oogenesis and fecundity due to environmental stressors including changes in nutrient availability, temperature, and social interactions have been linked to changes in the ecdysone titer in females (Rauschenbach et al. 2000; Gruntenko et al. 2003; Terashima and Bownes 2004; Meiselman et al. 2018). Furthermore, divergence in nonsynonymous SNPs is noted in our comparisons for *Eip75B*, which encodes an ecdysone-induced protein and for which knockdown is associated with a decrease in recombination rate (Hunter et al. 2016). These data would suggest a link between environment, fertility, recombination rate, and ecdysone signaling, however, a more

Table 3

Candidate Loci and Associated Nonsynonymous Variants for Genes Expressed in the Ovary and Previously Annotated to Encode Zinc Finger Domain-Containing DNA-Binding Proteins in the Cold versus Warm, Cold versus Temp, and Warm versus Temp Comparisons

Comparison	Chromosome	Site	Symbol	Variant	Putative Impact
Cold versus Warm	**2L**	**68395**	***dbr***	**c.581C>T**	**p.Ser194Leu**
	2L	16678421	*her*	c.971T>C	p.Val324Ala
	2L	19180774	*CG17568*	c.773C>A	p.Thr258Lys
	2L	19180805	*CG17568*	c.742C>A	p.Gln248Lys
	2L	19180882	*CG17568*	c.665C>T	p.Ala222Val
	2L	20713069	*Hr38*	c.401G>A	p.Gly134Asp
	2R	6123167	*EcR*	c.394T>C	p.Phe132Leu
	2R	10508895	*lola*	c.2365C>A	p.Leu789Met
	2R	10520628	*lola*	c.2013C>A	p.Asn671Lys
	2R	15147402	*CG8089*	c.354G>C	p.Glu118Asp
	3L	13998171	*CG17359*	c.1012T>C	p.Phe338Leu
	3L	17998691	*Eip75B*	c.441G>T	p.Met147Ile
	3L	18059399	*Eip75B*	c.186C>G	p.Ser62Arg
	3L	21436389	*Aef1*	c.1107C>A	p.His369Gln
	3L	21436444	*Aef1*	c.1162G>A	p.Ala388Thr
	3R	11620278	*CG14710*	c.69C>G	p.Ile23Met
	3R	11621603	*CG6808*	c.1213A>T	p.Thr405Ser
	X	13883147	*mamo*	c.2318C>T	p.Thr773Met
	X	21030274	*CG1529*	c.1522C>A	p.Pro508Thr
Cold versus Temp	2L	68395	*dbr*	c.581C>T	p.Ser194Leu
	2L	15767479	*wek*	c.693C>G	p.Asp231Glu
	2R	6123167	*EcR*	c.394T>C	p.Phe132Leu
	2R	6503610	*jing*	c.572C>T	p.Thr191Ile
	3L	12119081	*ssp*	c.646C>G	p.Leu216Val
	3L	12119112	*ssp*	c.677C>T	p.Thr226Ile
	3L	13998171	*CG17359*	c.1012T>C	p.Phe338Leu
	3L	14760593	*Trl*	c.817A>T	p.Asn273Tyr
	3L	21436389	*Aef1*	c.1107C>A	p.His369Gln
	3L	21436444	*Aef1*	c.1162G>A	p.Ala388Thr
	X	13883147	*mamo*	c.2318C>T	p.Thr773Met
Warm versus Temp	2L	3812368	*CG3407*	c.836C>G	p.Thr279Ser
	2L	16678421	*her*	c.971T>C	p.Val324Ala
	2L	16745093	*BuGZ*	c.1015T>A	p.Ser339Thr
	2L	19180882	*CG17568*	c.665C>T	p.Ala222Val
	2R	10508895	*lola*	c.2365C>A	p.Leu789Met
	2R	10520628	*lola*	c.2013C>A	p.Asn671Lys
	2R	10524037	*lola*	c.1960A>G	p.Thr654Ala
	2R	10525682	*lola*	c.1544G>A	p.Ser515Asn
	3L	12119081	*ssp*	c.646C>G	p.Leu216Val
	3L	12119112	*ssp*	c.677C>T	p.Thr226Ile
	3L	12120750	*CG11560*	c.603A>C	p.Glu201Asp
	3L	15620888	*CG7372*	c.851A>G	p.Lys284Arg
	3R	5647254	*MTA1-like*	c.17C>G	p.Pro6Arg
	3R	7308125	*rn*	c.203G>A	p.Gly68Asp
	3R	11621081	*CG14710*	c.814T>A	p.Tyr272Asn
	3R	11621603	*CG6808*	c.1213A>T	p.Thr405Ser
	3R	12355726	*MBD-R2*	c.418T>A	p.Ser140Thr
	3R	31728395	*ttk*	c.781A>T	p.Ser261Cys
	X	5064506	*ovo*	c.466G>A	p.Ala156Thr
	X	13883147	*mamo*	c.2318C>T	p.Thr773Met
	X	21030313	*CG1529*	c.1483G>C	p.Ala495Pro
	X	21030367	*CG1529*	c.1429C>G	p.Leu477Val

Table 4

Comparison to Genes Previously Identified in Mediating Recombination Rate in a Separate Population of Fly Lines (Hunter et al. 2016)

Genes overlapping Hunter et al. (2016) for each comparison					
Cold versus Warm		**Cold versus Temp**		**Warm versus Temp**	
Ubx	*CG33970*	*alph*	*Ubx*	*Oaz*	*Ubx*
Pk	*Oaz*	*CG33970*	*lola*	*bru2*	*pk*
Lola	*bru2*	*bru2*	*CG1273*	*CG4440*	*lola*
CG1273	*CG4440*	*CG4440*	*dpr6*	*cdi*	*dpr6*
dpr6	*cdi*	*cdi*	*MESR3*	*Eip75B*	*MESR3*
MESR3	*Eip75B*	*Eip75B*	*CG15365*	*Ptp61F*	*grp*
grp	*Ptp61F*	*Ptp61F*	*grp*	*jing*	*CG33970*
alph	*jing*	*jing*		*CG9650*	
	CG9650				
Associated nonsynonymous SNPs					
Comparison	Chromosome	Site	Symbol	Variant	Putative impact
Cold versus Warm	2L	12357227	*bru2*	c.638C>G	p.Ala213Gly
	2R	10508895	*lola*	c.2365C>A	p.Leu789Met
	2R	10520628	*lola*	c.2013C>A	p.Asn671Lys
	3L	1344496	*Ptp61F*	c.1479C>A	p.Asn493Lys
	3L	1417156	*Ptp61F*	c.695A>G	p.His232Arg
	3L	17998691	*Eip75B*	c.441G>T	p.Met147Ile
	3L	18059399	*Eip75B*	c.186C>G	p.Ser62Arg
Cold versus Temp	2R	6503610	*jing*	c.572C>T	p.Thr191Ile
	3L	1344496	*Ptp61F*	c.1479C>A	p.Asn493Lys
	3L	1412953	*Ptp61F*	c.1678T>A	p.Ser560Thr
	3R	29594308	*alph*	c.895C>G	p.His299Asp
	3R	29604521	*alph*	c.1104C>G	p.Ile368Met
Warm versus Temp	2R	10508895	*lola*	c.2365C>A	p.Leu789Met
	2R	10520628	*lola*	c.2013C>A	p.Asn671Lys
	2R	10524037	*lola*	c.1960A>G	p.Thr654Ala
	2R	10525682	*lola*	c.1544G>A	p.Ser515Asn

NOTE.—Genes common to both studies and associated nonsynonymous SNPs for the three comparisons (Cold vs. Warm, Cold vs. Temp, Warm vs. Temp).

in-depth analysis linking ecdysone signaling in the ovary to recombination rate is required.

Most intriguing in this set is *mamo* which encodes a BTB domain-containing chromatin-binding protein. Reductions in *mamo* expression in the female germline are associated with defects in meiosis during prophase I and abnormal karyosome structure. Furthermore, *mamo* appears to mediate its effect through the direct binding to specific DNA consensus sequences (Mukai et al. 2007; Hira et al. 2013). Given the appearance of *mamo* in all of our data sets and comparisons, it is tempting to speculate a role for this gene in recombination variation however, testing this hypothesis requires perturbation of *mamo* function in a manner that allows oogenesis to proceed intact.

Comparison to Previous Studies

Previous work examining recombination between visible markers on chromosome 3R and X in a different set of inbred lines revealed a prioritized set of 20 candidate genes with potential roles in regulating recombination rate (Hunter et al. 2016). If genetic variation underlying divergence in recombination between populations also contributes to variation within populations, it is expected that these candidates would likewise show up in our analyses. Examining our initial data set prior to filtering for genes with annotated ovary expression, we identified 17 of those 20 genes in at least one comparison (table 4) of which 13 had annotated expression in the ovaries. Of those 17 genes, nonsynonymous SNPs were noted in *lola*, *bru2*, *Ptp61F*, and *Eip75B* in the Cold versus Warm comparison, *jing*, *alph*, and *Ptp61F* in the Cold versus Temp comparison, and *lola* in the Warm versus Temp comparison (table 4) all of which are expressed in the ovaries. An examination of SNP allele frequencies in whole populations and individual replicates comprising whole populations likewise revealed no fixation of any of these SNPs (supplementary table S9, Supplementary Material online).

We also examined our data sets for differences in allele frequencies for SNPs that mapped to loci previously implicated to participate directly in meiotic recombination (reviewed in McKim et al. 2002; Hughes et al. 2018). Five genes were noted that fell into this category including *mei-218* and *ord* in the Cold versus Warm comparison, *c(2)M* and *mus312* in the Cold versus Temp comparison, and *c(3)G* in the Warm

versus Temp comparison. Both *c(2)m* and *c(3)G* were noted to contain putative nonsynonymous variants. The relatively small number of genes in this category in our analyses in comparison to the 25-plus genes that have been characterized would suggest that differences in recombination rate between the three populations are regulated by a distinct set of loci; the majority of which do not directly participate in the recombination process.

Conclusions

In summary, we have identified a fairly large number of candidate genes in our screen that potentially mediate recombination rate divergence between populations. Our data indicate that these genes are not core components of meiotic recombination machinery. Rather, our data indicate that genetic variation underlying divergence in recombination between populations is largely overlapping with genetic variation underlying variation in recombination within populations.

Several additional phenotypes have diverged among these populations including fecundity, developmental plasticity, thermal tolerance, and cell membrane composition (Cooper et al. 2012; Condon et al. 2014, 2015; Adrian et al. 2016; Le Vinh Thuy et al. 2016; Alton et al. 2017). Despite refinements of our data set to genes expressed in the ovaries and those with functions associated with DNA-binding or zinc finger transcription factors, we expect many of the remaining candidates may have diverged in the evolutionary context of these phenotypes as well as other traits associated with temperature. Moreover, we cannot discount the possibility that differences in allele frequency in our candidates stem from selection on other phenotypes related to thermal tolerance as demonstrated in earlier studies (Rolandi et al. 2018). However, our data are consistent with a role for these candidate genes in contributing to the observed divergence in recombination rate among experimental evolution regimes.

From a broader evolutionary perspective, it is difficult to determine whether these candidates are the result of direct selection of pre-existing recombination rate variation in response to temperature or result indirectly from selection of these other traits in response to different temperature regimes. This is further complicated by the broad pleiotropic roles of many of our candidates. Furthermore, the experimental design introduces a possible role for genetic drift in our associated allelic variation as differences in temperature can affect generation counts and population size and variation of these parameters were noted during the experimental evolution phase (Yeaman et al. 2010). Although genetic drift was largely ruled out in the context of thermal tolerance, cell size, and cell membrane composition (Cooper et al. 2012; Condon et al. 2015; Adrian et al. 2016), it was noted that drift may play a role in the observed variation in recombination rate at the level of replicate lines within a population (Kohl and Singh 2018). However, given that this study is focused exclusively on recombination rate variation at the pooled population level, a role for genetic drift seems less likely than direct or indirect selection. Regardless, functional validation of our candidates in the context of recombination rate variation is necessary. Finally, it should be noted that given our experimental design many of our candidates and pathways are likely distinct from those subject to selection in the context of recombination rate due other environmental factors such as nutrient availability (Reviewed in Modliszewski and Copenhaver 2017).

Future studies centered around the functional validation of the most intriguing candidates, such as *mamo*, *EcR*, and *Tollo,* through perturbation in function, expression of the alleles in question, quantification of expression levels, and antibody visualization of protein chromatin interactions during various stages of meiosis will allow us to determine the roles the candidates with higher potential play in recombination rate variation.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

Reviewers and editors provided helpful feedback that greatly improved the manuscript. The authors gratefully acknowledge Stephanie Ruzsa for generating the embryos and preparing the DNA libraries. The authors also thank Molly Burke providing valuable feedback on this manuscript. This work was supported in part by the National Science Foundation (Grant No. MCB-1412813) to N.D.S.

Data Availability

Raw FASTQ files have been deposited into the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra/docs/sradownload/) and are available under accession number (PRJNA605869). SNP observation files and scripts have been deposited into Figshare (https://doi.org/10.6084/m9.figshare.12940052.v1).

Literature Cited

Adrian AB, Comeron JM. 2013. The *Drosophila* early ovarian transcriptome provides insight to the molecular causes of recombination rate variation across genomes. BMC Genomics 14(1):794.

Adrian GJ, Czarnoleski M, Angilletta MJ Jr. 2016. Files evolved small bodies and cells at high or fluctuating temperatures. Ecol Evol. 6(22):7991–7996.

Alton LA, Condon C, White CR, Angilletta MJ Jr. 2017. Colder environments did not select for a faster metabolism during experimental evolution of *Drosophila melanogaster*. Evolution 71(1):145–152.

Barton NH. 1995. A general model for the evolution of recombination. Genet Res. 65(2):123–145.

Barton NH, Charlesworth B. 1998. Why sex and recombination? Science 281(5385):1986–1990.
Baudat F, et al. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327(5967):836–840.
Baudat F, Imai Y, Bd M. 2013. Meiotic recombination in mammals: localization and regulation. Nat Rev Genet. 14(11):794–806.
Beadle GW. 1932. A possible influence of the spindle fibre on crossing-over in *Drosophila*. Proc Natl Acad Sci U S A. 18(2):160–165.
Berg IL, et al. 2011. Variants of the protein PRDM9 differentially regulate a set of human meiotic recombination hotspots highly active in African populations. Proc Natl Acad Sci U S A. 108(30):12378–12383.
Brand CL, Cattani MV, Kingan SB, Landeen EL, Presgraves DC. 2018. Molecular evolution at a meiosis gene mediates species differences in the rate and patterning of recombination. Curr Biol. 28(8):1289–1295.
Brauner R, et al. 2016. Familial forms of disorders of sex development may be common if infertility is considered a comorbidity. BMC Pediatr. 16(1):195–204.
Brooks LD, Marks RW. 1986. The organization of genetic variation for recombination in *Drosophila melanogaster*. Genetics 114(2):525–547.
Burke MK, et al. 2010. Genome-wide analysis of a long-term evolution experiment in *Drosophila*. Nature 467(7315):587–590.
Buszczak M, et al. 1999. Ecdysone response genes govern egg chamber development during mid-oogenesis in *Drosophila*. Development 126(20):4581–4589.
Canty A, Ripley BD. 2020. boot: bootstrap R (S-Plus) Functions. R Package Version. 1:3–24.
Chan AH, Jenkins PA, Song YS. 2012. Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. PLoS Genet. 8(12):e1003090.
Charlesworth B, Barton NH. 1996. Recombination load associated with selection for increased recombination. Genet Res. 67(1):27–41.
Chinnici JP. 1971. Modification of recombination frequency in *Drosophila*. II. The polygenic control of crossing over. Genetics 69(1):85–96.
Chowdhury R, Bois PRJ, Feingold E, Sherman SL, Cheung VG. 2009. Genetic analysis of variation in human meiotic recombination. PLoS Genet. 5(9):e1000648.
Ciapponi L, et al. 2004. The *Drosophila* Mre11/Rad50 complex is requird to prevent both telomeric fusion and chromosome breakage. Curr Biol. 14(15):1360–1366.
Cingolani P, et al. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain *w1118; iso-2; iso-3*. Fly (Austin) 6(2):80–92.
Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genet. 8(10):e1002905.
Condon C, Cooper BS, Yeaman S, Angilleta MJ Jr. 2014. Temporal variation favors the evolution of generalists in experimental populations of *Drosophila melanogaster*. Evolution 68(3):720–728.
Condon C, et al. 2015. Indirect selection of thermal tolerance during experimental evolution of *Drosophila melanogaster*. Ecol Evol. 5(9):1873–1880.
Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33(18):2938–2940.
Cooper BS, Hammad LA, Fisher NP, Karty JA, Montooth KL. 2012. In a variable thermal environment selection favors greater plasticity of cell membranes in *Drosophila melanogaster*. Evolution 66(6):1976–1984.
Cubadda Y, et al. 1997. *u-shaped* encodes a zinc finger protein that regulates the proneural genes *achaete* and *scute* during the formation of bristles in *Drosophila*. Genes Dev. 11(22):3083–3095.
Dean FB, et al. 2002. Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A. 99(8):5261–5266.
Dumont BL, Broman KW, Payseur BA. 2009. Variation in genomic recombination rates among heterogeneous stock mice. Genetics 182(4):1345–1349.
Ferrandon D, Imler J-L, Hetru C, Hoffman JA. 2007. The *Drosophila* systemic immune response: sensing and signaling during bacterial and fungal infections. Nat Rev Immunol. 7(11):862–874.
Fledel-Alon A, et al. 2011. Variation in human recombination rates and its genetic determinants. PLoS One 6(6):e20321.
Fuyama Y. 1984. Gynogenesis in *Drosophila melanogaster*. Jpn J Genet. 59(1):91–96.
Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907 [q-bio.GN]. Available from: https://arxiv.org/abs/1207.3907.
Gorski MM, et al. 2004. Disruption of *Drosophila* Rad50 causes pupal lethality, the accumulation of DNA double-strand breaks and the induction of apoptosis in third instar larvae. DNA Repair (Amst). 3(6):603–615.
Gruntenko NE, Bownes M, Terashima J, Sukhanova MZh, Raushenbach IY. 2003. Heat stress affects oogenesis differently in wild-type *Drosophila virilis* and a mutant with altered juvenile hormone and 20-hydroxyecdysone levels. Insect Mol Biol. 12(4):393–404.
Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. Genet Res. 8(3):269–294.
Hinch AG, et al. 2011. The landscape of recombination in African Americans. Nature 476(7359):170–175.
Hira S, et al. 2013. Binding of *Drosophila* maternal Mamo protein to chromatin and specific DNA sequences. Biochem Biophys Res Commun. 438(1):156–160.
Huang W, et al. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* genetic reference panel lines. Genome Res. 24(7):1193–1208.
Hughes SE, Miller DE, Miller AL, Hawley RS. 2018. Female meiosis: synapsis, recombination, and segregation in *Drosophila melanogaster*. Genetics 208(3):875–908.
Hunter CM, Huang W, Mackay TFC, Singh ND. 2016. The genetic architecture of natural variation in recombination rate in *Drosophila melanogaster*. PLoS Genet. 12(4):e1005951.
Johnston SE, Bérénos C, Slate J, Pemberton JM. 2016. Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (*Ovis aries*). Genetics 203(1):583–598.
Kidwell MG. 1972. Genetic change of recombination value in *Drosophila melanogaster*. I. Artificial selection for high and low recombination and some properties of recombination-modifying genes. Genetics 70(3):419–432.
Kliman RM, Hey J. 1993. Reduced natural selection associated with low recombination in *Drosophila melanogaster*. Mol Biol Evol. 10(6):1239–1258.
Kofler R, Pandey RV, Schlöter C. 2011. Popoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27(24):3435–3436.
Kofler R, Schlotterer C. 2012. Gowinda: unbiased analysis of gene set enrichment for genome wide association studies. Bioinformatics 28(15):2084–2085.
Kohl KP, Singh ND. 2018. Experimental evolution across different thermal regimes yields genetic divergence in recombination fraction but no divergence in temperature associated plastic recombination. Evolution 72(4):989–999.
Kong A, et al. 2014. Common and low-frequency variants associated with genome-wide recombination rate. Nat Genet. 46(1):11–16.
Kuo HK, McMahan S, Rota CM, Kohl KP, Sekelsky J. 2014. *Drosophila* FANCM helicase prevents spontaneous mitotic crossovers generated by MUS81 and SLX1 nucleases. Genetics 198(3):935–945.

Lambing C, Franklin CH, Wang C-JR. 2017. Understanding and manipulating meiotic recombination in plants. Plant Physiol. 173(3):1530–1542.
Langley CH, Crepeau M, Carden C, Corbett-Detig R, Stevens K. 2011. Circumventing heterozygosity: sequencing the amplified genome of a single haploid *Drosophila melanogaster* embryo. Genetics 188(2):239–246.
Le Vinh Thuy J, VandenBrooks JM, Angilletta MJ Jr. 2016. Developmental plasticity evolved according to specialists-generalists trade-offs in experimental populations of *Drosophila melanogaster*. Biol Lett. 12(7):20160379.
Lee W-P, et al. 2014. MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping. PLoS One 9(3):e90581.
Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27(21):2987–2993.
Lindsley DL, Sandler L. 1977. The genetic analysis of meiosis in female *Drosophila melanogaster*. Philos T R Soc B. 277(955):295–312.
Lyne R, et al. 2007. Flymine: an integrated database for *Drosophila* and *Anopheles* genomics. Genome Biol. 8(7):R129.
Ma L, et al. 2015. Cattle sex-specific recombination and genetic control from a large pedigree analysis. PLoS Genet. 11(11):e1005387.
Manzano-Winkler B, McGaugh SE, Noor MA. 2013. How hot are *Drosophila* hotspots? Examining recombination rate variation and associations with nucleotide diversity, divergence, and maternal age in *Drosophila pseudoobscura*. PLoS One 8(8):e71582.
Marth GT, et al. 1999. A general approach to single-nucleotide polymorphism discovery. Nat Genet. 23(4):452–456.
McKim KS, Jang JK, Manheim EA. 2002. Meiotic recombination and chromosome segregation in *Drosophila* females. Annu Rev Genet. 36(1):205–232.
Meiselman MR, Kingan TG, Adams ME. 2018. Stress-induced reproductive arrest in *Drosophila* occurs through ETH deficiency-mediated suppression of oogenesis and ovulation. BMC Biol. 16(1):18.
Modliszewski JL, Copenhaver GP. 2017. Meiotic recombination gets stressed out: CO frequency is plastic under pressure. Curr Opin Plant Biol. 36:95–102.
Mukai M, et al. 2007. MAMO, a maternal BTB/POZ-Zn-finger protein enriched in germline progenitors is required for the production of functional eggs in *Drosophila*. Mech Dev. 124(7–8):570–583.
Oliver PL, et al. 2009. Accelerated evolution of the *Prdm9* speciation gene across diverse metazoan taxa. PLoS Genet. 5(12):e1000753.
Ortiz-Barrientos D, Chang AS, Noor MAF. 2006. A recombinational portrait of the *Drosophila pseudoobscura* genome. Genet Res. 87(1):23–31.
Otto SP. 2009. The evolutionary enigma of sex. Am Nat. 174(S1):S1–S14.
Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: an efficient Swiss army knife for population genomic analysis in R. Mol Biol Evol. 31(7):1929–1936.
Rauschenbach IY, Sukhanova MZh, Hirashima A, Sutsugu E, Kuano E. 2000. Role of the ecdysteroid system in the regulation of *Drosophila* reproduction under environmental stress. Dokl Biol Sci. 375(1/6):641–643.
Ritz KR, Noor MAF, Singh ND. 2017. Variation in recombination rate: adaptive or not? Trends Genet. 33(5):364–374.
Rolandi C, Lighton JRB, de la Vega GJ, Schilman PE, Mensch J. 2018. Genetic variation for tolerance to high temperatures in a population of *Drosophila melanogaster*. Ecol Evol. 8(21):10374–10383.
Sackton TB, et al. 2009. Population genomic inferences from sparse high-throughput sequencing of two populations of *Drosophila melanogaster*. Genome Biol Evol. 1:449–465.
Sandor C, et al. 2012. Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. PLoS Genet. 8(7):e1002854.
Schwartz JJ, Roach DJ, Thomas JH, Shendure J. 2014. Primate evolution of the recombination regulator PRDM9. Nat Commun. 5(1):4370.
Singh N. 2019. Wolbachia infection associated with increased recombination in *Drosophila*. G3 (Bethesda) 9(1):229–237.
Singh ND, Aquadro CF, Clark AG. 2009. Estimation of fine-scale recombination intensity variation in the white-echinus interval of *D. melanogaster*. J Mol Evol. 69(1):42–53.
Singh ND, et al. 2015. Fruit flies diversify their offspring in response to parasitic infection. Science 349(6249):747–750.
Singh ND, Macpherson JM, Jensen JD, Petrov DA. 2007. Similar levels of X-linked and autosomal nucleotide variation in African and non-African populations of *Drosophila melanogaster*. BMC Evol Biol. 7(1):202.
Smukowski Heil CS, Ellison C, Dubin M, Noor MAF. 2015. Recombining without hotspots: a comprehensive evolutionary portrait of recombination in two closely related species of *Drosophila*. Genome Biol Evol. 7(10):2829–2842.
Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017. Variation in recombination frequency and distributions across eukaryotes: patterns and processes. Philos Trans R Soc B. 372(1736):20160455.
Stefansson H, et al. 2005. A common inversion under selection in Europeans. Nat Genet. 37(2):129–137.
Stone EA. 2012. Joint genotyping on the fly: identifying variation among a sequenced panel of inbred lines. Genome Res. 22:966–974.
Tavares H. 2019. windowscanr: apply functions using sliding windows. R package version 0.1.
Terashima J, Bownes M. 2004. Translating available food into the number of eggs laid by *Drosophila melanogaster*. Genetics 167(4):1711–1719.
Tevosian SG, et al. 2002. Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. Development 129(19):4627–4634.
True JR, Mercer JM, Laurie CC. 1996. Differences in crossover frequency and distribution among three sibling species of *Drosophila*. Genetics 142(2):507–523.
Valentin J. 2009. Characterization of a meiotic control gene affecting recombination in *Drosophila melanogaster*. Hereditas 75(1):5–22.
Vlachos C, et al. 2019. Benchmarking software tools for detecting and quantifying selection in evolve and resequencing studies. Genome Biol. 20(1):169.
Wang Y, Copenhaver GP. 2018. Meiotic recombination: mixing it up in plants. Annu Rev Plant Biol. 69(1):577–609.
Yeaman S, Chen Y, Whitlock MC. 2010. No effect of environmental heterogeneity on the maintenance of genetic variation in wing shape in *Drosophila melanogaster*. Evolution 64(12):3398–3408.

Associate editor: Baer Charles