Hybrid Incompatibilities and Transgressive Gene Expression Between Two Closely Related Subspecies of Drosophila

Alwyn C. Go and Alberto Civetta*

Department of Biology, The University of Winnipeg, Winnipeg, MB, Canada

Genome-wide assays of expression between species and their hybrids have identified genes that become either over- or underexpressed relative to the parental species (i.e., transgressive). Transgressive expression in hybrids is of interest because it highlights possible changes in gene regulation linked to hybrid dysfunction. Previous studies in Drosophila that used long-diverged species pairs with complete or nearly complete isolation (i.e., full sterility and partial inviability of hybrids) and high-levels of genome misregulation have found correlations between expression and coding sequence divergence. The work highlighted the possible effects of directional selection driving sequence divergence and transgressive expression. Whether the same is true for taxa at early stages of divergence that have only achieved partial isolation remains untested.

Here, we reanalyze previously published genome expression data and available genome sequence reads from a pair of partially isolated subspecies of Drosophila to compare expression and sequence divergence. We find a significant correlation in rates of expression and sequence evolution, but no support for directional selection driving transgressive expression in hybrids. We find that most transgressive genes in hybrids show no differential expression between parental subspecies and used SNP data to explore the role of stabilizing selection through compensatory mutations. We also examine possible misregulation through cascade effects that could be driven by interacting gene networks or co-option of off-target cis-regulatory elements.

Keywords: regulatory co-option, network interactions, speciation, selection, transgressive gene expression, compensatory evolution, gene expression divergence

INTRODUCTION

Studies that have addressed the genetic basis of incompatibilities in hybrids between species, or diverging populations, have traditionally resorted to mapping loci and interactions between them (Coyne and Orr, 1989; Masly and Presgraves, 2007; Presgraves, 2008; Victoria Cattani and Presgraves, 2012; Dufresnes et al., 2016). This approach has been fruitful in that ultimately a few major protein-coding genes have been identified (Ting et al., 1998; Masly et al., 2006; Mihola et al., 2009; Phadnis and Allen Orr, 2009), and in some cases, the effect of these major genes require interactions with other genetic factors. Major genes often show patterns of rapid evolution between divergent populations or species (Ting et al., 1998; Presgraves et al., 2003;
suggested that, at least in part, changes in protein composition might exert effects on phenotype and function through alterations in patterns of expression of genes targeted by such proteins. Moreover, genome-wide surveys have provided evidence to support that many genes and complex systems of epistasis are linked to hybrid incompatibility phenotypes (Morán and Fontdevila, 2014; Turner and Harr, 2014; Turner et al., 2014; Fontdevila, 2016). While co-evolution among interacting genes keeps function within populations and species, hybridization between divergent isolated populations and incipient species brings together incompatible interloci allele interactions resulting in a reduction in hybrid fitness (Dobzhansky, 1937; Muller, 1942; Orr, 1996). The reduced fitness of hybrids serves as a postzygotic barrier among divergent taxa.

The role of divergence in the regulation of gene expression has been long acknowledged (King and Wilson, 1975), but not until recently has genome-wide divergence in gene expression during speciation been addressed. Recent reviews have summarized how changes in gene expression could impact hybrid phenotypes (Civetta, 2016; Mack and Nachman, 2017). Using genome-wide approaches, questions have been addressed as to the proportion of genome-wide misregulation in hybrids, the relative contribution of cis- vs. trans-regulatory elements in gene misregulation, and the identity of misregulated genes that might contribute to hybrid fitness breakdown (Ranz et al., 2004; Haerty and Singh, 2006; Renaut et al., 2009; Tirosi et al., 2009; McManus et al., 2010; Llopart, 2012; Coolon et al., 2014; Gomes and Civetta, 2015; Brill et al., 2016; Mack et al., 2016). Often, genome-wide assays of expression in hybrids reveal gene regulatory dysfunctions as patterns of transgressive gene expression (i.e., expression beyond levels found in parental species). This can be a consequence of directional selection or drift causing changes at cis- and trans-regulatory elements that drive divergence in expression between taxa and transgressive expression in hybrids. Previous studies have found positive correlations between protein-coding evolution and gene expression divergence between species of Drosophila (Castillo-Davis et al., 2004; Nuzhdin et al., 2004; Lemos et al., 2005; Artieri et al., 2007). Moreover, the finding of a similar significant positive correlation between non-synonymous (\(d_N\)) and non-synonymous/synonymous (\(d_N/d_S\)) divergence and gene expression differences between hybrids and parental species has been used to suggest sequence divergence driving regulatory incompatibilities and to highlight the potential effects of directional selection in gene expression during speciation (Artieri et al., 2007; Hunt et al., 2013). However, the species pairs used were typically long diverged with hybrids exhibiting complete or nearly complete isolation and high levels of genome misregulation (Ranz et al., 2004; Haerty and Singh, 2006; Artieri et al., 2007; McManus et al., 2010; Coolon et al., 2014). The use of divergent populations within species of copepods have found no significant relationship between hybrid transgressive expression and estimates of sequence divergence, and the authors offered an alternative physiological explanation for the detected pattern (Barreto et al., 2015, 2018).

There are, in fact, alternative explanations that could explain the lack of relationship between sequence and expression divergence. Mutations within taxa can work to compensate the effect of deleterious mutations on expression (i.e., stabilizing selection). The possibility that cis–trans mutations may cause compensation within species but lead to transgressive expression in hybrids is supported by studies that report abundant cis–trans epistasis (Mackay, 2014; Mackay and Moore, 2014; He et al., 2016; Vonesch et al., 2016). However, the strength of selection for a secondary compensatory mutation might be small (Bourguet, 1999). It is also possible for transgressive expression in hybrids to arise as a response to hybrid dysfunction within gene-interacting networks or metabolic pathways. While this could work to ameliorate fitness problems in hybrids, it could also exacerbate hybrid dysfunction. This might be particularly the case for fitness breakdown between diverging populations (Barreto et al., 2015, 2018). Finally, we speculate that newly arising mutations in trans regulatory elements that result from divergence between taxa or compensatory mutations within, could co-opt preexisting cis-regulatory elements among multiple genes thereby causing widespread misregulation.

Here, we used a pair of geographically separated subspecies of D. pseudoobscura, D. p. pseudoobscura, and D. p. bogotana that have diverged for at least 0.15 Myr (Schaeffer and Miller, 1991; Wang et al., 1997) and whose hybrids exhibit unidirectional male sterility where only male hybrids produced by D. p. bogotana females are sterile. We reanalyze previously published transcriptomics data (Gomes and Civetta, 2015) using a newer D. p. pseudoobscura genome release (r3.04) and updated mapping and expression analysis tools to explore relationships between genome expression and gene-coding sequence divergence. Our report identifies no relationship between sequence divergence and transgressive expression in hybrids suggesting a need for broader examinations of transgressive expression between recently diverged populations and species across taxa. We find that most transgressive genes in hybrids are not differentially expressed between subspecies. We explore explanations for transgressive expression other than incompatibilities in regulation arising from rapid divergence between subspecies, such as compensatory mutations, gene-interaction networks, and the co-option of multiple cis-regulatory elements by trans-regulatory elements. While we find some support for these alternative hypotheses, we acknowledge that they do not fully explain transgressive expression in hybrids. We discuss some caveats and offer other possible explanations in the hope that they will trigger further inquiry. Ultimately, full comprehension of transgressive expression in hybrids will require combining information on genome expression and sequencing with the identification of interactomes and a proper characterization of mechanism of trans effects on characterized cis-regulatory targets.

MATERIALS AND METHODS

RNA-Sequence Data

Raw RNA sequence data used in this analysis were from a genome-wide transcriptomics study of the Drosophila pseudoobscura subspecies pair and their reciprocal hybrids by...
Gomes and Civetta (2015). Briefly, RNA was extracted from the whole male reproductive tract. Biological replicates were obtained for the parental subspecies and their reciprocal F1 hybrids with each replicate containing 30–40 male reproductive tracts. cDNA libraries were prepared using the Illumina TruSeq Stranded mRNA sample preparation kit and multiplexed on a single lane of an Illumina HiSeq2000 platform with 100 bp paired-end sequencing. A quality check was performed on the raw reads using FastQC (Andrews, 2010). Read processing and adapter trimming were performed with Trimmmomatic (Bolger et al., 2014), and reads with a Phred score below 30 and a final length of 50 bp were excluded.

Mapping and Differential Expression Analysis

We mapped processed reads to the latest release (r3.04) of the D. p. pseudoobscura reference genome1 using STAR, chosen for its reliability (Dobin et al., 2013; Baruzzo et al., 2017) over the previously used TopHat approach (Gomes and Civetta, 2015). Read counting was performed at the gene level using featureCounts (Liao et al., 2014) with the reversely stranded (-s 2) and fragment counting (-p) parameters and the latest version of the D. p. pseudoobscura annotation serving as a guide.

Pairwise differential expression across all groups was performed using both DESeq2 (Love et al., 2014) and edgeR (Robinson et al., 2009). In the analysis using edgeR, genes with less than 1 count per million (CPM) in at least one group were excluded from further analysis, and the per gene counts for each sample were normalized using the TMM method (Robinson and Oshlack, 2010). The default settings were used to obtain normalized counts from the DESeq2 analysis. The consensus list of differentially expressed genes from both tools were used for all downstream analyses. Differentially expressed genes among the hybrids were identified as transgressive if their expression was significantly above or below the range found in the parental subspecies. Further, log2 fold-changes (lfc) thresholds of 0.5 and 1 were applied to increase our statistical yield of true positives (Schurch et al., 2016). All tools for the analysis were ran on Galaxy2.

Coding Sequence and Expression Divergence

Rates of coding sequence divergence between D. p. bogotana and D. p. pseudoobscura were estimated for differentially expressed genes between the parental subspecies and for transgressive genes in fertile and sterile F1 hybrids. Since the RNA-seq data provided only partial sequences from each gene analyzed, we retrieved genes between the parental subspecies and for transgressive genes D. p. pseudoobscura from the r.3.04

http://flybase.org/

http://usegalaxy.org
a significant result for the Fisher’s exact test indicated evidence for compensatory cis and trans mutations (McManus et al., 2010), while a non-significant result suggested a conservation in regulatory interactions and classified as non-compensatory.

**Interactions and Sequence Similarity**

Interactions among proteins were predicted using STRING (v11.0; Szklarczyk et al., 2019). Gene-Ontology and UniProt keyword enrichments were assessed from outputs using STRING and DAVID (v6.8; Huang et al., 2009a,b). We used the extended gene regions (which includes 2-kb 5′ and 3′) for genes that showed transgressive expression driven by trans regulatory elements (i.e., cis–trans divergent or compensatory) to perform a BLASTn against a database containing all transgressive genes and against another database with all D. p. pseudoobscura extended gene regions within the genome to identify similarities between upstream regions for plus/plus matches or between the upstream and downstream regions for plus/minus matches. We retained only hits that were lower than $1 \times 10^{-14}$ and unique among transgressive sequences and not shared with other genes in the genome. Retained hits had E-values lower than $8 \times 10^{-13}$, with nucleotide alignments of at least 173 base pairs and identities higher than 64%.

**RESULTS**

The re-analysis of our previously published data (Gomes and Civetta, 2015) by mapping reads onto a newer released genome assembly and using more recently developed analytical pipelines found similar results in terms of lack of bias in mapping, low proportion of differentially expressed genes between subspecies, and significant excess of transgressive expression in sterile relative to fertile hybrids (Supplementary Material).

**Transgressive Gene Expression in Hybrids Does Not Correlate With Accelerated Rates of Evolution as Expected Under a Scenario of Divergent Selection Between Subspecies**

Under the assumption that regulatory evolution and structural protein evolution are under similar selective pressures, a correlation is expected between expression difference and nucleotide sequence evolution. Of the 819 differentially expressed genes between the parental subspecies, 604 (73.7%) were protein-coding genes with the remaining 215 (26.3%) being non-coding RNAs or coding genes without full coding sequences available for both subspecies. The percentage of differentially expressed protein-coding genes between subspecies increases significantly when a less stringent lfc threshold of 0.5 was applied (82.7%; $Z = 5.51, P < 0.001$) (Supplementary Figure 1). We found a significant correlation for expression differences between subspecies and non-synonymous ($d_N$) sequence divergence ($N = 604$; Spearman’s $\rho = 0.091, P = 0.026$) but not between differences in expression and synonymous substitutions ($d_S$) (Spearman’s $\rho = -0.046, P = 0.261$). The $d_N/d_S$ ratio was also positively correlated with expression differences ($\rho = 0.108, P = 0.011$). Using the less stringent lfc threshold of 0.5, $d_N$, $d_S$, and $d_N/d_S$ were all significantly correlated with gene expression divergence between subspecies ($N = 1,801$; $\rho = 0.121, P = 2.39 \times 10^{-7} ; \rho = 0.065, P = 0.005$; and $\rho = 0.096, P = 8.4 \times 10^{-5}$, respectively) (Figure 1A). These results are overall in agreement with previous findings in Drosophila and other organisms confirming that protein sequence and expression divergence are influenced by similar selective processes (Castillo-Davis et al., 2004; Nuzhdin et al., 2004; Khaitovich et al., 2005; Artieri et al., 2007; Ortiz-Barrientos et al., 2007).

Given that protein-coding sequence differentiation serves as a good predictor of expression divergence, some studies
have explored correlations between rates of protein divergence with expression of misregulated genes in hybrids. Misregulated genes with transgressive expression in hybrids are of interest in speciation as they associate with hybrid disrupted phenotypes (Moehring et al., 2007; Catron and Noor, 2008; Sundararajan and Civetta, 2011; Gomes and Civetta, 2015; Brill et al., 2016; Civetta, 2016). Significant positive correlations are suggestive of either directional selection or relaxation of selective constraints fueling regulatory incompatibilities (Artieri et al., 2007; Hunt et al., 2013; Barreto et al., 2015). Of the 44 transgressive genes in the hybrids, 35 had available sequence data for the estimation of coding sequence divergence. The analysis showed no significant correlations between sequence divergence and expression difference ($N = 35$; $d_{N}$, $\rho = 0.078$, $P = 0.655$; $d_{S}$, $\rho = 0.457$, $P = 0.144$).

![FIGURE 2](scenarios.png)

**FIGURE 2** | Scenarios of regulatory divergence for cis and trans regulatory elements. (A) Gene 1 shows compensatory cis and trans mutations, denoted by asterisks, wherein $D$. $p$. bogotana experiences an initial mutation in cis followed by a mutation in trans restoring gene expression to similar levels between parental subspecies. Gene 2 shows similar levels of expression in parental subspecies. In the hybrid background, the $D$. $p$. pseudoobscura trans factor for gene 1 interacts with the $D$. $p$. bogotana trans factor for gene 2 leading to a conformation change. This new trans factor complex can now bind optimally to the cis region of genes 1 and 2 (red arrows) resulting in transgressive expression (i.e., expression above parental levels). The allelic ratio of gene 2 in the hybrid is equal to the allelic ratio of the two subspecies, and the gene is classified as non-compensatory through SNP analysis. (B) Gene 3 shows divergence in cis in one subspecies and trans in the other subspecies. This leads to suboptimal binding in both subspecies and differential expression. The regulatory incompatibilities persist within the hybrid background leading to unequal allelic ratios. Gene 3 is classified as cis–trans divergent by SNP analysis. Gene 4 shows a situation of cis-only divergence between the parental subspecies. Regulatory incompatibilities would occur in $D$. $p$. bogotana but not $D$. $p$. pseudoobscura resulting in differential expression between the subspecies. Similar interactions for this gene would occur in the hybrid resulting in equal allelic ratio. Gene 4 is classified as cis-only by SNP analysis.
Alternative Explanations for Transgressive Expression in Hybrids: Compensatory Mutations, Interaction Networks, and Transcriptional Drive by Sequence Similarity Among Targets

One possibility for a lack of correlation between transgressive expression in hybrids and sequence divergence is that transgressive expression might be a consequence of occasional deleterious mutations that are followed by compensatory DNA changes to overcome detrimental effects on gene expression (i.e., a side effect of stabilizing selection between divergent taxa) (Figure 2A–Gene 1). Our data show that 32 out of 44 (72.72%) transgressive genes in the hybrids were not differentially expressed between parental subspecies. The low number of transgressive genes is likely a consequence of our stringent use of a twofold change (lfc = 1) in expression threshold to maximize our statistical yield of true positives. Given the low sample size, we decided to continue using a less stringent lfc threshold of 0.5 and found, as with the more stringent threshold, a large proportion of transgressive genes without differential expression between parental subspecies (79%, 207/262). If genes without differential expression between subspecies are under stabilizing selection favoring compensatory mutations to buffer deleterious mutations and restore expression to similar levels among parental subspecies, we expect their rate of sequence divergence to be lower than those of genes experiencing divergence in regulation, and thus expression, between subspecies. Our data shows no significantly lower rates of change (dN and dS/dN) for genes with transgressive expression in hybrids and no differential expression between parents (Mann–Whitney FDR corrected P-values) (Table 1).

We used informative SNPs to identify genes with transgressive expression in hybrids driven by compensatory mutations or cis–trans divergence (Figures 2A,B–Genes 1 and 3). Twenty-five percent of the transgressive genes (65/262) had non-informative SNPs to allow us to classify parent of origin for the alleles found in the hybrids. Of the remaining 197 transgressive genes, we found that for 65% of them, transgressive expression could be explained by compensatory mutations (97 genes) or cis–trans divergence (31 genes) (Figures 2A,B–Genes 1 and 3). The remaining are cases in which the transgressive gene shows similar ratios of subspecies allele expression in parents and hybrids. Of these, 62 were classified as non-compensatory and 7 as having experienced cis divergence (Figures 2A,B–Genes 2 and 4).

We explored whether transgressive expression in hybrids for genes that do not show evidence of compensatory or cis–trans mutations could be a cascade triggered by interactions in a shared gene network and/or pathway (Bader et al., 2015; Barreto et al., 2015). This will predict clusters of interacting and functionally related proteins to be misregulated in the hybrids.

| Table 1 | Average evolutionary rates (±SD) for differentially expressed genes between parental subspecies that do not show transgressive expression in hybrids ([P1 ≠ P2]κT), transgressive genes that show differential expression between subspecies ([P1 ≠ P2]κ), and transgressive genes that do not show differential expression between subspecies ([P1 = P2]κ).

We detected a protein–protein interaction (PPI) network of 90 genes (34% of the 262 transgressive genes) (Figure 3) with a significant (i.e., more interactions than randomly expected) PPI enrichment (P = 4.29 × 10−2). We found no evidence of known functional enrichment in the network, but a significant overrepresentation of “Signal” genes based on UniProt keywords (FDR corrected P = 1.25 × 10−7). More importantly, when we partitioned the network analysis by sterile vs. fertile hybrids, the PPI analysis was significant for sterile hybrids (PPI enrichment P = 1.06 × 10−2, 79 nodes) but not for fertile hybrids (PPI enrichment P = 0.106, four nodes). Twenty-two genes in the network were cis or non-compensatory, thus their misregulation could be driven by interactions with other misregulated genes in the network (Figure 3). We found no significant PPI for transgressive genes differentially expressed between subspecies (P = 0.597). Finally, we explored whether transgressive expression in hybrids could be a consequence of what we refer to as “transcriptional drive.” That is, the ability of a modified trans factors to control and drive the transcription of multiple prior non-target genes with cis-sequence similarity (Figure 2A–red arrows). We found 46 genes (18%–46/262) with possible evidence of co-option by newly evolved trans mutations. Of these genes, 15 were classified as compensatory, 10 had cis–trans divergence, 9 were non-compensatory, and 12 had non-informative SNPs for classification (Supplementary Table 1).

**DISCUSSION**

Genome-wide, our results are in agreement with previous reports of correlated evolution between sequence and expression divergence (Castillo-Davis et al., 2004; Nuzhdin et al., 2004; Khaitovich et al., 2005; Lemos et al., 2005; Artieri et al., 2007; Hunt et al., 2013; Whittle et al., 2014; Barreto et al., 2015), but provide no support for positive selection or relaxation of selective constraints as drivers of change causing misregulation and transgressive expression in hybrids. It is possible that the lack of correlation for hybrids results from the relatively...
low divergence between the subspecies. However, we did find a significant correlation for expression differences between subspecies. Genes with no differential expression between subspecies and transgressive expression in hybrids did not show overall evidence of lower sequence divergence than transgressive genes with differential expression between subspecies. This result is unexpected under a scenario of compensation favoring mutations that restore divergence in gene expression between...
Go and Civetta

Go and Civetta Transgressive Gene Expression in Hybrids

parental subspecies (i.e., stabilizing selection). We used SNPs to tease apart regulatory divergence among transgressive genes in hybrids. Transgressive expression results from divergence in cis and trans regulatory elements, leading to differential expression between parental species as well as hybrids. Alternatively, such changes can be buffered by compensatory mutations within lineages to restore levels of expression to similar levels between species but cause misexpression in hybrids (Landry et al., 2005; McManus et al., 2010; Mack and Nachman, 2017). Studies of divergence in gene expression between species provides support for changes in transcript levels being often deleterious, with large mutational effects, and equilibrium levels of genetic variation maintained by stabilizing selection (Rifkin et al., 2003; Lemos et al., 2005; Hodgins-Davis et al., 2015). Our study shows that the majority (79%) of transgressive genes in hybrids between D. p. pseudoobscura and D. p. bogotana were not differentially expressed between the subspecies, and the SNP analysis supports a good proportion of transgressive expression caused by compensatory changes (49%) during early stages of species divergence, with another (16%) caused by cis–trans divergence.

Informative SNPs are limited between closely related subspecies. Thus 25% of transgressive genes could not be analyzed this way. Moreover, for any gene, not all reads have informative SNPs imposing some analytical limitations. While this might lead to an underestimation, our result of 49% compensatory evolution for a pair of very closely related subspecies of Drosophila is expected when compared to estimates of 73% compensatory evolution for hybrids between more distantly related species of D. simulans and D. sechellia (Coolon et al., 2014) and 67% for yeast (Wang et al., 2015). The proportion of compensatory mutations within lineage (49%) is larger than cis–trans divergence between lineages (16%) and suggests that hybrids between closely related taxa might be more vulnerable to a breakdown of co-adaptations within species than misregulation caused by divergent evolution. Two important caveats are that we used one line per subspecies, which might underestimate true divergence between these subspecies by overlooking possible shared polymorphisms, and that there is inherent bias toward a possible overestimate of the role of compensatory evolution when using an ASE approach (Fraser, 2019; Zhang and Emerson, 2019). Therefore, it is important to explore possible alternative explanations for a large proportion of transgressive genes, which could not be explained by cis–trans compensation or divergent cis–trans evolution.

We found that genes with transgressive expression in hybrids that experienced divergence in regulation between subspecies produced proteins that did not show enrichment for interactions. On the other hand, transgressive genes with no evidence of divergence between subspecies were enriched for protein interactions, particularly for the sterile hybrids. The overrepresentation of protein interactions among transgressive genes may be due to possible tissue or cell type atrophy in the hybrids. However, these two subspecies are very closely related, and previous work have shown no evidence of tissue atrophy in the hybrids (Gomes and Civetta, 2014). Furthermore, we did not find an overrepresentation of underexpressed genes as expected under tissue/cell atrophy. This result suggests that in some cases, misregulation and transgressive expression could be a cascade effect driven by networks of interacting proteins and that such domino effect could work to exacerbate initial incompatibilities in hybrids between early stage diverging lineages. The role of gene-network effects is expected under the Bateson–Dobzhansky–Muller model of speciation (Turner et al., 2014), and while there has been some support for gene-networks buffering allelic variation among yeast strains (Bader et al., 2015), its importance in speciation is largely unexplored. Finally, we entertained the idea that newly arising trans mutations in either divergent or compensatory cases could possibly generate a cascade effect of misregulation of targets that might have not experienced cis-regulatory mutations between divergent taxa (Figure 2A–red arrows). We explored the idea of “transcriptional drive by sequence similarity among targets” by seeking sequence similarity within proximal (2,000 bp) putative cis-regulatory elements between transgressive genes showing evidence of cis–trans divergence or compensation and those showing no evidence of such sequence divergence. Our analysis showed some support for this idea with 18% of genes being possibly co-opted. However, only nine genes classified as non-compensatory appear as possible targets. One important limitation is that we only addressed sequence similarities between nearby upstream sequence regions of compensatory or cis–trans transgressive genes and upstream sequence regions of other transgressive genes, leaving unexplored the possibility that misregulation could be exerted by more distant cis-regulatory elements.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Drosophila bogotana SRA at GenBank: SRX091468. D. pseudoobscura extracted from FlyBase (http://flybase.org/). Expression data and other data can be accessed on Dryad (doi: 10.5061/dryad.2z34tmpjv).

AUTHOR CONTRIBUTIONS

AC designed the study. AG conducted the data analysis. Both authors contributed to writing and editing the manuscript.

FUNDING

This work was supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC-RGPIN-2017-04599) to AC and an NSERC Canada Graduate Scholarship (CGS M) to AG.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.599292/full#supplementary-material
REFERENCES

Andrews, S. (2010). FastQC A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed September 28, 2019).

Arterci, C. G., Haerty, W., and Singh, R. S. (2007). Association between levels of coding sequence divergence and gene misregulation in Drosophila male hybrids. J. Mol. Evol. 65, 697–704. doi: 10.1007/s00239-007-9048-2

Bader, D. M., Wilkening, S., Lin, G., Tekkedil, M. M., Dietrich, K., Steinmetz, L. M., et al. (2015). Negative feedback buffers effects of regulatory variants. Mol. Syst. Biol. 11:785. doi: 10.15252/msb20145844

Barreto, F. S., Pereira, R. J., and Burton, R. S. (2015). Hybrid dysfunction and physiological compensation in gene expression. J. Mol. Evol. 82, 613–622. doi: 10.1007/moblie/s00239-015-9833-9

Barreto, F. S., Watson, E. T., Lima, T. G., Edmands, S., Li, W., et al. (2018). Genomic signatures of mitonuclear coevolution across populations of Tityopus californicus. Nat. Evol. 2, 1250–1257. doi: 10.1038/nmeval.4106

Blankenberg, D., Von Kuster, G., Bouvier, E., Baker, D., Afgan, E., Stoler, N., et al. (2014). Dissemination of scientific software with Galaxy ToolShed. Genome Biol. 15:403. doi: 10.1186/gb4161

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bourguet, D. (1999). The evolution of dominance. Heredity 83(pt 1), 1–4. doi: 10.1038/sj.hdy.6858560

Brill, E., Kang, L., Michalak, K., Michalak, P., and Price, D. K. (2016). Hybrid sterility and evolution in Hawaiian Drosophila: differential gene and allel-specific expression analysis of backcross males. Heredity 117, 100–108. doi: 10.1038/hdy.2016.31

Castillo-Davis, C. I., Hartl, D. L., and Achaz, G. (2004). cis-Regulatory and protein selection and variability among functional categories. J. Mol. Evol. 29, 15–21.

Charif, D., and Lobry, J. R. (2007). "SeqinR 1.0-2: a contributed package to the oxfordjournals.molbev.a026334 journal.pone.0003009 Drosophila versus organ in Gomes, S., and Civetta, A. (2014). Misregulation of spermatogenesis genes in Drosophila hybrids is lineage-specific and driven by the combined effects of sterility and fast male regulatory divergence. J. Evol. Biol. 27, 1775–1783. doi: 10.1111/jeb.12428

Gomes, S., and Civetta, A. (2015). Hybrid male sterility and genome-wide misexpression of male reproductive proteases. Sci. Rep. 5:11976. doi: 10.1038/srep11976

Hyaphy, W., and Singh, R. S. (2006). Gene regulation divergence is a major contributor to the evolution of Dobhansky-Muller incompatibilities between species of Drosophila. Mol. Biol. Evol. 23, 1707–1714. doi: 10.1093/molbev/msi339

He, X., Zhou, S., St. Armour, G. E., Mackay, T. F. C., and Anholt, R. R. H. (2016). Epistatic partners of neurogenic genes modulate Drosophila olfactory behavior. Genes Brain Behav. 15, 280–290. doi: 10.1111/gbb.12279

Hodgins-Davis, A., Rice, D. P., Townsend, J. P., and Novembre, J. (2015). Gene expression evolves under a house-of-cards model of stabilizing selection. Mol. Biol. Evol. 32, 2130–2140. doi: 10.1093/molbev/msv094

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13. doi: 10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57. doi: 10.1038/nprot.2008.211

Hunt, B. G., Ometto, L., Keller, L., and Goodisman, M. A. D. (2013). Evolution at two levels in fire ants: the relationship between patterns of gene expression and protein sequence evolution. Mol. Biol. Evol. 30, 263–271. doi: 10.1093/molbev/msr234

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. doi: 10.1093/molbev/msr010

Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H., et al. (2005). Evolution: parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 309, 1850–1854. doi: 10.1126/science.1108296

King, M. C., and Wilson, A. C. (1975). Evolution at two levels in humans and chimpanzees. Science 188, 107–116. doi: 10.1126/science.1090005

Landry, C. R., Wittkopp, P. J., Taubes, C. H., Ranz, J. M., Clark, A. G., and Hartl, D. L. (2005). Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of Drosophila. Genetics 171, 1813–1822. doi: 10.1353/genetics.2005.047449

Lemos, B., MielkeJohn, C. D., Cáceres, M., and Hartl, D. L. (2005). Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. Evolution 59, 126–137. doi: 10.1111/j.0014-3820.2005.tb09090.x

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595. doi: 10.1093/bioinformatics/btp698

Liao, Y., Smyth, G. K., and Shi, W. (2014). FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. doi: 10.1093/bioinformatics/bti1656

Llopart, A. (2012). The rapid evolution of X-linked male-biased gene expression and the large-X effect in Drosophila yakuba, D. santomea, and their hybrids. Mol. Biol. Evol. 29, 3873–3886. doi: 10.1093/molbev/msr190

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550. doi: 10.1186/s13059-014-0550-8

Mack, K. L., Campbell, P., and Nachman, M. W. (2016). Gene regulation and speciation in house mice. Genome Res. 26, 451–461. doi: 10.1101/gr.195743.115

Mack, K. L., and Nachman, M. W. (2017). Gene regulation and speciation. Trends Genet. 33, 68–80. doi: 10.1016/j.tig.2016.11.003

Mackay, T. F. C. (2014). Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat. Rev. Genet. 15, 22–33. doi: 10.1038/nrg3627

Mackay, T. F. C., and Moore, J. H. (2014). Why epistasis is important for tackling complex human disease genetics. Trends Genet. 30, 59–64. doi: 10.1016/j.tig.2013.11.002

Mackay, T. F. C. (2014). Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat. Rev. Genet. 15, 22–33. doi: 10.1038/nrg3627

Mackay, T. F. C., and Moore, J. H. (2014). Why epistasis is important for tackling complex human disease genetics. Trends Genet. 30, 59–64. doi: 10.1016/j.tig.2013.11.002

Maheshwari, S., and Barbash, D. A. (2011). The genetics of hybrid incompatibilities. Annu. Rev. Genet. 45, 331–355. doi: 10.1146/annurev-genet-110410-132514
McManus, C. J., Coolon, J. D., Duff, M. O., Eipper-Mains, J., Graveley, B. R., Masly, J. P., Jones, C. D., Noor, M. A. F., Locke, J., and Orr, H. A. (2006). Go and Civetta Transgressive Gene Expression in Hybrids. *Genetics* 172, 89–99. doi: 10.1534/genetics.106.110644

Sundararajan, V., and Civetta, A. (2011). Male sex interspecies divergence and down regulation of expression of spermatogenesis genes in *Drosophila* sterile hybrids. *J. Mol. Evol. 72*, 80–89. doi: 10.1007/s00239-010-9404-5

Schurch, N. J., Schofeld, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., et al. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA 22*, 839–851. doi: 10.1016/j.rna.2015.07.002

Muller, H. J. (1942). Isolating mechanisms and evolution in temperature. *Biol. Synp. 6*, 71–125.

Nuzhdin, S. V., Wayne, M. L., Harmon, K. L., and McIntyre, L. M. (2004). Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. *Mol. Biol. Evol. 21*, 1308–1317. doi: 10.1093/molbev/msi128

Morán, T., and Fontdevila, A. (2014). Genome-wide distribution of hybrid sterility in *Drosophila* conforms a polygenic threshold architecture. *J. Hered. 105*, 381–396. doi: 10.1093/jhered/esu003

Ortiz-Barrientos, D., Counterman, B. A., and Noor, M. A. F. (2007). Gene expression divergence and the origin of hybrid dysfunctions. *Genetica* 129, 71–81. doi: 10.1007/s10709-006-0034-1

Phadnis, N., and Allen Orr, H. (2009). A single gene causes both male sterility and sege. *Science 323*, 376–379. doi: 10.1126/science.1163934.A

Presgraves, D. C. (2008). Sex chromosomes and speciation in *Drosophila*. *Trends Genet. 24*, 336–343. doi: 10.1016/j.tig.2008.04.007

Presgraves, D. C., Balagopalan, L., Abmayr, S. M., and Orr, H. A. (2003). Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature 423*, 715–719. doi: 10.1038/nature01679

Ranz, J. M., Namgyal, K., Gibson, G., and Hartl, D. L. (2004). Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaster* and *Drosophila simulans*. *Genome Res. 14*, 373–379. doi: 10.1101/gr.29804

Renaut, S., Nolte, A. W., and Bernatchez, L. (2009). Gene expression divergence and hybrid misexpression between lake whitefish species pairs (*Coregonus* spp.). *Genome Biol. Evol. 2*, 925–936. doi: 10.1093/molbev/msp017

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2009). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp166

Robinson, D. M., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol. 11*, R25. doi: 10.1186/gb-2010-11-3-125

Schaeffer, S. W., and Miller, E. L. (1991). Nucleotide sequence analysis of Adh genes estimates the time of geographic isolation of the Bogota population of *Drosophila pseudoobscura*. *Proc. Natl. Acad. Sci. U. S. A. 88*, 6097–6101. doi: 10.1073/pnas.88.14.6097

Sun, Y., and Johannesson, H. (2014). Dynamics of transcriptome evolution in the model eukaryote *Neurospora*. *J. Evol. Biol. 27*, 1125–1135. doi: 10.1111/jeb.12386

Zhang, X., and Emerson, J. J. (2019). Inferring compensatory evolution of cis- and trans-regulatory variation. *Trends Genet. 35*, 1–3. doi: 10.1016/j.tig.2018.11.003

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Go and Civetta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.