Genome structure predicts modular transcriptome responses to genetic and environmental conditions

Stephanie Mark[1], Joerg Weiss[1], Eesha Sharma[2], Ting Liu[1], Wei Wang[1], Julie M. Claycomb[2], Asher D. Cutter[1#]

[1]Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, ON Canada
[2]Department of Molecular Genetics, University of Toronto, Toronto, ON Canada
[#]Corresponding author: asher.cutter@utoronto.ca

Abstract

Understanding the plasticity, robustness, and modularity of transcriptome expression to genetic and environmental conditions is crucial to deciphering how organisms adapt in nature. To test how genome architecture influences transcriptome profiles, we quantified expression responses for distinct temperature-adapted genotypes of the nematode *Caenorhabditis briggsae* when exposed to chronic temperature stresses throughout development. We found that 56% of the 8795 differentially-expressed genes show genotype-specific changes in expression in response to temperature (genotype-by-environment interactions, GxE). Most genotype-specific responses occur under heat stress, indicating that cold versus heat stress responses involve distinct genomic architectures. The 22 co-expression modules that we identified differ in their enrichment of genes with genetic versus environmental versus interaction effects, as well as their genomic spatial distributions, functional attributes, and rates of molecular evolution at the sequence level. Genes in modules enriched for simple effects of either genotype or temperature alone tend to evolve especially rapidly, consistent with disproportionate influence of adaptation or weaker constraint on these subsets of loci. Chromosome scale heterogeneity in nucleotide polymorphism, however, rather than the scale of individual genes, predominate as the source of genetic differences among expression profiles, and natural selection regimes are largely decoupled between coding sequences and non-coding flanking sequences that contain *cis*-regulatory elements. These results illustrate how the form of transcriptome modularity and genome structure contribute to predictable profiles of evolutionary change.

Keywords: Caenorhabditis, transcriptome, regulatory evolution, robustness, temperature, genotype-by-environment interactions, molecular evolution

Introduction

Evolutionary adaptation to varying environmental conditions starts with genetic variability, often with alternate alleles affecting gene regulation and expression. Consequently, understanding the plasticity, robustness, and modularity of transcriptome responses to genetic and environmental conditions is crucial to deciphering how organisms adapt in nature (Ungerer *et al.* 2007). Gene expression represents the most basic level at which phenotypic plasticity to a perturbation can manifest, and therefore underpins the degree of robustness of higher level phenotypes in response to the same perturbation (de Visser *et al.* 2003; Flatt 2005). Because the transcriptome changes in response both to extrinsic factors (e.g. environmental inputs) and to factors that are intrinsic to the organism itself (e.g. genetic background) (Gagneur *et al.* 2013; Grishkevich & Yanai 2013), we must consider both extrinsic and intrinsic contributions in the dynamism of genetic network composition and its genomic architecture. Consequently, it is crucial to determine how much of the genome is expressed differentially in a plastic manner with sensitivity to environmental conditions versus a genetically deterministic manner independently of environmental conditions versus a non-additive combination of both (Grishkevich & Yanai 2013; Knowles *et al.* 2017). Moreover, it remains generally unclear how modular are distinct gene expression responses and what characteristics of the genome predict their composition and molecular evolution. These questions frame some of the key outstanding issues in connecting transcriptome activity to environmental heterogeneity and the molecular evolution of genomes.

Temperature conditions represent a pervasive extrinsic, environmental factor that influences gene expression and can help reveal the relative roles of plasticity versus robustness of transcriptome profiles (Causton *et al.* 2001; Smith *et al.* 2013). If expression plasticity is adaptive, then we expect organisms to modulate their transcriptomes under chronic developmental exposure to heat or cold stress in a coordinated way to maintain fitness. However, homeostasis may break down at environmental extremes and lead to non-adaptive changes in gene expression that simply reflect a ‘broken’ biological system. Pathways associated with the heat shock response are implicated in physiological buffering to acute heat stress (Lindquist & Craig 1988), but chronic sublethal heat stress may not activate this same stress response. By characterizing profiles of transcriptome change to temperature conditions, we can test the robustness of plastic responses to genetic divergence that reflects genomic evolution in the control over gene expression.

Allelic differences can be thought of as a kind of perturbation, a genetic perturbation, that can expose the sensitivity of gene networks in terms of expression changes (Hecker *et al.* 2009). Expression modulated by *cis*-regulatory alleles may minimize adverse pleiotropic effects and, consequently, modest effects of *cis*-regulatory SNPs might only be pronounced when they accrue over long periods of time to give rise to the kinds of expression differences that accumulate between species (Carroll 2008; Stern & Orgogozo 2008; Wittkopp & Kalay 2012). By contrast, changes to *trans*-acting regulators like transcription factors may lead to many downstream pleiotropic consequences. Consequently, large *trans*-acting effects might make up a substantial fraction of the genetic variability for gene expression differences among individuals within a species and yet rarely contribute to expression differences between species (Wittkopp *et al.* 2004; Smith & Kruglyak 2008; Stern & Orgogozo 2008; Wittkopp *et al.* 2008; Tirosh *et al.* 2009), because most changes that affect fitness are deleterious and eventually get eliminated by natural selection (Keightley & Lynch 2003). The intermediate timescale of adaptive divergence between populations of the same species thus has the potential to expose whether distinct regulatory architecture must be invoked to describe transcriptome changes across the extremes of timescales from polymorphism within a single population to divergence between species.

In this context, extensive transcriptome analysis of the nematode *C. elegans* in response to heat shock and knock-out mutation began with microarrays (Kim *et al.* 2001), with more recent studies using recombinant inbred lines of wild strains to map polymorphic loci that contribute genotype-dependent responses to temperature (Li *et al.* 2006; Grishkevich *et al.* 2012; Snoek *et al.* 2017; Snoek *et al.* 2019). For example, Li et al. (2006) found that among 496 detectable expression quantitative trait loci (eQTL), *trans*-eQTL were nearly 8-times as likely as *cis*-eQTL to show genotype-by-temperature responses, with subsequent study reinforcing this pattern (Snoek *et al.* 2017). Moreover, eQTL are found disproportionately on SNP-dense chromosome arms in *C. elegans* (Rockman *et al.* 2010). Grishkevich et al. (2012) reported that constitutively-expressed genes in *C. elegans* tend to have short intergenic regions, consistent with simple *cis*-regulatory controls, and that genes with genotype-dependent expression or genotype-by-environment interactions have longer intergenic regions, consistent with complex *cis*-regulation and a larger mutational target. It remains unknown whether natural selection might be important in shaping genetic variation in these features of *C. elegans*, and whether these properties are general across species.

Here we quantified transcriptome expression for *C. briggsae* nematodes from populations with distinct genetic backgrounds adapted to temperature differences associated with their origins in Tropical versus Temperate latitudes (Prasad *et al.* 2011; Stegeman *et al.* 2013; Poullet *et al.* 2015). Global collections and population genomic analyses of *C. briggsae* wild isolates from Tropical and Temperate regions show that they form distinct phylogeographic groups (Cutter *et al.* 2006; Jovelin & Cutter 2011; Felix *et al.* 2013; Thomas *et al.* 2015). Given this ecological context, along with resources like recombinant inbred line (RIL) libraries and chromosome-scale genome assembly (Ross *et al.* 2011; Stegeman *et al.* 2019), *C. briggsae* represents a valuable system to understand the links between temperature and genetic background in differential gene expression. The exemplar Tropical and Temperate genotypes used as RIL parents, the focus of the present study, exhibit diverse temperature-dependent phenotypic differences consistent with adaptive differentiation of the phylogeographic groups overall, including for fecundity (~2-fold difference at 14°C, ~4-fold difference at 30°C), motility, and gamete development traits (Prasad *et al.* 2011; Stegeman *et al.* 2013; Poullet *et al.* 2015; Stegeman *et al.* 2019). By rearing these animals at hot and cold sublethal temperatures near their fertile limits, as well as under benign thermal conditions, we characterize genotypic and environmentally-induced differential gene expression across the genome. We then describe transcriptome complexity in terms of co-expression modularity to reflect transcriptome plasticity and robustness to environmental and genetic context, demonstrating distinctive genomic spatial distributions, functional attributes, and rates of molecular evolution at the sequence level.

Materials and Methods

Experimental design and sequencing

To quantify the genome-wide effects of rearing temperature and genetic background on gene expression, we isolated and sequenced mRNA transcriptomes from *C. briggsae* young adult hermaphrodites of two isogenic strains (AF16 = “Tropical” strain, HK104 = “Temperate” strain) that were reared under “chronic” exposure to 14°C (~150h), 20°C (~65h), and 30°C (~48h) throughout their development from egg to adult. Previous generations of both genotypes had been raised at benign 20°C prior to establishment of eggs for rearing at the treatment temperatures following stage synchronization via standard *Caenorhabditis* sodium hypochlorite (“bleaching”) protocol (Stiernagle 1999), avoiding potential transgenerational effects of stressful temperature on gene expression. After reaching adulthood (checked for young gravid adult

hermaphrodites), total RNA was isolated with Trizol extraction and isopropanol precipitation (Tu *et al.* 2015) from mass isogenic cultures of each strain at each rearing temperature with three biological replicates (2 genotypes x 3 rearing temperatures x 3 replications = 18 samples). The mRNA was then separated from small RNA fractions of less than approximately 200 nucleotides using the mirVana kit from Ambion as per the manufacturer's instructions, and prepared for single-end 100bp sequencing of TruSeq libraries via Illumina HiSeq 2000 (Genome Quebec, Canada) with each of the 18 barcoded samples sequenced across 2 lanes to control for lane effects (Fang & Cui 2011).

We obtained an average of 51.4 million reads per sample (range: 34.5 - 73.4 million) for 925.3 million total reads. Sequences are available in NCBI in project accession PRJNA509247. Over 96% of reads were retained after cleaning and trimming of raw FASTQ files with Trimmomatic 0.36 (894.4 million reads retained), using a seed-mismatch rate of 2, a simple clip threshold of 10, discarding reads <60bp long, and trimming bases from 5' and 3' ends if they had phred33 scores lower than 3 (Bolger *et al.* 2014).

Read mapping and expression counts

For each sample, we mapped reads to the *C. briggsae* genome (WS253) using STAR (Dobin *et al.* 2013), setting the maximum intron size to 5000 bp which includes 99.6% of all intron annotations in the *C. briggsae* reference genome. We applied a liberal mismatch rate of 10 to accommodate potential mapping efficacy differences between the AF16 and HK104 strains due to their genetic differences; the reference genome is based on the AF16 strain, so this liberal parameter choice minimizes the potential for mapping to bias towards the Tropical genotype that could inflate inference of differential expression due to genotype. Over 90% of the 894.4 million total reads mapped to unique locations, in all samples (except one replicate of HK104 at 30°C with 73.86% of 48.4 million reads mapping uniquely), with an average of 45.9 million reads mapping per sample to unique locations (Supplementary Table S1).

We then counted the number of reads that mapped to each exon annotated in the WS253 reference genome with htseq-count (Anders *et al.* 2013) and summed over all exons in a gene to give a raw measure of expression for each gene in each sample. For our analysis, we neglected alternative splicing isoforms, treating them as contributing to expression levels for the same gene, and set the "mode" parameter in htseq-count to "intersection-nonempty" to resolve ambiguity for overlapping genes (Anders *et al.* 2013). Among mapped reads, 82-85% were

assigned successfully to a particular gene among the 23,267 genes annotated in the WS253 *C. briggsae* genome in all samples (again excepting one replicate of HK104 at 30°C, with 24.0 million = 58% of reads assigned to genes). Among the reads that were not assigned to genes, most (9% on average) could not be associated with any exon or were counted in multiple locations (8% on average) and less than 0.1% were ambiguous.

Differential expression analysis

We first visualized gene expression counts in a Multi-Dimensional Scaling (MDS) plot (Nikolayeva & Robinson 2014), which showed strong clustering of most biological replicates within a treatment and differentiation among treatments (Supplementary Figure S1). We then retained only the subset of genes with at least 1 cpm (gene read count per million; using the "cpm" function in edgeR (Robinson *et al.* 2010)) in 3 or more libraries (i.e. in one biological replicate) to exclude 7068 genes with extremely low expression that could bias downstream analysis. It is possible that the genes filtered out at this step might exhibit higher expression at different developmental stages, males, or alternative environmental conditions than those assessed here. To test for statistical evidence of differential expression, we next transformed the expression counts using limma and voom, which performs well in controlling Type I error and in detecting true positives (Smyth 2005; Law *et al.* 2014; Ritchie *et al.* 2015). Preliminary analysis (not shown) found limma to be more conservative than edgeR for our dataset, so we elected to use limma for downstream analysis. Upon applying the voom transformation from the limma package to the remaining set of 16,199 genes, a Q-Q plot showed that the data closely approximated a normal distribution (Supplementary Figure S1).

We then tested these 16,199 genes for differential expression using limma by fitting a linear model to the expression profile for each gene as: expression ~ strain + temperature + strain*temperature interaction. We first tested for significance of the interaction term, and then tested for significance of the main effect terms only if the interaction was non-significant. The model intercept was set as expression for the Tropical strain at 20°C and P-values were adjusted for multiple testing using the Benjamini-Hochberg correction with significance inferred for a false discovery rate (FDR) of 0.05 (Benjamini & Hochberg 1995). To distinguish which genes responded to hot versus cold rearing conditions for genes with a significant effect of temperature (either main effect or interaction effect), we performed post-hoc tests on the individual temperature coefficients (FDR = 0.05). We then classified genes into five mutually exclusive categories based on whether they showed significant differential expression due to

191 genotype (strain) only ("G only" genes), temperature only ("T only" genes), both genotype and
192 temperature as independent main effects (i.e. additive effects; "G&T" genes), a non-additive
193 interaction between genotype and temperature ("GxT" genes), or no differential expression ("no
194 DE" genes).

195 ***Co-expression clustering of gene expression profiles***

196 To capture distinct stereotypical profiles of gene expression differences in response to our
197 temperature and genotype treatments, we performed a co-expression clustering analysis using
198 the Weighted Gene Correlation Network Analysis (WGCNA) package (Langfelder & Horvath
199 2008). Because WGCNA works best with normally distributed expression values, we again used
200 the voom-transformed expression values for the 16,199 filtered genes. A preliminary
201 hierarchical clustering analysis of the samples rejected batch effects as a source of
202 heterogeneity among samples, instead identifying both genotype and temperature as likely and
203 biologically interesting sources of variation in the data (Supplementary Figure S1). We
204 determined the best soft-thresholding power parameter for our data to be 30 (R^2 correlation with
205 a scale-free network topology = 0.75) based on fits across a range of values from 1 to 42
206 (Supplementary Figure S2), which also yielded an acceptable level of mean connectivity (k =
207 115), which is central to the assumptions of the WGCNA model (Zhang & Horvath 2005).

208 Running WGCNA yielded 124 initial clusters of genes with similar patterns of expression, which
209 we consolidated further by merging similar modules, defined as those with a correlation of 0.75
210 or higher with each other (Supplementary Figure S2). This procedure produced 22 co-
211 expression modules plus one pseudo-module (M0) containing the 37 genes that could not be
212 grouped based on expression pattern. The characteristic expression profile of genes in a
213 module is represented by WGCNA as the first principal component in expression space, termed
214 the "module eigengene" (Langfelder & Horvath 2007), which we plotted for each genotype
215 separately as the module eigengene expression values averaged across the three biological
216 replicates as a function of rearing temperature.

217 We performed statistical overrepresentation tests of Gene Ontology (GO) terms associated with
218 gene lists of each co-expression module using PANTHER (Mi *et al.* 2010), using all four
219 PANTHER lists available for *C. briggsae*: Pathways, GO-slim Molecular Function, GO-slim
220 Biological Process, and GO-slim Cellular Components. P-values were adjusted for multiple
221 testing with the Bonferroni correction.

Genomic enrichment analysis

C. briggsae chromosomes are defined by distinct recombination domains (high recombination arms, low recombination centres, and small tip regions with little detectable recombination), which also correlate with the density of coding genes, repetitive elements and single nucleotide polymorphism (SNPs) (Hillier *et al.* 2007; Ross *et al.* 2011; Thomas *et al.* 2015). We therefore tested whether gene profiles of differential expression or module affiliation were enriched in particular chromosomal regions using Bonferroni-adjusted G-tests, defining arm-center boundaries as in Ross et al. (2011). Analyses of upstream intergenic lengths were log-transformed prior to analysis with ANOVA, excluding genes with overlapping positions in the genome annotation. We used the transcription factor gene designations from (Haerty *et al.* 2008). We also cross-referenced differential expression categories and co-expression module membership with Wormbase-defined *C. elegans* orthologs found have sex-biased differential expression by Ortiz et al. (2014), which we used to test for enrichment with G-tests.

SNP and molecular evolution analysis

Genotype-dependent differences in expression could result from allelic differences in the local vicinity of genes (*cis*-acting effects; e.g. variants in promoter or nearby enhancer elements) or in distant regulators (*trans*-acting effects; e.g. variants in the regulation or functional sequence of transcription factors or miRNAs) (Rockman & Kruglyak 2006). The allelic differences contributing to local *cis*-acting regulation are likely to occur in the upstream promoter regions for those genes showing genotype-dependent expression (Grishkevich *et al.* 2012), though there are additional important roles of downstream and intronic regulatory elements in gene expression (Merritt *et al.* 2008). Therefore, we quantified the incidence of single-nucleotide polymorphisms (SNP) between the AF16 and HK104 genomic backgrounds in 500bp upstream and downstream flanking regions of coding sequences, as promoter regions tend to be in close proximity to coding sequences in *Caenorhabditis* (Saito *et al.* 2013).

We called single nucleotide variants between AF16 and HK104 based on Illumina paired-end sequencing of HK104 to ~33x coverage using identical methods of Thomas et al. (2015), yielding 761,531 SNPs and 173,341 indels. Sequences are available in NCBI in project accession PRJNA509247. We calculated the per-bp density of SNPs (π) in the pairwise comparison of AF16 and HK104 in a 500bp window upstream (and downstream) of coding sequences, excluding genes internal to operons (and using just the 5'-most or 3'-most operon gene for upstream or downstream sequence, respectively). 1070 operons comprising 2573

genes were identified based on orthology and synteny with annotated *C. elegans* operons, as in Tu et al. (2015). We also calculated the per-bp incidence of SNPs for different genomic features on a per-gene basis, including non-synonymous sites, synonymous sites, and introns, in addition to the 500bp flanking regions, after masking non-covered and low-quality sites. The effective number of codons (ENC) metric of biased codon usage was calculated for each gene in the *C. briggsae* reference genome WS253 with codonw (J. Peden, http://codonw.sourceforge.net). We used 6911 coding sequence divergence values (dN/dS') for 1-1 orthologs between *C. briggsae* and *C. nigoni* from Thomas et al. (2015).

Results

Widespread genotype- and temperature-dependent differential gene expression

We tested for differential expression across the *C. briggsae* transcriptome in response to three rearing temperature conditions and two genotypes, based on gene expression quantification from 45.7 million uniquely-mapped RNA sequence reads for each triplicate sample on average (824 million total mapped reads; Supplementary Table S1). Over half (54%, n=8795) of *C. briggsae* genes analyzed showed significant differential expression due to genotype, temperature, or both (16,199 genes tested after quality filtering for the genome's 21,827 annotated coding genes). The majority of these genes had a significant genotype-specific response to temperature (n=4919 "GxT genes"; 56% of 8795 differentially expressed genes; 30.4% of all genes analyzed; Supplementary File S1) (Figure 1A). In contrast to this "complex" GxT pattern, the remaining 3876 differentially expressed genes exhibited a "simple" dependence on genotype, temperature, or additive effects of both (8.8% "G genes", n=770; 23% "T genes", n=1987; 13% "G&T genes", n=1119 genes). Although 64% more genes overall exhibited a simple plastic response to temperature than a deterministic response to genotype (1987 + 1119 = 3106 vs 770 + 1119 = 1889), the abundance of genes with a complex GxT interaction of both factors highlights the important roles of both environmental plasticity and genetic determinism in transcriptome profiles (Figure 1A).

Distinct genetic responses to chronic heat versus cold stress

Genes with expression influenced by chronic cold stress (14°C) responded differently than genes affected by chronic heat stress (30°C) in terms of the number of genes involved, whether genes increased or decreased expression, and the magnitude of expression change. In

particular, cold stress affected expression of 74% of those genes with simple effects of temperature relative to benign conditions at 20°C (2308 of the 3106 T and G&T genes), whereas it was heat stress that altered expression of the plurality of GxT genes (2393 of 4919 genes, 49%) (Figure 1B). Among all these genes that responded to temperature in some way, more genes showed reduced expression at cool temperatures and elevated expression at warm temperatures, compared to benign conditions (Figure 1C; 1.05-fold reduction for T plus G&T and 1.3-fold reduction for GxT at 14°C, 6.8-fold elevated for T plus G&T and 1.2-fold elevated for GxT at 30°C). In terms of the magnitude of differential expression, chronic cold and heat stress were similar for genes with simple expression dynamics (T and G&T genes; Figure 1D; 6.1 to 6.5-fold increase for heat and cold; 3.0 to 3.7-fold decrease for heat and cold). The magnitudes of elevated expression change for GxT genes, however, were much larger under chronic heat stress than under chronic cold stress (hot 8.57-fold vs. cold 3.73-fold increase). Reciprocally, GxT genes that decreased expression under chronic cold stress had a larger magnitude change than under chronic heat stress (hot 6.50-fold vs. cold 9.19-fold decrease). These observations support the idea that distinct genetic networks mediate response to cold versus heat stress, rather than control by a single shared temperature stress response.

Co-expression modules define gene sets with distinct sensitivities to temperature and genotype

We defined 22 co-expression clusters in the *C. briggsae* transcriptome with WGCNA to capture modules showing distinctive patterns of differential gene expression in response to temperature and genotype differences (Figure 2, Figure 3, Supplementary File S1). The stereotypical expression profile for genes in each co-expression module is represented by its “module eigengene,” defined by the first principal component in expression space (Figure 3). These eigengene profiles illustrate how a given module reflects a dominant trend of genotype-dependence (e.g. M10), temperature-dependence (e.g. M12), additive effects of genotype and temperature (G&T, e.g. M4), or genotype-specific sensitivity to temperature (GxT, e.g. M22) (Figure 3). An average of 46% of genes in a module showed individually-significant differential expression, ranging from a low of just 6% (M13) to a high of 84% (M15) (Figure 3, Supplementary Figure S3). Genes with temperature- and genotype-specific differential expression are concentrated within distinct subsets of modules (Figure 3). Moreover, modules differ in sequence characteristics and in their enrichment with sex-related differential gene expression, as described below.

Rapid molecular evolution in modules sensitive to genotype or temperature alone

Genotype-dependent expression profiles predominate in just two modules (M7 and M10), which together include 44% (n=342) of all 770 genes with significant 'genotype-only' differential expression. Their eigengene expression profiles show limited dynamics across temperatures, with expression for the Temperate HK104 genotype consistently higher than Tropical AF16 in M7 and consistently lower in M10 (Figure 3). M10 is enriched for gene ontology (GO) terms related to extracellular constituents (Supplementary File S2). GO term enrichment in M7 indicates disproportionate representation of genes with nervous system function, including 11 GABA and 11 acetylcholine receptor activity genes, such as the ortholog of *C. elegans* nicotinic acetylcholine receptor *acr-9*. This nervous system enrichment of M7 is salient due to the HK104 and AF16 strains of *C. briggsae* differing in rearing-dependent thermal taxis and locomotion (Stegeman *et al.* 2013; Stegeman *et al.* 2019), a suite of behaviors under neural control.

Genes in module M10 have several other special features compared to other modules: rapidly-evolving protein coding sequences (high dN/dS'), high density of SNPs in replacement sites despite lowest SNP density in introns, the highest enrichment in arm regions of autosomes, enrichment on the X-chromosome, and exceptional rarity in operons (Figure 4, Figure 5A). We observed that genes in M10 also have the least consistent expression among replicates, with very few gene members having orthologs with "oogenic" expression according to Ortiz et al. (2014) (Figure 4A; Figure 5A). These features imply weaker canalization of expression of genes in M10, reflecting either weaker purifying selection or perhaps recent adaptive divergence in average expression levels that has not yet fine-tuned expression variability.

By contrast to the pronounced genotype-dependent differential expression in modules M7 and M10, two other modules each were comprised of >50% 'temperature-only genes' (M12, M15), although they accounted for just 12% (n=247) of the 1987 total T-only gene set (Figure 3). In both modules, eigengene expression is highest at high rearing temperatures across genotypes (Figure 3). In addition, modules M6 and M4 also contained a large fraction of temperature-only genes, and as large modules they also contain a large count of such genes (Figure 3). Modules M12 and M15 have genes with the highest average rates of evolution (dN/dS') and that occur only rarely in operons (Figure 5). They also have among the lowest average expression levels and codon usage bias (Figure 5A). Module M12 is highly enriched (3.7-fold) for orthologs with an oogenic gene classification in *C. elegans* (Ortiz *et al.* 2014), whereas M15 is depleted of such genes by having 2.6-fold fewer than expected (Figure 4A). Module M12 GO terms show

enrichment for genes associated with chromatin, like the ortholog of *C. elegans cec-7*, but with just 8 such genes of the 245 in M12, it is unclear how distinctive a property this is. More enigmatically, M15 shows no GO term enrichment, providing little clue as to whether these heat-sensitive and rapidly-evolving genes act in related functional pathways (Supplementary File S2).

Sperm gene function associated with both temperature- and genotype-dependence

Two co-expression modules were especially enriched in genes with additive effects of both genotype and temperature (M4, M5; G&T genes), accounting for over half (56%) of all such genes genome-wide (Figure 3). Their eigengene profiles show high expression at low temperatures, with the Tropical AF16 genotype having consistently higher expression than Temperate HK104 in M5 and vice versa for M4 (Figure 3). Interestingly, we found that module M5 is 3.3-fold enriched for orthologs of "spermatogenic" genes from Ortiz et al. (2014), a level unlike any other module (Figure 4A). Genes in M5 also are rare on the X-chromosome and nearly absent from operons, as expected for sperm-related genes (Reinke *et al.* 2000; Reinke & Cutter 2009; Albritton *et al.* 2014), and with fewer transcription factors (TFs) than most modules (Figure 4B; Figure 5A). Moreover, GO term enrichment in M5 indicates a prominent role of genes with phosphatase/kinase activity and glycogen metabolism (Supplementary File S2), including the orthologs of *C. elegans gsp-3/4* and *aagr-1*. Previous expression studies have reported male-biased and sperm-related genes to be enriched for genes with phosphatase/kinase GO terms (Reinke *et al.* 2004; Thomas *et al.* 2012), and some glycoproteins play crucial roles in sperm competitiveness in *C. briggsae* (Yin *et al.* 2018). Sperm fertility is known to show temperature sensitivity differently between the AF16 and HK104 genotypes of *C. briggsae* (Prasad *et al.* 2011; Poullet *et al.* 2015). Thus, the M5 expression pattern implies that universally higher expression for a suite of sperm-related genes, rather than a GxT profile, is associated with the greater sperm fertility at high temperatures observed in the AF16 genetic background.

Modules enriched for GxE and non-differential expression involved in core biological processes

Eight modules contained an especially large set of GxT genes (M1, M2, M3, M9, M14, M16, M18, M22), indicating a prominent influence of genotype-specific responses to temperature (Figure 3). These eight modules accounted for 71% (n=3477) of all GxT genes genome-wide.

379 The eigengene profiles for four of them show dramatic 'crossing reaction norms' such that the
380 Temperate HK104 and Tropical AF16 genotypes exhibit opposite expression responses to
381 rearing temperature (M9, M14, M18, M22; Figure 3). The known genotype-dependence in *C.*
382 *briggsae* for how sensitive oogenesis is to temperature, with strong reductions of mitotic and
383 meiotic cell counts in the gonad of HK104 animals (Poullet *et al.* 2015), suggests prime
384 candidates among the GxT genes in M9 that has 2.7-fold enrichment for orthologs with oogenic
385 roles (Ortiz *et al.* 2014) (Figure 4A). The other four modules show a much more exaggerated
386 eigengene expression response for just one of the genotypes specifically under high 30°C
387 conditions (M1, M2, M3, M16; Figure 3), rather than crossing reaction norms.

388 GO terms for core biological processes like mitochondria-related, ribosome-related, and/or
389 translation-related function are enriched in M14, M18 and M22, with M1, M2 and M9 enriched
390 for chromatin and transcription-related GO terms (Supplementary File S2). Genes in modules
391 M18 and M22 also are enriched for orthologs of "sex neutral" genes (Ortiz *et al.* 2014), are
392 enriched in operons, and include few TFs (Figure 4A; Figure 5A). Consistent with these
393 modules involving core biological functions, we also observed the distinctive features of M18
394 and M22 in having genes with the highest average expression and strongest codon usage bias,
395 while also having the strongest protein sequence conservation (lowest dN/dS') and the lowest
396 incidence of replacement-site SNPs (Figure 5; Supplementary Figure S5).

397 The seven remaining co-expression modules consisted primarily of genes that lacked
398 individually significant differential expression, though their module eigengene profiles
399 nevertheless suggest important effects of genetic background and temperature on the
400 stereotypical expression profile (M8, M11, M13, M17, M19, M20, M21). Several of these
401 modules showed GO term enrichment for various metabolic processes (M8, M11, M17, M21)
402 and transcriptional or translational functions (M8, M11, M13, M20). Among these modules, M8,
403 M19 and M20 are extremely enriched for orthologs of "oogenic" genes from Ortiz et al. (2014),
404 but include very few operonic genes (Figure 4A; Figure 5A). M20 also has the highest incidence
405 of TFs (29%) among all co-expression modules, has low average expression, and is enriched
406 for genes on autosomal arms and on the X-chromosome (Figure 4, Figure 5A, Supplementary
407 Figure S5). Genome-wide, TFs are more likely to show no differential expression than other
408 kinds of genes (no DE for 54.5% of TFs vs. 45.2% of other genes; G-test χ^2=29.2, P<0.0001).
409 Module M21 is distinctive in having the highest incidence of genes in operons (60.4%), which
410 are extremely rare on autosomal arms and the X-chromosome (Figure 4, Figure 5A). The 96
411 genes in M21 have extremely consistent expression across replicates, with most showing no

412 individually-significant differential expression due to either temperature or genotype (Figure 3;
413 Figure 5A). In *C. elegans*, these features are typical of genes that are expressed constitutively
414 across development (Cutter *et al.* 2019).

415 ***Genomic position and differential gene expression***

416 We hypothesized that genomic architectural and molecular evolutionary features might lead to
417 local enrichment of genes with genotype-dependent differential-expression. For example, SNP
418 variation is greater in the high recombination arm domains of autosomes in *C. briggsae*
419 (Thomas *et al.* 2015), and the X-chromosome exhibits a variety of distinctive features compared
420 to autosomes (Ross *et al.* 2011; Cutter 2018). Therefore, we tested for non-random distributions
421 of differentially-expressed genes along chromosomes and between chromosomes. We found
422 that autosome arm domains contained 22% more genes with genotype-dependent expression
423 than expected by chance, and also were slightly enriched for GxT genes (1.04-fold; Figure 1B).
424 Chromosome arms of *C. elegans* also have been reported to contain a disproportionate
425 representation of genes with genotype-dependent differential expression (Denver *et al.* 2005;
426 Rockman *et al.* 2010; Grishkevich *et al.* 2012). By contrast, it was center domains that
427 contained 15% more G&T genes than expected (Figure 1B). Temperature-only genes and
428 genes with no differential expression were randomly distributed between arm and center
429 domains (Figure 1B). Among the 22 co-expression modules, we observed 9 modules to have
430 significant enrichment in arms and 5 enriched in center domains of autosomes (Figure 4B).
431 Thus, gene expression profiles are not spatially independent and genome structural features
432 yield predictable patterns of differential expression within and between chromosomes.

433 We also found the X-chromosome to be enriched for genes with significant differential
434 expression due to genetic background (G-only genes) as well as for genes with no individually-
435 significant differential expression (X under-representation for G&T and GxT genes; Figure 1B).
436 X-linked biases also held true for co-expression modules (X enriched for 8 modules, autosomes
437 enriched for 8 modules; Figure 4B). Genes from module M21, in particular, are virtually absent
438 from the X-chromosome (Figure 4B), likely associated with the prevalence of operon genes in
439 this co-expression module that also tend to be exceptionally rare on the X-chromosome
440 (Blumenthal *et al.* 2002; Reinke & Cutter 2009). Chromosomes II and IV were distinctive in
441 having no module with significant enrichment or under-enrichment of genes (Supplementary
442 Figure S4). Other autosomes, however, were especially enriched (or under-enriched) for genes
443 in particular modules, for example, genes from M21 were 2.3-fold enriched on Chromosome III

and genes from M16 were 2.0-fold enriched on Chromosome V (Supplementary Figure S4). Given the extreme enrichment on Chromosome V for M16 and its genotype-specific expression response at 30°C (Figure 3), it is notable that a quantitative trait locus (QTL) mapping study in *C. briggsae* found QTL on Chromosome V to control differences in heat-sensitive movement behaviors (Stegeman *et al.* 2019).

In *C. elegans*, loci with genotype-dependent expression tend to have longer upstream intergenic regions, interpreted as being consistent with more complex *cis*-regulation of these genes (Grishkevich *et al.* 2012). We observed a similar pattern in *C. briggsae*, with median upstream length of 1367bp for G-only genes versus 1074bp for T-only genes (ANOVA $F_{4,15414}$=5.84, P<0.0001, Tukey post-hoc tests on log-transformed upstream intergenic length show G-only > T-only). After partitioning the genomic locations of differentially-expressed genes to account for their non-random distributions in the genome, however, we found that only those G-only genes in autosomal centers have significantly longer upstream intergenic regions compared to T-only genes (arms ANOVA $F_{4,5653}$=0.10, P=0.98; centers $F_{4,6410}$=5.50, P=0.0002, Tukey post-hoc tests on log-transformed upstream intergenic length show G-only > T-only). However, genes in autosomal centers with no differential expression also had longer upstream intergenic lengths than T-only genes and were not significantly different in length to GxT genes or G&T genes. We also find significant variation among co-expression modules in upstream intergenic length (arm ANOVA $F_{22,5635}$=13.59, P<0.0001; center ANOVA $F_{22,6392}$=15.48, P<0.0001), but observe no clear trend between length and the relative composition of genotype- or temperature-dependent genes. Thus, our analysis of *C. briggsae* upstream length distributions does not strongly support the notion that loci with genotype-dependent differential expression have more complex *cis*-regulatory controls.

Genome structure drives SNP associations with differential expression

We quantified the incidence of single-nucleotide polymorphisms (SNP) for the 761,531 SNPs between the AF16 and HK104 genomic backgrounds in 500bp upstream and downstream flanking regions of coding sequences, as promoter regions tend to be in close proximity to coding sequences in *Caenorhabditis* (Saito *et al.* 2013). We found zero upstream SNPs for 26.2% of the 16,167 genes that had expression and genomic coverage in both AF16 and HK104 (23.0% G-only, 30.9% G+T, 25.6% GxT). Such genes should have no role for *cis*-acting SNPs, suggesting this value as a lower-bound estimate for the incidence of entirely *trans*-acting regulatory differences that may alter genotype-dependent expression. Moreover, of

476 differentially-expressed genes affected by genotype, 32.6% have zero downstream SNPs,
477 20.3% have zero intronic SNPs, and 18.6% have zero SNPs in the coding sequence, also
478 consistent with a major role of *trans*-regulatory control being responsible for the genotype-
479 dependence of differential expression. Consistent with this idea, *C. elegans* shows a
480 predominant role of *trans*-regulatory control in genotype-dependent differential expression to
481 acute heat stress (Snoek *et al.* 2017).

482 We further predicted that an important role of *cis*-acting SNPs would be most evident by their
483 enrichment in association with G-only genes (as well as G&T genes and GxT genes), whereas
484 SNPs would be underrepresented in genes with no differential expression or T-only profiles.
485 Genome-wide, we did observe significant differences among differential expression categories
486 in the incidence of SNPs in upstream (ANOVA $F_{4,16141}$=7.63, P<0.0001), downstream
487 ($F_{4,16141}$=4.75, P=0.0008), and intronic portions of coding genes ($F_{4,16141}$=7.82, P<0.0001).
488 Overall, G-only genes have significantly higher SNP densities than other expression classes at
489 replacement sites, synonymous sites, introns and flanking sequences and genes with a GxT
490 pattern of differential expression had a greater density of SNPs than T-only genes only in
491 introns. These results mirror the report by Grishkevich et al. (2012) for *C. elegans* that SNPs are
492 enriched in promoters of genes with genotype-dependent differential expression.

493 Our findings therefore superficially support the idea of a key role for *cis*-acting SNPs controlling
494 genotype-dependent differential expression. However, we observed that this trend is driven
495 primarily by the enrichment of G-only genes in chromosome arms (Figure 1B), where SNPs are
496 disproportionately abundant for both functionally-constrained and unconstrained sites (Thomas
497 *et al.* 2015). When we account for genomic region, SNP density remains elevated for G-only
498 genes among genes in autosomal centers but not in arms (ANOVA $F_{4,6729}$=3.60, P=0.0062, G-
499 only > other gene classes with Tukey HSD post-hoc test; Figure 5C). Thus, genome structure is
500 an important determinant of inferences about *cis*-acting regulators of genotype-dependent
501 differential expression. We hypothesize that the SNPs in upstream regions of genes in the “SNP
502 deserts” of chromosome centers are more likely to represent causal regulatory variants that
503 modulate gene expression.

504 ***Molecular evolution is decoupled between coding and regulatory sequence***
505 ***regions***

Replacement-site SNPs are rarest in the coding sequences of non-differentially expressed genes (in both chromosome arms and centers), consistent with these genes having strongest selective constraint that most effectively eliminates new mutations (Figure 5C). Weaker selective constraint among genes with genotype-dependent differential expression that allows mutations to accumulate could result in their excess of coding SNPs. However, replacement-site divergence between species, which reflects a longer timescale of evolution, is no different between G-only genes and non-differentially expressed genes (median G-only dN/dS' = 0.0580, no DE dN/dS' = 0.0511; no significant difference from Tukey's post-hoc test on log-transformed values). These contrasting patterns for the scale of divergence between phylogeographic groups and between species suggest that relaxed selection on G-only genes might be evolutionarily recent or that adaptive divergence between Temperate and Tropical phylogeographic groups of *C. briggsae* contribute disproportionately to loci with genotype-dependent differential expression.

Consistent with the idea that genes and modules with many SNPs are subject to weaker selective constraints, co-expression modules with high average coding SNP density have low average expression and weak codon usage bias (Figure 5A; Figure 5B; Supplementary Figure S5). Associations were weaker for non-coding flanking regions (Supplementary Figure S5). Modules M10, M16, M12 and M15 were most enriched for coding SNPs (Figure 5B), also exhibiting among the lowest average expression and codon bias. Modules with high coding SNP density also show high long-term molecular evolutionary divergence between species (dN/dS'; Figure 5B), further implicating their constituent genes being subject to weaker selective constraints, or potentially, a greater incidence of adaptive divergence.

Finally, we tested for molecular evolutionary correspondence between coding and non-coding sequence. First, we found that SNPs and interspecies divergence correlate positively across genes for replacement sites, consistent with concordant pressures of purifying selection at both short and long evolutionary timescales on coding sequences (log-transformed $\pi_{nonsynonymous}$ and dN/dS', $F_{1,5741} = 646.7$, $P<0.0001$). However, interspecies divergence in coding sequence did not correlate with SNP density in non-coding sequences (Supplementary Figure S6). When we analyzed average values for co-expression modules instead of per-gene values, however, we observe positive correlations of non-coding SNP density with both coding SNPs and interspecies divergence (Supplementary Figure S6), suggesting that the distinct gene contents and genomic locations of genes among modules partly contributes to the coding-noncoding correspondence at the module level. Overall, these observations support the idea that selection

pressures are largely decoupled between coding sequences and non-coding flanking sequences that contain regulatory elements.

Muted differential expression role among heat shock proteins

We hypothesized that if heat shock proteins (hsp) modulate transcriptomic responses to chronic temperature stress then we would detect disproportionate differential expression for hsp genes. Of the 24 hsp genes in our expression dataset, only 8 showed significant differential expression, which is less than the genome overall (33% vs. 54%; G-test $\chi^2 = 4.267$, df = 1, P = 0.039), and similar to the genome in differential expression categories (Fisher Exact Test, P = 0.35). This suggests that hsp genes may play a lesser role in temperature stress experienced chronically across development, despite their profound importance to maintaining homeostasis in the face of acute heat stress (Lindquist & Craig 1988); even with acute heat shock, however, few genes show consistent upregulation in *C. elegans* (GuhaThakurta *et al.* 2002).

Discussion

The *C. briggsae* transcriptome shows widespread differential expression arising from distinct chronic temperatures over development and from distinct genotypes representative of phylogeographic groups from Tropical versus Temperate parts of the world, altering expression for over half of its genes. Genotype-specific responses to temperature represent the most common kind of differential expression (i.e. non-additive genotype-by-environment interactions), with less than a quarter of differentially-expressed genes being sensitive to temperature alone or genotype alone. Our temperature and genotype conditions cluster transcriptome responses into 22 co-expression modules, each comprised of genes with distinctive functional and evolutionary properties that reveal an important role for genome structure in transcriptomic patterns of differential expression.

The influence of genome structure in differential expression profiles

We found that genome structure plays an important role in shaping the landscape of differential expression of the *C. briggsae* transcriptome, and in the molecular evolutionary features of the corresponding genes. Transcriptome profiles cluster within and between chromosomes, making them susceptible to cryptic correlations with other non-random genomic features. For example, genes showing genotype-dependent expression were enriched on *C. briggsae* chromosomal arms, genomic regions that also are rich in SNPs and with high rates of recombination (Ross *et*

al. 2011; Thomas *et al.* 2015). This pattern is reminiscent of the excess of eQTL and loci with genotype-dependent expression on *C. elegans* chromosome arms (Denver *et al.* 2005; Rockman *et al.* 2010; Grishkevich *et al.* 2012). One possibility is that direct selection drives this pattern, as could occur from either adaptive divergence being more prevalent or from purifying selection being weaker for genes on high recombination arms. Alternatively, it might result as a byproduct of linked selection, known to be potent in the *C. briggsae* genome (Cutter & Choi 2010; Thomas *et al.* 2015), whereby elimination of polymorphisms from low recombination centers simply leads to few loci with the potential to show genotype-dependent differential expression.

The higher recombination rate of arm regions means that natural selection favoring a given allele at one locus will be subject to less interference from selection at other loci in the genome (Hill & Robertson 1966; Comeron *et al.* 2008; Cutter & Payseur 2013). Experiments implicate temperature-related adaptive divergence between *C. briggsae* genotypes from Tropical and Temperate latitudes (Prasad *et al.* 2011). Consequently, gene-specific adaptation to distinct ecological conditions should operate more efficiently for genes on arms, which might yield the enrichment of genotype-dependent expression on arms as well as the more rapid sequence evolution of genes on arms. However, it is difficult to exclude a role of linked selection, as the high self-fertilization in *C. briggsae* leaves a substantial imprint on genomic patterns of variation for both synonymous and non-synonymous polymorphisms (Cutter & Choi 2010; Thomas *et al.* 2015). Moreover, if our observations of genotype-dependent differential expression depend primarily on a small number distant *trans*-acting upstream regulators that influence many target loci (rather than local *cis*-acting allelic variants for many genes), then the bias toward chromosome arms of differentially-expressed genes might simply be a byproduct of non-random distributions of gene functions encoded across the genome.

Sequence conservation and regulatory controls

The three co-expression modules with the strongest coding sequence conservation also have the highest make-up of GxT genes, which exhibit crossing reaction norms and very high average expression as well as functional enrichment of core biological processes (M14, M18, M22). This finding of especially strong purifying selection implicates either adaptive plasticity in the expression control of these modules or unusually low robustness of expression levels to perturbation from both genotypic and environmental sources. At the other end of the spectrum, modules displaying the highest average rates of sequence evolution have the lowest overall

601 expression and the most pronounced dependence on genotype alone (M10) or temperature
602 alone (M12 and M15). These findings are consistent with expression level as a key determinant
603 of rates of coding sequence evolution, with faster molecular evolution of weakly expressed
604 genes. Similarly, *C. elegans* genes showing non-interaction differential expression tend to have
605 low expression levels (Grishkevich *et al.* 2012). These results can be explained by weaker
606 purifying selection on low-expression genes, though it remains possible that adaptive change
607 might also play a disproportionate role in their molecular evolution.

608 Among genes located in chromosome centers, those with genotype-dependent differential
609 expression are enriched for SNPs in upstream non-coding regions, consistent with local *cis*-
610 acting alleles affecting their expression. However, long-term coding sequence divergence
611 correlates poorly with non-coding SNP density across genes, implying that the strength of
612 selection on coding sequence variation may be decoupled from *cis*-regulatory genetic variation
613 (Castillo-Davis *et al.* 2004; Jordan *et al.* 2005; Lemos *et al.* 2005; Liao & Zhang 2006; Tirosh &
614 Barkai 2008) or that regulatory elements are too sparse within flanking DNA to leave a clear
615 selective signature with our approach. Nevertheless, the abundance of loci with zero upstream
616 SNPs suggests that distant *trans*-regulatory control is a profound source of genetic variation in
617 the differential expression patterns that we quantified, consistent with studies of short
618 evolutionary timescales in other systems (Stern & Orgogozo 2008; Wittkopp & Kalay 2012). For
619 example, eQTL analysis for both *C. elegans* and yeast implicates a stronger role for *trans*-
620 relative to *cis*-regulation of genotype-environment interactions (Li *et al.* 2006; Smith & Kruglyak
621 2008).

622 ***Gene function in differential expression profiles***

623 We found that module M5 was unique in having a large representation of sperm-related genes
624 among its orthologs to *C. elegans* genes with spermatogenesis-enriched function. It includes
625 overrepresentation of genes with phosophatase/kinase activity and is associated with glycogen
626 metabolism, which previous studies show to be especially important in sperm function (Reinke
627 *et al.* 2004; Thomas *et al.* 2012; Wu *et al.* 2012; Yin *et al.* 2018). Both genotype and
628 temperature were important determinants of expression profiles in M5 (Figure 3), implicating the
629 potential for both adaptive divergence and phenotypic plasticity to influence gene responses.
630 Sperm-dependent fertility appears to be especially sensitive to high temperature, with Tropical
631 and Temperate genotypes of *C. briggsae* differing in sensitivity (Prasad *et al.* 2011; Poullet *et al.*
632 2015), though module M5 shows additive contributions for genetic and temperature effects. As

633 expected for sperm genes (Reinke *et al.* 2000; Reinke & Cutter 2009; Albritton *et al.* 2014),
634 genes from M5 are especially rare on the X-chromosome and virtually absent from operons.

635 Opposite to the rarity of operon genes in sperm-enriched module M5, fully 60% of the genes in
636 module M21 occurred in operons and yet less than 17% of them had individually significant
637 differential expression. Overall, genes in operons were much less likely to show significant
638 differential expression than non-operon genes (43% of operon vs. 56% of non-operon genes;
639 Fisher exact text P<0.0001). *C. elegans* operon genes, most of which are conserved in *C.*
640 *briggsae* (Qian & Zhang 2008), are known to show high expression during growth, as for gonad
641 tissue (Reinke & Cutter 2009) and following growth-arrested states (Zaslaver *et al.* 2011), and
642 generally have non-dynamic expression profiles across ontogeny (Cutter *et al.* 2019). These
643 observations are consistent with operon genes being disproportionately robust to both
644 environmental and genetic perturbation.

645 ***Plasticity versus adaptive divergence in expression profiles***

646 In *C. elegans* and *C. remanei*, plasticity dominates the transcriptome response to temperature
647 stress, at least in terms of acute heat shock (Jovic *et al.* 2017; Sikkink *et al.* 2019). We also
648 found in *C. briggsae* that over 90% of differentially-expressed genes changed at least in part
649 due to temperature, but more commonly due to chronic cool rather than warm conditions. If
650 environment-dependent expression responses reflect adaptive plasticity, then our observations
651 suggest stronger canalization of stereotyped cool-rearing expression responses. While the large
652 number of such differentially-expressed genes does not pinpoint the key determinants of
653 temperature-dependent adaptive divergence, we can nevertheless largely rule out the nearly
654 9500 genes in the genome that show temperature-only effects or no differential expression. Our
655 analysis of *C. briggsae* finds a stronger signal of genotype-dependent differential gene
656 expression than the *C. remanei* study, perhaps reflecting the longer period of divergence
657 between AF16 and HK104 than between the experimental evolution lines for *C. remanei*, in
658 addition to technical differences between the studies (Sikkink *et al.* 2019). Warm conditions
659 causing pervasive genotype-specific responses in *C. briggsae* might reflect adaptive evolution
660 by the distinct genetic backgrounds from Tropical and Temperate regions (Prasad *et al.* 2011).
661 Phylogenetic comparative analysis of differential expression among genotypes and
662 environments could prove fruitful in deciphering whether shared gene networks across species
663 provide common substrate for adaptive divergence and adaptive plasticity in organismal
664 responses to chronic and acute temperature stress.

Conclusions

Genome-wide differential gene expression is sensitive to both extrinsic temperature conditions and to intrinsic genomic background in the nematode *C. briggsae*, with co-expression modules defining distinctive functional features, genomic distributions and molecular evolutionary patterns of their constituent genes. Most genotype-specific responses occur under heat stress, indicating that cold versus heat stress responses involve distinct genomic architectures. Co-expression modules associated with reproductive function, and which exhibit strong sensitivity to both temperature and genotype, provide candidates for adaptive divergence between Temperate and Tropical phylogeographic groups of *C. briggsae*. The fastest-evolving protein coding sequences correspond to a predominant influence of temperature alone or genotype alone, and have overall low levels of expression across conditions. However, chromosome-scale patterning of nucleotide differences is a key predictor of SNP content of genes, undermining gene-centric causes and *cis*-regulatory inferences for SNP differences across differential-expression classes of genes. These findings highlight the powerful way that genome structure can influence transcriptome profiles to make them susceptible to cryptic correlations with other non-random genomic features.

Acknowledgements

We are grateful to Rajarshi Ghosh and Leonid Kruglyak for helping to generate and share HK104 genomic sequence. We also thank two anonymous reviewers for constructive feedback on earlier drafts of the manuscript. This work was supported by funds from the Natural Sciences and Engineering Research Council of Canada and a Canada Research Chair to ADC and to JMC.

References

Albritton SE, Kranz A-L, Rao P*, et al.* (2014) Sex-biased gene expression and evolution of the X chromosome in nematodes. *Genetics* **197**, 865-883.

Anders S, McCarthy DJ, Chen Y*, et al.* (2013) Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protocols* **8**, 1765-1786.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300.

Blumenthal T, Evans D, Link CD*, et al.* (2002) A global analysis of *Caenorhabditis elegans* operons. *Nature* **417**, 851-854.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.

Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25-36.

Castillo-Davis CI, Hartl DL, Achaz G (2004) *cis*-Regulatory and protein evolution in orthologous and duplicate genes. *Genome Research* **14**, 1530-1536.

Causton HC, Ren B, Koh SS*, et al.* (2001) Remodeling of yeast genome expression in response to environmental changes. *Molecular Biology of the Cell* **12**, 323-337.

Comeron JM, Williford A, Kliman RM (2008) The Hill-Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* **100**, 19-31.

Cutter AD (2018) X exceptionalism in Caenorhabditis speciation. *Molecular Ecology* **0**.

Cutter AD, Choi JY (2010) Natural selection shapes nucleotide polymorphism across the genome of the nematode *Caenorhabditis briggsae*. *Genome Research* **20**, 1103-1111.

Cutter AD, Félix M-A, Barriere A, Charlesworth D (2006) Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates of *Caenorhabditis briggsae*. *Genetics* **173**, 2021-2031.

Cutter AD, Garrett RH, Mark S, Wang W, Sun L (2019) Molecular evolution across developmental time reveals rapid divergence in early embryogenesis. *bioRxiv*, 518621.

Cutter AD, Payseur BA (2013) Genomic signatures of selection at linked sites: unifying the disparity among species. *Nature Reviews Genetics* **14**, 262-274.

de Visser JAGM, Hermisson J, Wagner GP*, et al.* (2003) Evolution and detection of genetic robustness. *Evolution* **57**, 1959-1972.

Denver DR, Morris K, Streelman JT*, et al.* (2005) The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nature Genetics* **37**, 544-548.

Dobin A, Davis CA, Schlesinger F*, et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**.

Fang Z, Cui X (2011) Design and validation issues in RNA-seq experiments. *Briefings in Bioinformatics* **12**, 280-287.

Felix M-A, Jovelin R, Ferrari C*, et al.* (2013) Species richness, distribution and genetic diversity of *Caenorhabditis* nematodes in a remote tropical rainforest. *BMC Evolutionary Biology* **13**, 10.

Flatt T (2005) The evolutionary genetics of canalization. *The Quarterly Review of Biology* **80**, 287-316.

Gagneur J, Stegle O, Zhu C*, et al.* (2013) Genotype-Environment Interactions Reveal Causal Pathways That Mediate Genetic Effects on Phenotype. *PLoS Genetics* **9**, e1003803.

Grishkevich V, Ben-Elazar S, Hashimshony T*, et al.* (2012) A genomic bias for genotype–environment interactions in <em>C. elegans</em&gt. *Molecular Systems Biology* **8**.

Grishkevich V, Yanai I (2013) The genomic determinants of genotype × environment interactions in gene expression. *Trends in Genetics* **29**, 479-487.

GuhaThakurta D, Palomar L, Stormo GD*, et al.* (2002) Identification of a novel cis-regulatory element involved in the heat shock response in *Caenorhabditis elegans* using microarray gene expression and computational methods. *Genome Research* **12**, 701-712.

Haerty W, Artieri C, Khezri N, Singh RS, Gupta BP (2008) Comparative analysis of function and interaction of transcription factors in nematodes: extensive conservation of orthology coupled to rapid sequence evolution. *BMC Genomics* **9**, 399.

Hecker M, Lambeck S, Toepfer S, van Someren E, Guthke R (2009) Gene regulatory network inference: Data integration in dynamic models—A review. *BioSystems* **96**, 86-103.

Hill WG, Robertson A (1966) Effect of linkage on limits to artificial selection. *Genetical Research* **8**, 269-294.

748 Hillier LW, Miller RD, Baird SE*, et al.* (2007) Comparison of *C. elegans* and *C. briggsae* genome
749 sequences reveals extensive conservation of chromosome organization and synteny.
750 *PLoS Biology* **5**, e167.
751 Jordan IK, Mariño-Ramírez L, Koonin EV (2005) Evolutionary significance of gene expression
752 divergence. *Gene* **345**, 119-126.
753 Jovelin R, Cutter AD (2011) MicroRNA sequence variation potentially contributes to within-
754 species functional divergence in the nematode *Caenorhabditis briggsae*. *Genetics* **189**,
755 967-976.
756 Jovic K, Sterken MG, Grilli J*, et al.* (2017) Temporal dynamics of gene expression in heat-
757 stressed Caenorhabditis elegans. *PLoS ONE* **12**, e0189445.
758 Keightley PD, Lynch M (2003) Toward a realistic model of mutations affecting fitness. *Evolution*
759 **57**, 683-685.
760 Kim SK, Lund J, Kiraly M*, et al.* (2001) A gene expression map for *Caenorhabditis elegans*.
761 *Science* **293**, 2087-2092.
762 Knowles DA, Davis JR, Edgington H*, et al.* (2017) Allele-specific expression reveals interactions
763 between genetic variation and environment. *Nature Methods* **14**, 699.
764 Langfelder P, Horvath S (2007) Eigengene networks for studying the relationships between co-
765 expression modules. *BMC Systems Biology* **1**, 54.
766 Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network
767 analysis. *BMC Bioinformatics* **9**, 559.
768 Law CW, Chen Y, Shi W, Smyth GK (2014) voom: precision weights unlock linear model
769 analysis tools for RNA-seq read counts. *Genome Biology* **15**, R29.
770 Lemos B, Bettencourt BR, Meiklejohn CD, Hartl DL (2005) Evolution of proteins and gene
771 expression levels are coupled in *Drosophila* and are independently associated with
772 mRNA abundance, protein length, and number of protein-protein interactions. *Molecular*
773 *Biology and Evolution* **22**, 1345-1354.
774 Li Y, Alvarez OA, Gutteling EW*, et al.* (2006) Mapping determinants of gene expression
775 plasticity by genetical genomics in *C. elegans*. *PLoS Genetics* **2**, e222.
776 Liao B-Y, Zhang J (2006) Evolutionary conservation of expression profiles between human and
777 mouse orthologous genes. *Molecular Biology and Evolution* **23**, 530-540.
778 Lindquist S, Craig EA (1988) The Heat-Shock Proteins. *Annual Review of Genetics* **22**, 631-
779 677.
780 Merritt C, Rasoloson D, Ko D, Seydoux G (2008) 3' UTRs are the primary regulators of gene
781 expression in the *C. elegans* germline. *Current Biology* **18**, 1476-1482.
782 Mi H, Dong Q, Muruganujan A*, et al.* (2010) PANTHER version 7: improved phylogenetic trees,
783 orthologs and collaboration with the Gene Ontology Consortium. *Nucleic Acids Research*
784 **38**, D204-D210.
785 Nikolayeva O, Robinson MD (2014) edgeR for differential RNA-seq and ChIP-seq analysis: an
786 application to stem cell biology. In: *Stem Cell Transcriptional Networks: Methods and*
787 *Protocols* (ed. Kidder BL), pp. 45-79. Springer New York, New York, NY.
788 Ortiz MA, Noble D, Sorokin EP, Kimble J (2014) A new dataset of spermatogenic vs. oogenic
789 transcriptomes in the nematode *Caenorhabditis elegans*. *G3: Genes|Genomes|Genetics*
790 **4**, 1765-1772.
791 Poullet N, Vielle A, Gimond C, Ferrari C, Braendle C (2015) Evolutionarily divergent thermal
792 sensitivity of germline development and fertility in hermaphroditic Caenorhabditis
793 nematodes. *Evolution and Development* **17**, 380-397.
794 Prasad A, Croydon-Sugarman M, Murray RL, Cutter AD (2011) Temperature-dependent
795 fecundity associates with latitude in *Caenorhabditis briggsae*. *Evolution* **65**, 52-63.
796 Qian W, Zhang J (2008) Evolutionary dynamics of nematode operons: Easy come, slow go.
797 *Genome Research* **18**, 412-421.

798 Reinke V, Cutter AD (2009) Germline expression influences operon organization in the
799 *Caenorhabditis elegans* genome. *Genetics* **181**, 1219-1228.
800 Reinke V, Gil IS, Ward S, Kazmer K (2004) Genome-wide germline-enriched and sex-biased
801 expression profiles in *Caenorhabditis elegans*. *Development* **131**, 311-323.
802 Reinke V, Smith HE, Nance J*, et al.* (2000) A global profile of germline gene expression in *C.*
803 *elegans*. *Molecular Cell* **6**, 605-616.
804 Ritchie ME, Smyth GK, Phipson B*, et al.* (2015) limma powers differential expression analyses
805 for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47-e47.
806 Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential
807 expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.
808 Rockman MV, Kruglyak L (2006) Genetics of global gene expression. *Nature Reviews Genetics*
809 **7**, 862.
810 Rockman MV, Skrovanek SS, Kruglyak L (2010) Selection at linked sites shapes heritable
811 phenotypic variation in *C. elegans*. *Science* **330**, 372-376.
812 Ross JA, Koboldt DC, Staisch JE*, et al.* (2011) *Caenorhabditis briggsae* recombinant inbred line
813 genotypes reveal inter-strain incompatibility and the evolution of recombination. *PLoS*
814 *Genetics* **7**, e1002174.
815 Saito TL, Hashimoto S-i, Gu SG*, et al.* (2013) The transcription start site landscape of C.
816 elegans. *Genome Research* **23**, 1348-1361.
817 Sikkink KL, Reynolds RM, Ituarte CM, Cresko WA, Phillips PC (2019) Environmental and
818 evolutionary drivers of the modular gene regulatory network underlying phenotypic
819 plasticity for stress resistance in the nematode Caenorhabditis remanei. *G3:*
820 *Genes|Genomes|Genetics* **9**, 969-982.
821 Smith EN, Kruglyak L (2008) Gene–environment interaction in yeast gene expression. *PLoS*
822 *Biology* **6**, e83.
823 Smith S, Bernatchez L, Beheregaray LB (2013) RNA-seq analysis reveals extensive
824 transcriptional plasticity to temperature stress in a freshwater fish species. *BMC*
825 *Genomics* **14**, 375.
826 Smyth GK (2005) limma: Linear Models for Microarray Data. In: *Bioinformatics and*
827 *Computational Biology Solutions Using R and Bioconductor* (eds. Gentleman R, Carey
828 VJ, Huber W, Irizarry RA, Dudoit S), pp. 397-420. Springer New York, New York, NY.
829 Snoek BL, Sterken MG, Bevers RPJ*, et al.* (2017) Contribution of trans regulatory eQTL to
830 cryptic genetic variation in C. elegans. *BMC Genomics* **18**, 500.
831 Snoek BL, Volkers RJM, Nijveen H*, et al.* (2019) A multi-parent recombinant inbred line
832 population of *C. elegans* allows identification of novel QTLs for complex life history traits.
833 *BMC Biology* **17**, 24.
834 Stegeman GW, Baird SE, Ryu WS, Cutter AD (2019) Genetically distinct behavioral modules
835 underlie natural variation in thermal performance curves. *bioRxiv*, 523654.
836 Stegeman GW, Bueno de Mesquita M, Ryu WS, Cutter AD (2013) Temperature-dependent
837 behaviours are genetically variable in the nematode *Caenorhabditis briggsae*. *Journal of*
838 *Experimental Biology* **216**, 850-858.
839 Stern DL, Orgogozo V (2008) The loci of evolution: how predictable is genetic evolution?
840 *Evolution* **62**, 2155-2177.
841 Stiernagle TL (1999) Maintenance of *C. elegans*. In: *C. elegans: A Practical Approach* (ed.
842 Hope IA). Oxford University Press, New York.
843 Thomas CG, Li RH, Smith HE*, et al.* (2012) Simplification and desexualization of gene
844 expression in self-fertile nematodes. *Current Biology* **22**, 2167-2172.
845 Thomas CG, Wang W, Jovelin R*, et al.* (2015) Full-genome evolutionary histories of selfing,
846 splitting and selection in *Caenorhabditis*. *Genome Research* **25**, 667-678.
847 Tirosh I, Barkai N (2008) Evolution of gene sequence and gene expression are not correlated in
848 yeast. *Trends in Genetics* **24**, 109-113.

849 Tirosh I, Reikhav S, Levy AA, Barkai N (2009) A yeast hybrid provides insight into the evolution
850 of gene expression regulation. *Science* **324**, 659.
851 Tu S, Wu MZ, Wang J*, et al.* (2015) Comparative functional characterization of the CSR-1 22G-
852 RNA pathway in Caenorhabditis nematodes. *Nucleic Acids Research* **43**, 208-224.
853 Ungerer MC, Johnson LC, Herman MA (2007) Ecological genomics: understanding gene and
854 genome function in the natural environment. *Heredity* **100**, 178.
855 Wittkopp PJ, Haerum BK, Clark AG (2004) Evolutionary changes in cis and trans gene
856 regulation. *Nature* **430**, 85.
857 Wittkopp PJ, Haerum BK, Clark AG (2008) Regulatory changes underlying expression
858 differences within and between Drosophila species. *Nature Genetics* **40**, 346-350.
859 Wittkopp PJ, Kalay G (2012) Cis-regulatory elements: molecular mechanisms and evolutionary
860 processes underlying divergence. *Nature Reviews Genetics* **13**, 59-69.
861 Wu J-c, Go AC, Samson M*, et al.* (2012) Sperm development and motility are regulated by PP1
862 phosphatases in *Caenorhabditis elegans*. *Genetics* **190**, 143.
863 Yin D, Schwarz EM, Thomas CG*, et al.* (2018) Rapid genome shrinkage in a self-fertile
864 nematode reveals sperm competition proteins. *Science* **359**, 55-61.
865 Zaslaver A, Baugh LR, Sternberg Paul W (2011) Metazoan operons accelerate recovery from
866 growth-arrested states. *Cell* **145**, 981-992.
867 Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network
868 analysis. *Statistical Applications in Genetics and Molecular Biology* **4**, 1-45.

869

Data accessibility

Data used for analysis is provided in NCBI in project accession PRJNA509247 for transcriptome and genome sequences, with online supplement summary tables to be submitted to Dryad, strains are publicly available from the Caenorhabditis Genetics Center.

Author contributions

SM, JW, JMC and ADC designed research; SM, JW, ES, TL and WW performed research; JMC and ADC contributed reagents/analytic tools; SM, TL, WW and ADC analyzed data; SM, JMC and ADC wrote and edited the manuscript. All authors read and approved the final manuscript.

Figures

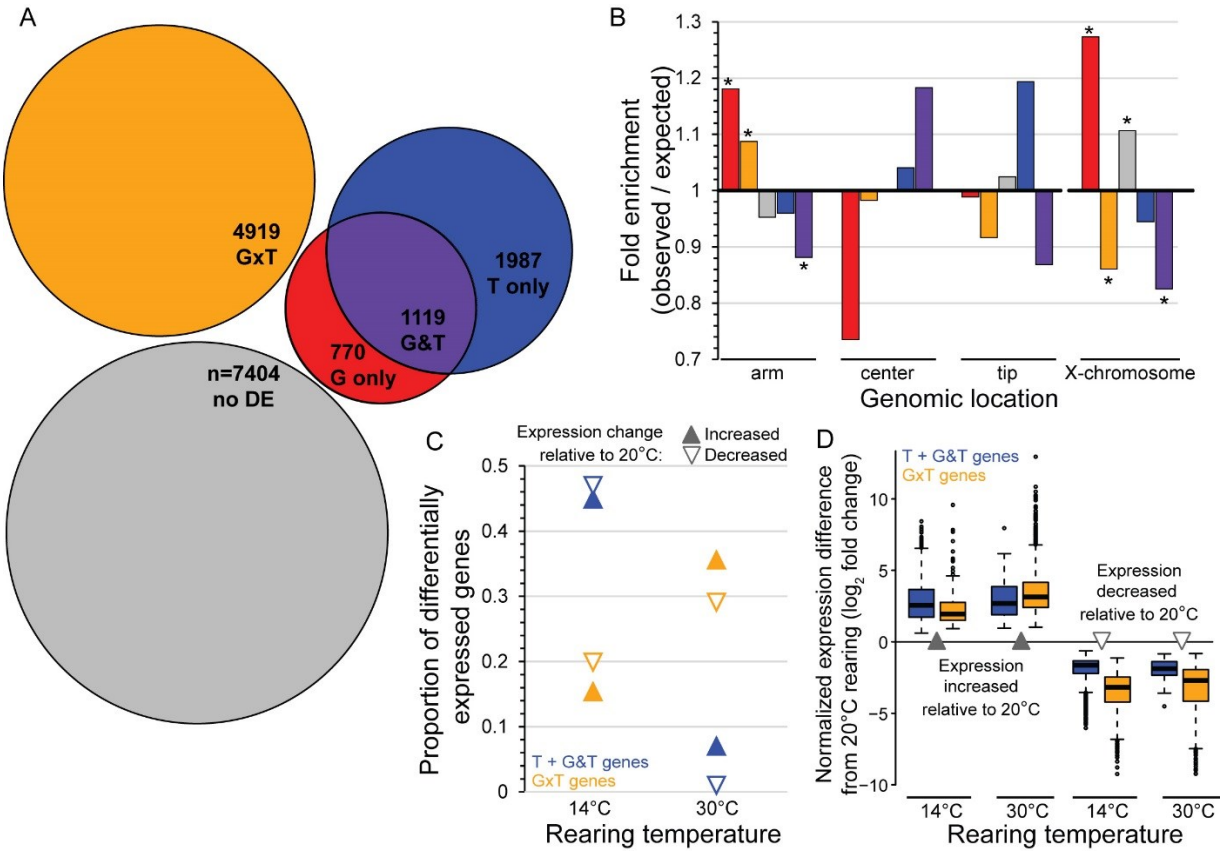


Figure 1. (**A**) Differential expression (DE) analysis identified 54% of 16,199 genes to have significant differential expression (8795 genes at 5% FDR in limma, colored area proportional to gene number). Most DE genes had significant interaction effects (55.9% “GxT”), whereas 12.6% of DE genes had independent additive effects of both genotype and rearing temperature (“G&T”). Other DE genes showed effects of genotype alone or rearing temperature treatments alone (8.8% “G only”; 22.6% “T only”). (**B**) G-only and GxT genes are significantly enriched on autosomal arms, whereas G&T genes are enriched in autosomal centers (colors as in A; * indicates G-test Bonferroni corrected P<0.05). Genes with G-only or no differential expression (“no DE”) are enriched on the X-chromosome, whereas genes with GxT and G&T patterns of differential expression are underrepresented on the X-chromosome (* indicates G-test Bonferroni corrected P<0.05). (**C**) Similar proportions of GxT genes increase vs decrease expression at a given stressful rearing temperature relative to benign 20°C (filled vs empty orange triangles within a temperature condition), but fewer GxT genes show expression differences for cool rearing than for hot rearing (orange triangles for 14°C vs 30°C). By contrast, genes with a non-interacting effect of temperature on expression (T only and G&T genes) show disproportionate response to cool rearing (blue triangles for 14°C vs all other triangles). (**D**) The magnitude of expression change is similar for genes with a non-interacting effect of temperature (T only and G&T genes) under chronic cold stress and chronic heat stress (blue boxes for 14°C vs 30°C; median with interquartile range, whiskers show 1.5x interquartile range). For GxT genes, however, the magnitude of expression increase is greater under heat stress than cold stress (orange boxes for 14°C vs 30°C).

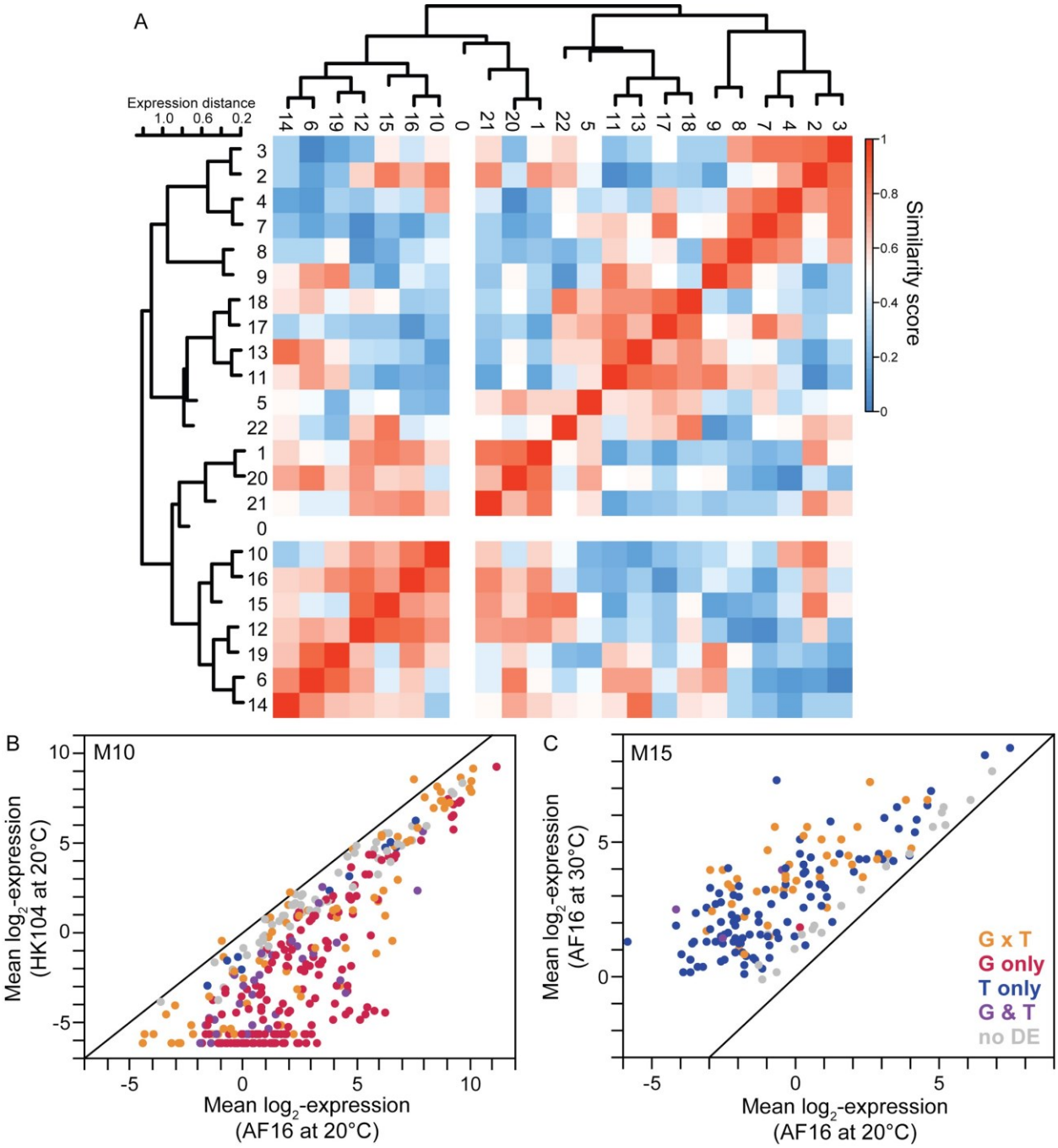


Figure 2. (**A**) WGCNA analysis yielded 23 co-expression modules for 16,199 genes (including module M0 with genes that did not cluster), after merging modules with expression similarity distance <0.25 from an initial set of 124 co-expressed gene sets. The dendrogram and heatmap summarize module expression profile similarity. For example, (**B**) module M10 is comprised of 338 genes with disproportionate representation of genotypic differences in expression, reflected as higher expression by the AF16 (Tropical) genotype. (**C**) Module M15, by contrast, is enriched in genes with individually significant differential expression due to rearing temperature, as reflected in higher expression when reared at 30°C. See also Figure 3.

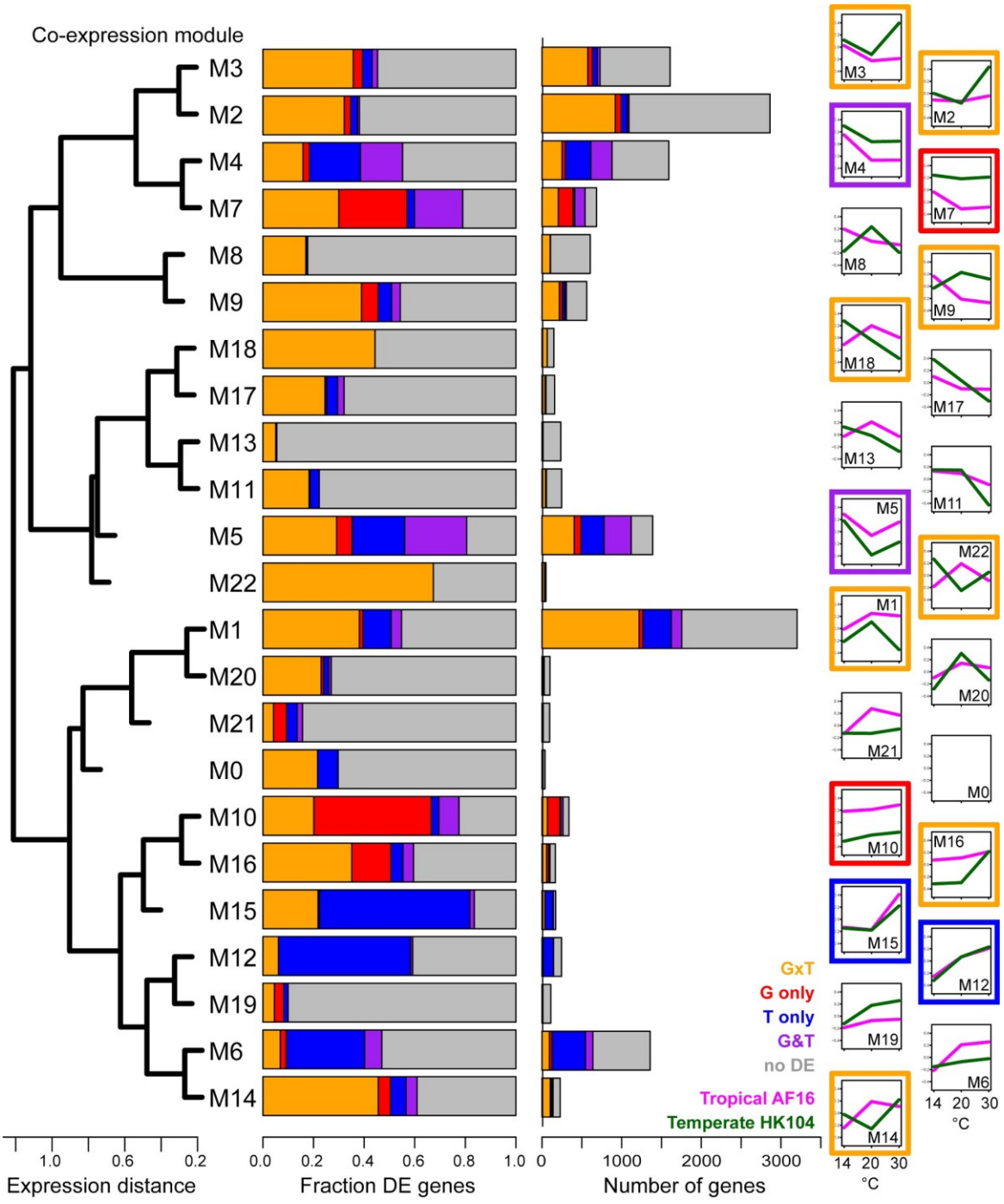


Figure 3. Co-expression clustering of 16,199 genes into 23 modules, reflected in eigengene plots of normalized log_2-transformed expression profiles across rearing temperature treatments (14°C, 20°C, 30°C) for each genotype (Tropical AF16, Temperate HK104). Modules range in size from 3203 genes (M1) to 49 genes (M22). Module compositions contained distinctive representation of genes with individually-significant patterns of differential expression (cf. Figure 1); modules with highest incidence of T-only genes indicated with blue (M12, M15), G-only genes in red (M7, M10), G&T genes (M4, M5), and GxT genes with crossing or non-crossing reaction norms (M9, M14, M18, M22; M1, M2, M3, M16) (Supplementary Figure S3).

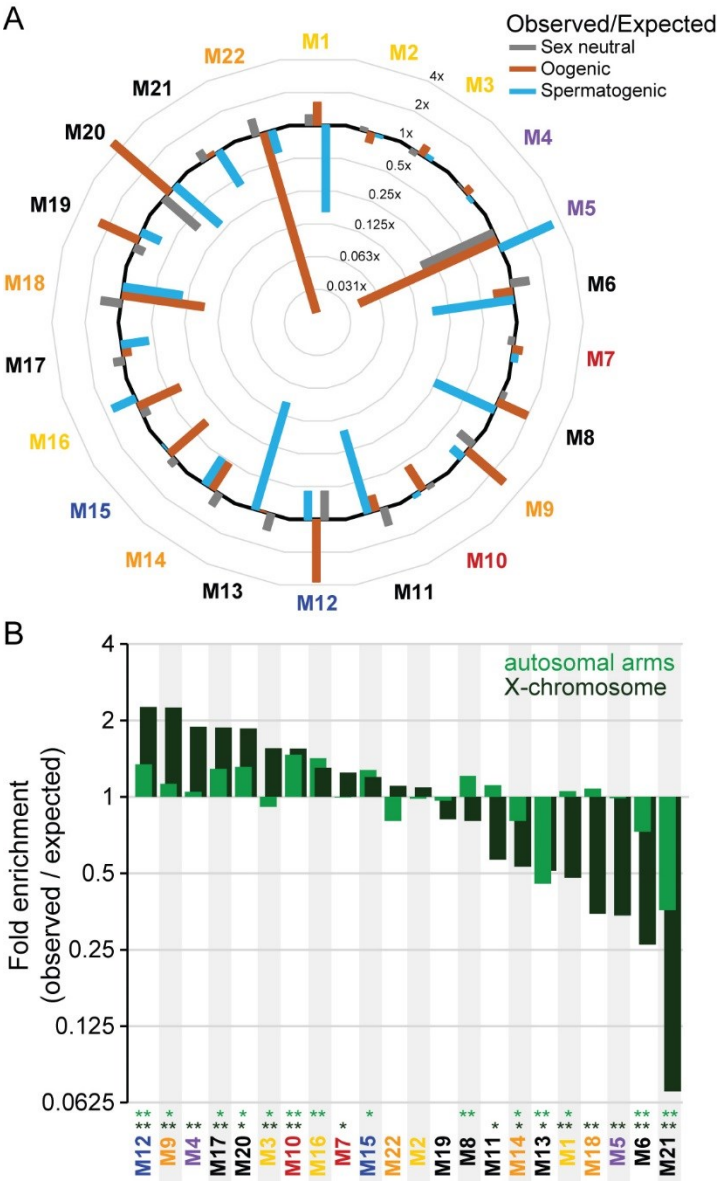


Figure 4. (**A**) Module gene over- and under-enrichment for *C. elegans* orthologs with spermatogenic, oogenic, or sex neutral expression profiles from Ortiz et al. (2014). Log-2 interval scale for observed/expected number of genes per module in radial plot, with black line indicating a value of 1 (outer curve indicates 4-fold enrichment, innermost curve indicates 2^{-5} under-enrichment). (**B**) Module gene over- and under-enrichment across the genome shows biases toward autosomal arms (values < 1 indicate enrichment in autosomal centers) and for X-linkage (values < 1 indicate enrichment on autosomes). Significant enrichment indicated by * (P<0.05 after Benjamini-Hochberg adjustment; FDR = 0.05) and ** (P<0.001 after Benjamini-Hochberg adjustment; FDR = 0.05). Coloring of module names in A and B corresponds to differential enrichment patterns indicated in Figure 3 (blue, T only; red, G only; purple, G&T; orange = GxT with crossing reaction norm profile; yellow, GxT with non-crossing reaction norm; gray, black, differential expression).

A

Module ID	Average expression	SD expression	Codon bias (ENC)	Median dN/dS'	% operonic	% TF	Average peptide length
M22	7.98	2.73	40.39	0.024	32.7	0.0	194.6
M18	7.74	2.33	40.87	0.019	46.3	2.0	260.7
M11	6.45	1.85	47.86	0.034	35.2	0.8	630.3
M21	6.28	1.09	51.63	0.046	60.4	1.0	629.2
M13	6.23	1.52	50.46	0.045	46.2	1.3	591.2
M14	6.23	2.21	47.69	0.033	40.4	1.3	238.5
M17	6.17	2.36	44.28	0.035	14.5	1.9	361.6
M8	5.20	1.93	51.04	0.047	25.1	10.9	1047.0
M6	4.94	2.04	51.04	0.060	44.6	3.3	258.2
M1	4.36	2.79	51.49	0.073	26.8	7.8	489.7
M4	3.38	2.59	48.75	0.040	3.2	4.9	546.2
M9	3.20	2.15	51.46	0.046	4.6	14.9	1077.0
M0	2.77	2.28	48.51	0.077	8.1	0.0	356.6
M2	2.41	2.70	50.01	0.060	5.9	4.4	289.6
M3	2.25	2.44	50.66	0.046	3.4	5.0	449.9
M19	2.10	2.03	51.14	0.043	7.3	9.1	313.6
M16	1.87	2.26	50.37	0.079	1.2	3.6	305.8
M5	1.69	2.57	50.09	0.072	2.4	2.3	369.7
M20	1.52	2.34	50.71	0.061	2.0	29.1	440.8
M7	1.37	2.43	50.13	0.055	2.2	3.9	354.4
M12	1.22	2.33	52.36	0.089	4.1	5.7	299.1
M10	0.91	3.46	50.11	0.088	1.8	3.3	338.4
M15	0.84	2.40	52.07	0.092	0.6	2.4	303.9

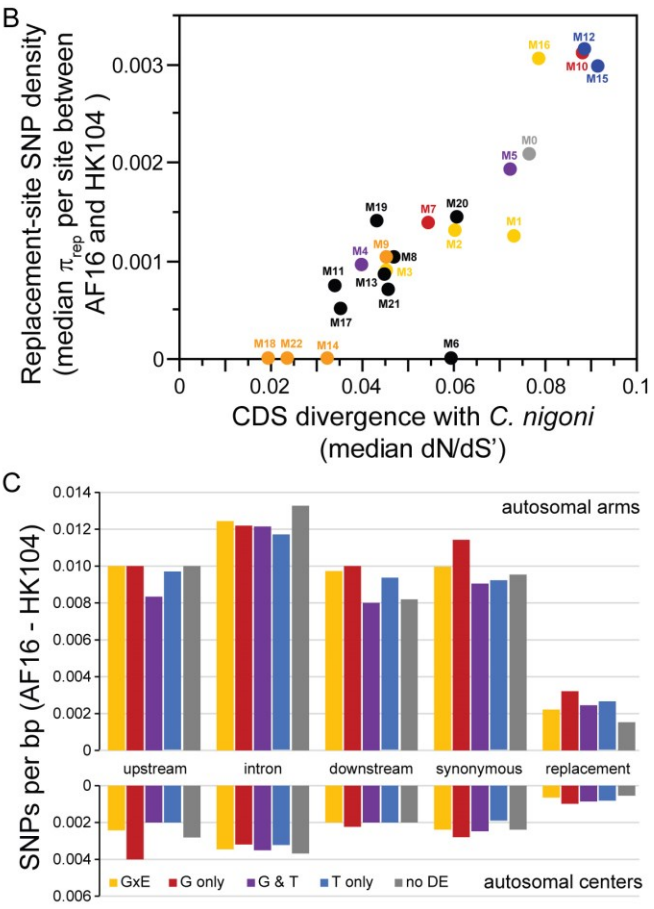


Figure 5. (**A**) Heatmap of module features, sorted by average normalized expression. Modules with high expression profiles tend to contain genes with stronger codon usage bias, greater sequence constraint (low dN/dS'), and more operons. (**B**) Modules with gene orthologs having little coding sequence divergence between *C. briggsae* and *C. nigoni* also have low densities of replacement-site SNPs in coding sequences. Coloring of module names in A and B corresponds to differential enrichment patterns indicated in Figure 3 (blue, T only; red, G only; purple, G&T; orange = GxT with crossing reaction norm profile; yellow, GxT with non-crossing reaction norm; gray, black, differential expression). (**C**) Among genes with distinctive profiles of individually-significant differential expression, linkage to autosomal arms versus centers represents the primary driver of SNP variation with little difference among DE categories for a given genomic site type (1kb upstream of CDS, intronic, 1kb downstream of CDS, synonymous sites).

Supplementary Information

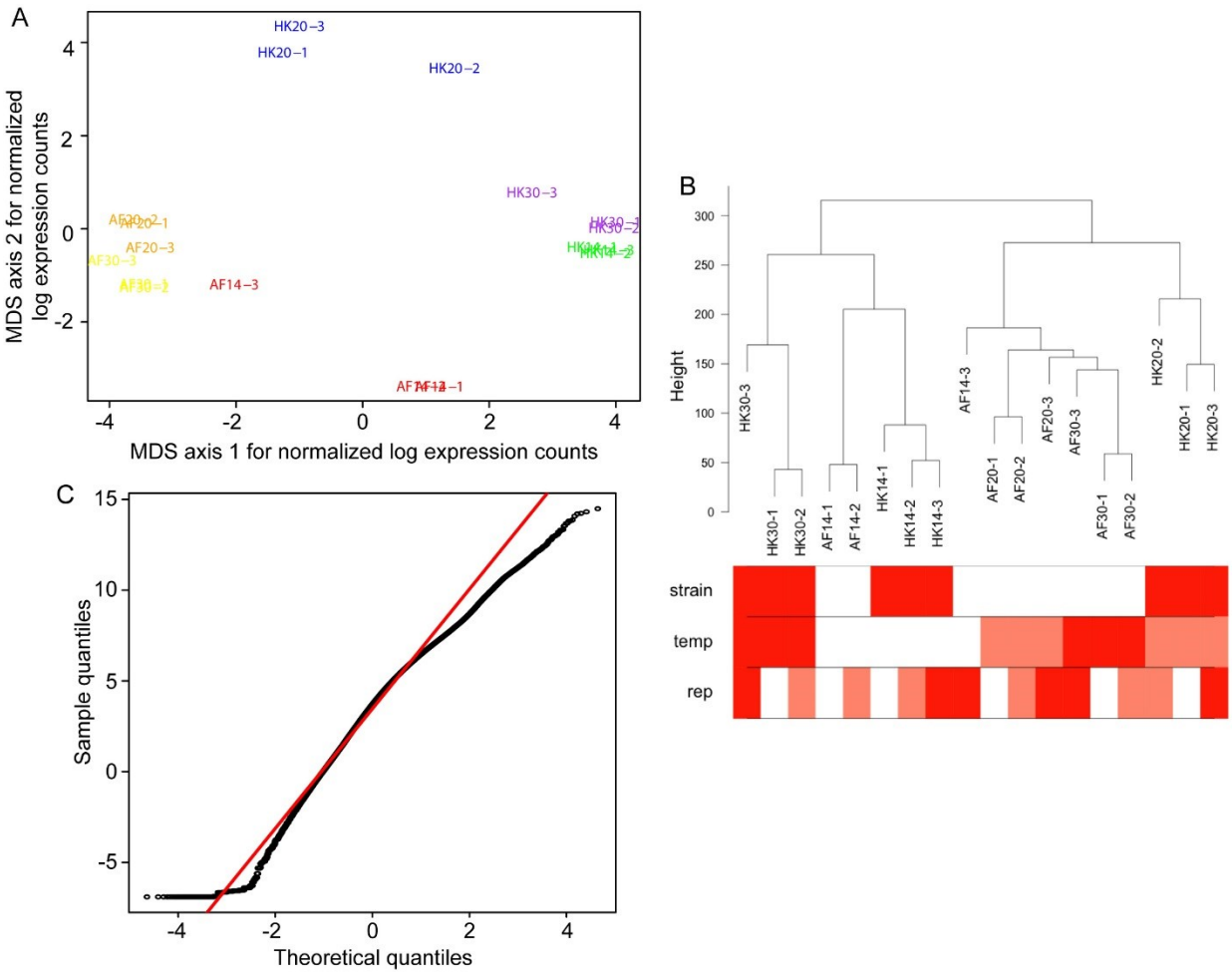
Supplementary File S1. "SuppFile_voomNormFiltLog2Expr_DE_modules.csv" contains log-2 normalized expression values for each of the 16,199 genes analyzed in each replicate sample, as well as the category of differential expression (G-only, T-only, G&T, GxT, noDE) and name of the co-expression module (M0 through M22). Columns labeled with sample name (treatment-replicate), where treatment is a combination of genotype and rearing temperature for each of three biological replicates ("AF" = Tropical AF16 genotype, "HK" = Temperate HK104 genotype; 14=14°C rearing, 20=20°C rearing, 30=30°C rearing).

Supplementary File S2. "SuppTable_GO.xlsx" contains lists of gene ontology term enrichment with summary statistics for different PANTHER GO-slim categories for each co-expression module.

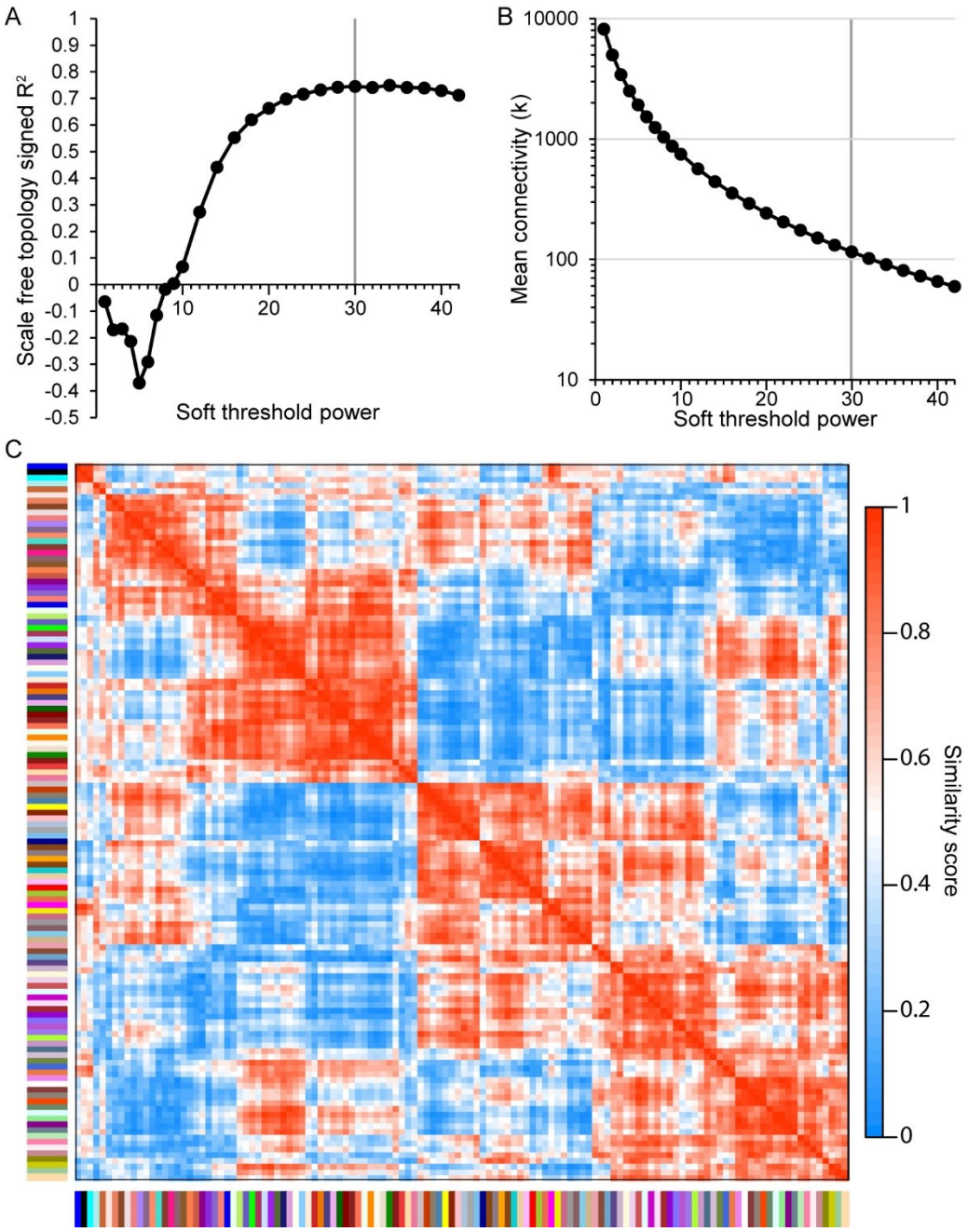
961 **Supplementary Table S1**. Number and percentage of reads that mapped to unique genomic
962 locations with STAR.
963

Sample (treatment-rep)*	Number of uniquely mapped reads	% of reads that uniquely mapped
AF14-1	32007586	93.35%
AF14-2	47012080	93.84%
AF14-3	54523984	94.15%
AF20-1	57787850	94.17%
AF20-2	62122726	93.43%
AF20-3	43596994	93.75%
AF30-1	66295835	93.74%
AF30-2	35618842	93.20%
AF30-3	33087508	90.10%
HK14-1	51264144	93.39%
HK14-2	43616826	93.65%
HK14-3	48035023	92.74%
HK20-1	30262908	90.76%
HK20-2	49528582	93.95%
HK20-3	41332043	94.26%
HK30-1	51537189	91.44%
HK30-2	40748065	92.22%
HK30-3	35632338	73.62%

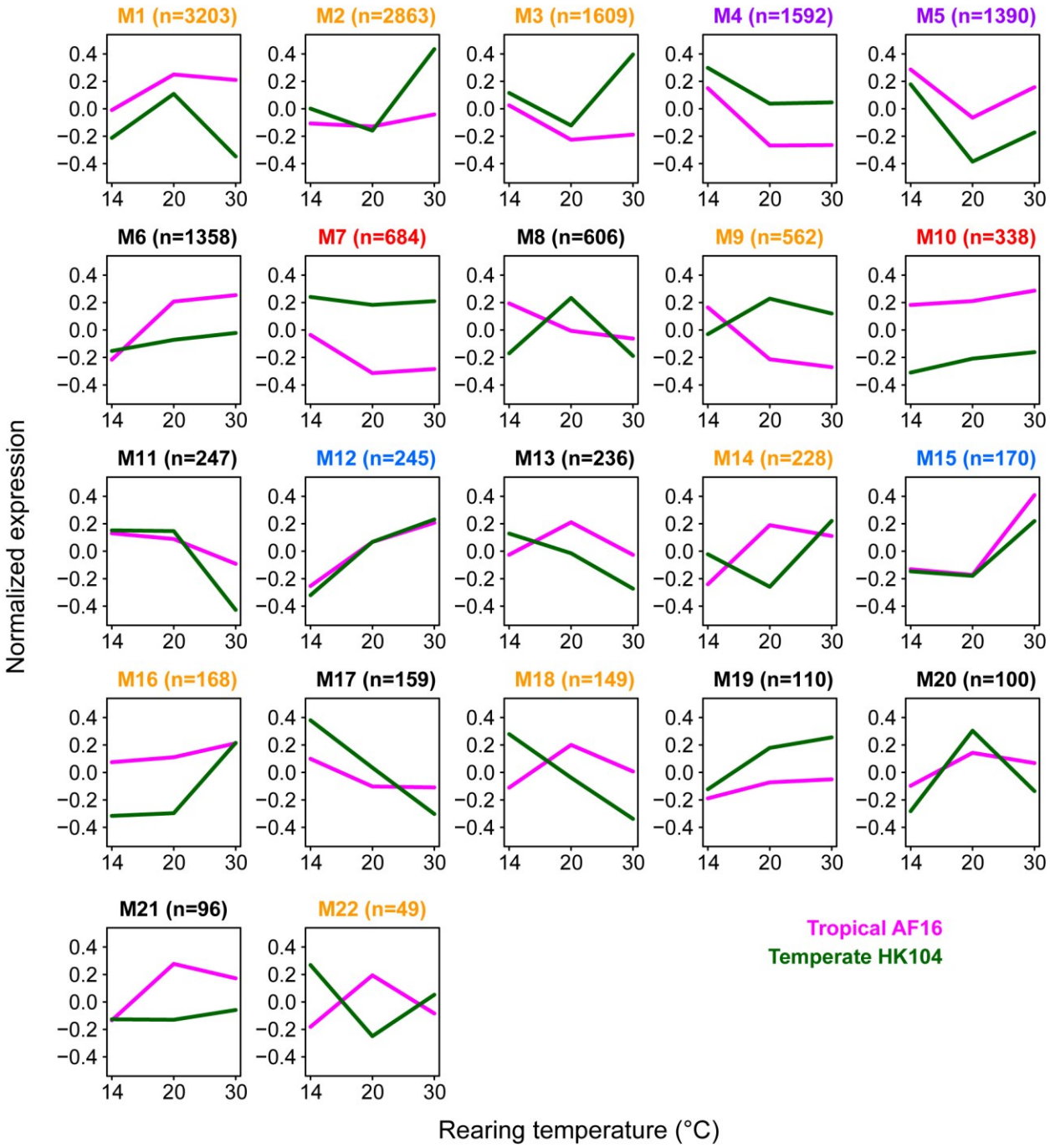
964 *AF=AF16, HK=HK104 genotypes; 14=14°C, 20=20°C, 30=30°C rearing



Supplementary Figure S1. (A) Multi-dimensional scaling plot (MDS) of filtered, normalized, and log-transformed count data for each sample. Sample labels (treatment-replicate) are colored by experimental treatment for a given genotype and rearing temperature combination ("AF" = Tropical AF16 genotype, "HK" = Temperate HK104 genotype; 14=14°C rearing, 20=20°C rearing, 30=30°C rearing). The x- and y-axes show the two principal components that explain most variation. Biological replicates that cluster together in the plot are more similar to each other, indicating consistency across replicates. (B) After filtering out genes with very low to no counts, TMM normalization for different library sizes, and log-transforming count data, a dendrogram reveals similarity of samples within strain and within temperature ("temp"), but not replicate ("rep"); red shading along a row indicates shared strain, temperature or replicate value. This suggests data heterogeneity among samples due to experimental treatments, not batch effects. (C) Quantile-quantile plot for normalized, voom-transformed count data shows good approximation to a normal distribution (red line).



Supplementary Figure S2. (A) Analysis of soft-thresholding powers revealed 30 to be the power at which the scale-free fit is maximized (R^2 = 0.75) and most closely approximates a scale-free network. (B) Analysis of a range of soft-thresholding powers revealed a value of 30 to have a mean connectivity of at least 100 (k = 115), while also maximizing fit to a scale-free network. (C) Clustering of 16,199 genes with WGCNA into a preliminary set of 124 co-expression clusters (red in heatmap indicates maximum similarity and blue no similarity; color bars along the x- and y-axes correspond to the 124 clusters). Clusters with similarity distance <0.25 were merged, leading to the 22 co-expression modules (plus pseudo-module M0) analyzed in the main text.



Supplementary Figure S3. Co-expression module eigengene plots of normalized, log_2-transformed expression across temperature treatments for each genotype (pink AF16, green HK104). Module names colored as in Figure 3 of the main text according to representation of individually significant differentially expressed genes within the module.

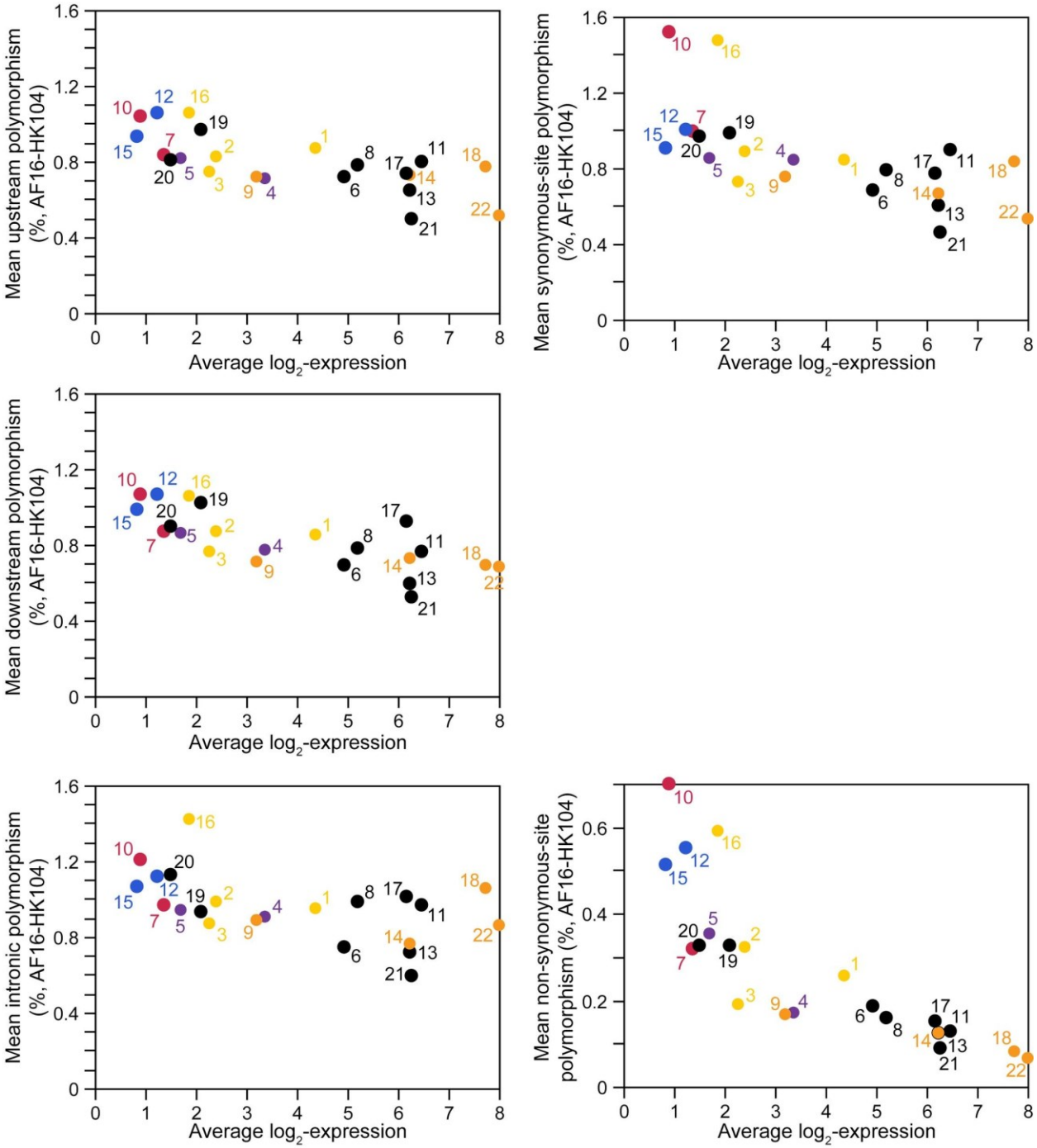
A Observed / Expected Number of Genes per Chromosome

Module	I	II	III	IV	V	X
M1	**1.433**	1.127	**1.249**	1.072	**0.745**	**0.468**
M2	**0.612**	0.876	**0.716**	1.011	**1.506**	1.110
M3	**0.619**	0.911	**0.740**	0.882	**1.230**	**1.596**
M4	**0.611**	0.954	**0.762**	0.872	0.981	**1.866**
M5	**1.471**	1.161	**1.314**	0.996	0.797	**0.327**
M6	**1.297**	1.140	**1.498**	1.119	0.799	**0.258**
M7	0.784	1.071	**0.495**	0.954	1.286	1.273
M8	1.282	0.953	**1.493**	1.072	**0.550**	0.783
M9	0.976	0.674	0.711	0.826	**0.643**	**2.255**
M10	**0.482**	0.920	0.664	0.709	1.359	1.478
M11	1.167	1.259	1.322	1.021	0.776	0.582
M12	0.665	0.932	0.500	0.926	0.721	**2.244**
M13	1.221	1.075	1.269	1.256	0.706	0.476
M14	1.154	1.057	1.224	1.300	0.841	0.548
M15	0.663	0.784	1.081	1.298	0.950	1.176
M16	0.448	0.642	0.567	0.563	**2.043**	1.153
M17	0.552	0.838	0.514	0.952	1.174	**1.807**
M18	1.682	0.852	1.370	1.185	0.644	0.419
M19	1.196	0.923	0.742	1.433	0.734	0.909
M20	1.095	1.047	0.793	0.857	0.441	**1.941**
M21	1.566	1.123	**2.268**	0.854	0.473	**0.065**
M22	1.662	0.906	0.833	1.544	0.308	1.020

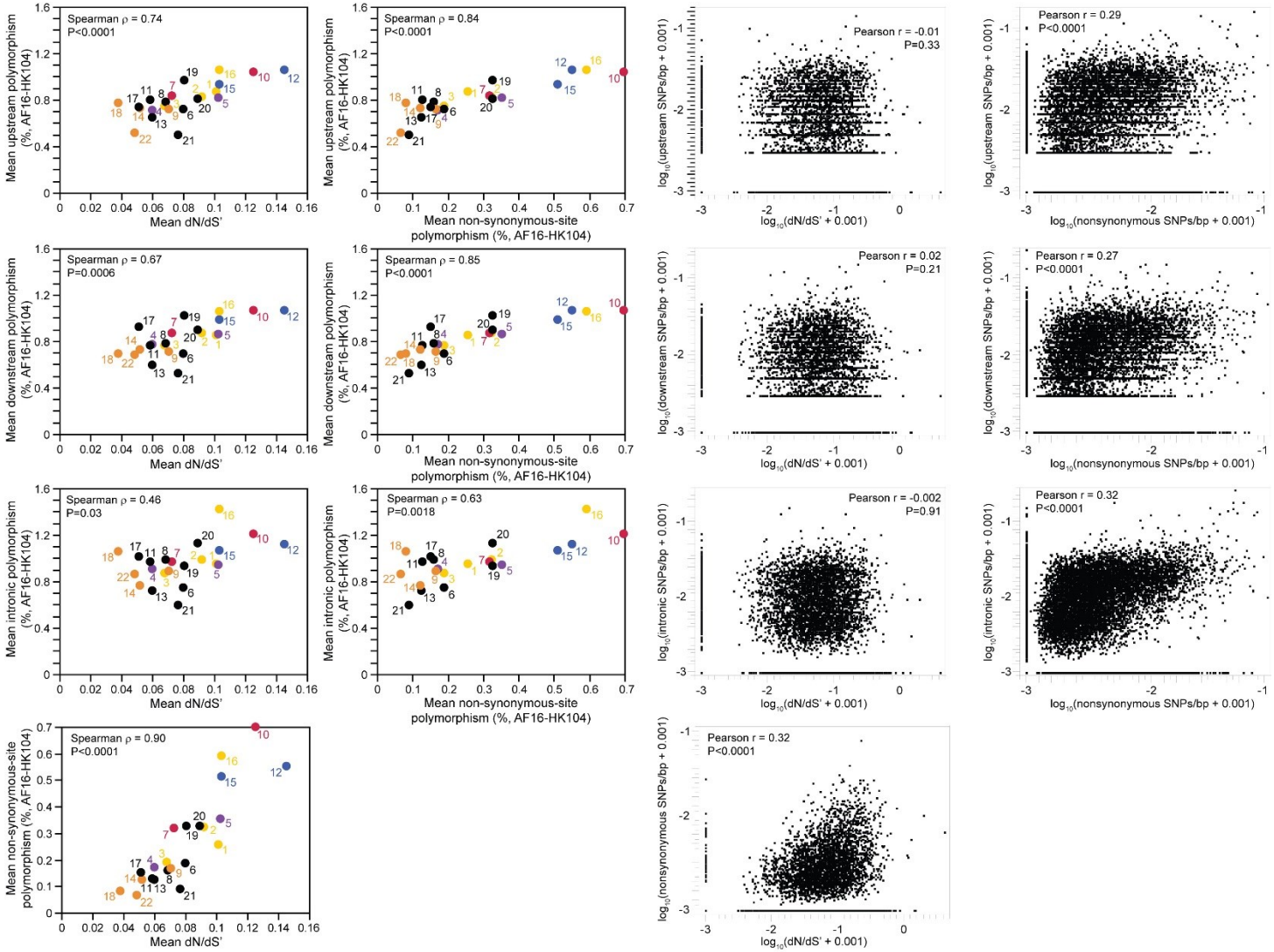
B

	I	II	III	IV	V	X
T only	1.054	1.108	1.018	0.994	0.915	0.950
G only	0.819	0.952	0.722	0.832	1.195	1.236
G&T	1.170	1.083	1.064	1.099	0.843	0.832
GxT	0.984	1.053	0.918	1.038	**1.111**	**0.859**
no DE	0.989	0.928	1.069	0.979	0.952	**1.108**

Supplementary Figure S4. Enrichment (observed/expected) number of genes on each chromosome for modules (A) and differential-expression categories (B). Values in black bold text indicate significant enrichment or under-enrichment after Bonferroni multiple-test correction (χ^2 test df=1 with α=0.05/132 for modules, α=0.05/30 for categories).



Supplementary Figure S5. SNP density in upstream, downstream, and intronic non-coding locations of genes and at synonymous and non-synonymous sites within coding sequences, averaged for genes within each co-expression module as a function of average module expression. Correlation across modules: module mean $\pi_{nonsynonymous}$ × average expression Spearman ρ = -0.92, P<0.0001; $\pi_{nonsynonymous}$ × ENC ρ = 0.418, P=0.048; module mean π × average expression Spearman ρ, $\rho_{upstream}$ = -0.72, P<0.0001; $\rho_{downstream}$ = -0.77, P<0.0001; $\rho_{intronic}$ = -0.54, P<0.0082. Points are labeled and colored with module name as in Figure 3 of the main text.



Supplementary Figure S6. SNP density in upstream, downstream, and intronic non-coding locations of genes and at non-synonymous sites within coding sequences, averaged for genes within each co-expression module (left panels) or per gene (right panels) as a function of average module interspecies divergence (dN/dS') or non-synonymous site substitution. Modules are labeled and colored as in Figure 3 of the main text.