In: Speciation: Natural Processes, Genetics and Biodiversity ISBN: 978-1-62618-389-6
Editor: Pawel Michalak

Chapter 5

Transcriptional Differentiation Across the Four Subspecies of *Drosophila mojavensis*

Luciano M. Matzkin[1] *and Therese A. Markow*[2,3]
[1]University of Alabama in Huntsville,
Department of Biological Sciences, Alabama, US
[2]University of California San Diego, Section of Molecular and Cell Biology, California, US
[3]Center for Research and Advanced Studies, National Laboratory of Genomics for Biodiversity, University of California San Diego, California, US

Abstract

Local adaptation can play a fundamental role in the isolation of populations. While less well-studied than differentiation in sequence variation, changes in transcriptional variation during speciation also are fundamental to the evolutionary process. *Drosophila mojavensis* offers an unprecedented opportunity to examine the role of transcriptional differentiation in local adaptation. *Drosophila mojavensis* is a cactophilic fly composed of four ecologically distinct subspecies that inhabit the deserts of western North America. Each of the four subspecies utilizes necrotic tissue of different cactus host species characterized by distinct chemical profiles. The subspecies in Baja California, Mexico uses *Stenocereus gummosus* (Agria), in mainland Sonora it uses *S. thurberi* (Organ Pipe), in the Mojave Desert the host is *Ferocactus cylindraceus* (Red Barrel) and in Santa Catalina Island, USA, *Opuntia littoralis* (Prickly Pear) is the host. In this chapter we examine how the adaptation to the different environmental conditions across the four subspecies have shaped their transcriptional profiles. Using complete *D. mojavensis* genome microarrays we examined the transcriptome of third instar larvae from all four subspecies reared in standard laboratory media free of necrotic cactus-derived compounds. This experimental strategy focused on differences between constitutively expressed genes and not genes induced by necrotic cactus-derived compounds. The subspecies exhibited significant differential expression of genes that likely underlie the

adaptation to different cactus hosts, such as detoxification genes (Glutathione S-transferases, Cytochrome P450s and UDP-Glycosyltransferases) and chemosensory genes (Odorant Receptors, Gustatory Receptors and Odorant Binding Proteins).

Introduction

Increasing levels of genetic isolation between populations can lead to the formation of new species. The mechanisms involved could be broadly characterized as pre-zygotic or post-zygotic (Coyne and Orr 2004).

Although natural selection can play and active role in maintaining the isolation (e.g. hybrid inviability), natural selection does not necessary have to be implicated in the genetic divergence of the populations (Dobzhansky-Muller incompatibility model) (Dobzhansky 1937; Muller 1942).

In certain cases local adaptation can amplify the divergence across populations and hence accelerate the speciation process (Funk, Nosil, and Etges 2006; Nosil 2007; Schluter and Conte 2009; Nosil 2012).

Among many other characters, the pattern of variation of the transcriptome can be shaped by the adaptation to local ecological conditions. Knowledge of the regulatory differentiation present in ecologically distinct populations therefore informs our understanding the role of transcriptional changes in speciation.

Cactophilic *Drosophila* offer a powerful system to assess how a combination of geographic isolation, local adaptation and in some cases sexual selection can play a critical role in speciation.

One such example is that of the North American endemic cactophilic species, *D. mojavensis*. Similar to all cactophiles, *D. mojavensis* feeds, oviposits and develops in the necrotic tissues of certain cactus species (Fellows and Heed 1972; Heed 1978; Heed 1982; Ruiz, Heed, and Wasserman 1990). In general all *Drosophila* are saprophytic, and just like others, *D. mojavensis* feeds on the several yeast species that are known to inhabit the cactus necroses (Starmer 1982b; Starmer 1982a; Fogleman and Starmer 1985; Starmer et al. 1990).

In the process of consuming yeasts the flies also ingest and are exposed to the tissues and chemical composition of the cactus host (Kircher 1982). *Drosophila mojavensis* has relatively recently (<0.5 million years ago) diverged from its cactophilic sister species, *D. arizonae* (Heed 1978; Machado et al. 2007; Matzkin 2008; Smith et al. 2012). Currently, *D. mojavensis* is composed of four geographically (Figure 15) and ecologically distinct subspecies (Pfeiler, Castrezana, and Reed 2009).

Each subspecies utilizes a distinct cactus host, each characterized by distinct microflora and chemical profiles. Two of the subspecies, Baja California and mainland Sonora (hereafter Sonora), utilize columnar cacti belonging to the same genus, *Stenocereus* (*S. gummosus* and *S. thurberi*, respectively). The subspecies in the Mojave Desert utilizes Red Barrel (*Ferocactus cylindraceus*), while the subspecies in Santa Catalina Island utilized Prickly Pear (*Opuntia littoralis*).

Given the toxic nature of some of the compounds found in the cactus necroses that are inhabited by *D. mojavensis*, much of the earlier transcriptional work has focused on the dietary induction of cactus-related compounds (Matzkin et al. 2006; Matzkin 2012). This

approach has helped identify genes whose expression and pattern of sequence variation appear to have been shaped by local adaptation (Matzkin 2008).

In this chapter our goal was to remove the possible confounding effects of the induction of genes in response to cactus-derived compounds, and thus focus on transcriptionally fixed differences across the subspecies.

Using *D. mojavensis* specific microarrays, we investigated the transcriptional profile of third instar larvae (reared in cactus-free media) from 22 isofemale lines representative of all four subspecies. We observed that genes associated with metabolism are a major component of the transcriptional differences across the host subspecies.

Surprisingly, several chemosensory genes were among those who exhibited significant within subspecies variation, although many of these also exhibited between subspecies differences.

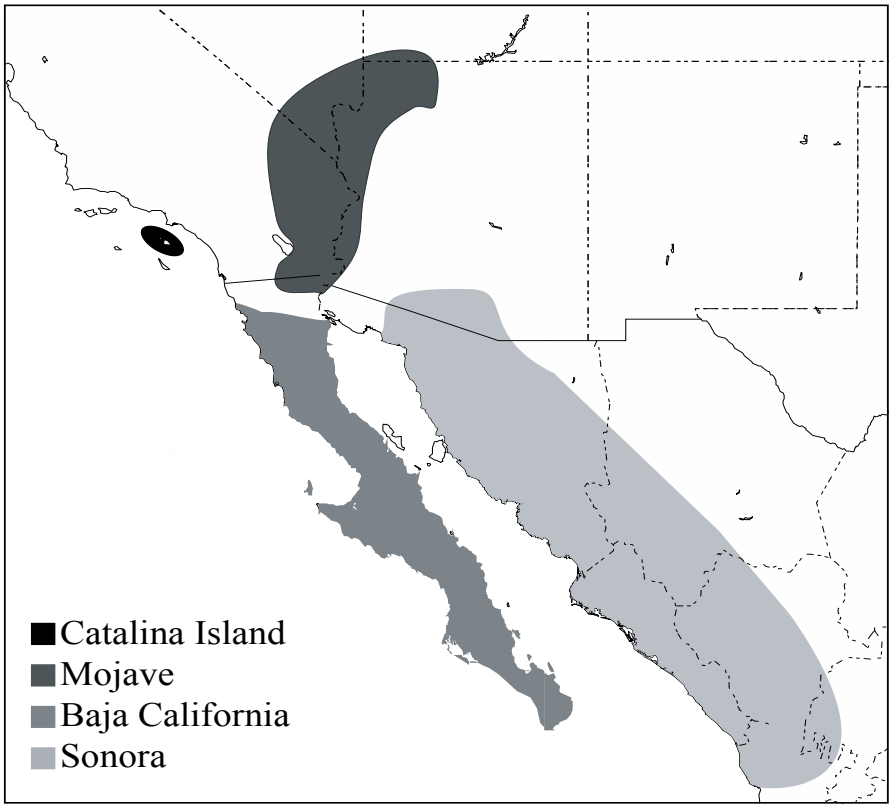

Figure 15. Distribution of the four *D. mojavensis* subspecies.

Analysis of Transcriptional Divergence

A total of 22 isofemale lines were used in this chapter, six each for the Catalina, Sonora and Mojave subspecies and four from Baja California (Table 4). Isofemale lines were reared in banana-molasses media in vials.

A generation prior to the experiment five females and five males were placed in a banana-molasses 8-dram vial with granules of live yeast for 24 hours. After this period the ten adults were removed and placed in new vials for next 24 hours. A total of five of these replicate vials were established per each of the 22 isofemale lines. Approximately eight days after oviposition third instar wandering larvae were collected. RNA was extracted for a total of two replicates per each of the 22 isofemale lines.

All samples were hybridized onto a custom *D. mojavensis* microarray based on the previously sequenced *D. mojavensis* genome (*Drosophila* 12 Genomes Consortium 2007). As originally described in Bono *et al.* (2011) the array consists of 71,998 60 oligonucleotide probes representing 14,519 annotated *D. mojavensis* genes. The large majority (96%) of the genes in the microarray were represented by 6 probes each. Expression intensities were first normalized using the Robust Multichip Average (RMA) method (Bolstad et al. 2003; Irizarry et al. 2003). Statistical analysis of the $\log_2$ transformed data was performed using a two-step mixed model ANOVA (Wolfinger et al. 2001).

The first step of this method is a global (data set wide) analysis that removes probe- and hybridization-specific effects, while the second step is a gene-specific analysis with subspecies and lines-within-subspecies as factors (also including microarray as a random factor). Given the 14,519 tests performed, significance was determined using the False Discovery Rate (FDR) method (Storey and Tibshirani 2003).

Table 4. Collection sites and cactus host for lines used

Line	Collecting site	Subspecies	Cactus host
ANZA-0402-3 ANZA-0402-4 ANZA-0402-5 ANZA-0402-8 ANZA-0402-10 ANZA-0402-17	Anza-Borrego Desert (CA, USA)	Mojave	Red Barrel (*F. cylindraceus*)
CI-1002-3 CI-1002-6 CI-1002-8 CI-1002-9 CI-1002-23 CI-1002-27	Santa Catalina Island Conservancy (CA, USA)	Catalina Island	Prickly Pear (*O. littoralis*)
MJBC-35 MJBC-103 MJBC-155 MJBC-216	La Paz (BC, Mexico)	Baja California	Agria (*S. gummosus*)
OPNM-407-2 OPNM-407-3 OPNM-407-5 OPNM-407-6 OPNM-407-8 OPNM-407-10	Organ Pipe National Monument (AZ, USA)	Mainland Sonora	Organ Pipe (*S. thurberi*)

Two-way hierarchical clustering using the Ward method was performed for the differentially expressed genes. Clustering was first performed on mean expression intensity for each isofemale line and then by isofemale line.

Analysis of the overrepresentation gene ontology (GO) terms was determined using the gene ontology enrichment analysis and visualization tool, GOrilla (Eden et al. 2007; Eden et al. 2009). Annotations of the *D. mojavensis* genome were based on the most current FlyBase annotation set (version FB2012-02) and those described in Matzkin (2012). Of the 14,519 *D. mojavensis* genes in the microarray we were able to obtain orthologous calls to *D. melanogaster* for just 10,685 genes. This smaller set of genes was used as the background gene set to test for overrepresentation of GO terms. For the purpose of examining overrepresented GO terms we set the *P*-value at less than 1×10^{-4}.

The entire transcriptional profile for each isofemale line can be observed in Figure 16. All expression data have been placed in the Gene Expression Omnibus under series entry #GSE41155.

Although there are three exceptions (MJBC103, CI27 and ANZA10) all the isofemale lines tend to group together according to host subspecies (Figure 16). Genes with significant expression differences between and within subspecies are shown in Table 5. Regardless of the FDR cutoff used, the number of genes that exhibited a significant subspecies effect was roughly more than twice that showing significant within subspecies variation. We were interested in examining the most robust of differences between the subspecies and within subspecies and hence chose to use an FDR cutoff of 0.001 to assign significance. Even with this conservative FDR level, a total of 3,092 genes exhibited a significant subspecies effect, of which 655 also had a significant within subspecies effect (Figure 17).

Among the genes with significant between subspecies differences include those involved in detoxification (Glutathione S-transferases, Cytochrome P450s, and UDP-Glycosyltransferases) and chemosensory (Odorant Receptors, Gustatory Receptors and Odorant Binding Proteins) (Table 6).

For 660 genes we observed only a within subspecies effect (Figure 17). To examine subspecies pair differences, we performed a *post-hoc* test setting the FDR at 0.1% (Table 7). Comparisons using the mainland Sonora subspecies have the greatest number of observed transcriptional differences, while the comparison between Catalina Island and Baja California had the fewest (Table 7).

Of the 3,092 genes with significant between subspecies differences, 2,033 had and orthologous calls to *D. melanogaster*. Results of the analysis of overrepresented Molecular Function and Biological Process GO terms are shown Figure 18. Most of the overrepresented GO terms appear to be associated with some aspect of metabolism.

Also overrepresented were members of important detoxification gene families such as Glutathione S-transferases and Cytochrome P450s (within the heme-binding Biological Process GO terms). On the other hand, GO terms of genes that exhibited significant within subspecies variation (1,315 of which 744 had orthologous calls to *D. melanogaster*) were largely associated with chemosensory perception (Figure 19).

The chromosomal location of the significant genes (Table 8) is significantly different ($\chi^2 = 14.2$, df = 4, $P = 0.0068$) from the expectation given the distribution of all genes in the *D. mojavensis* genome.

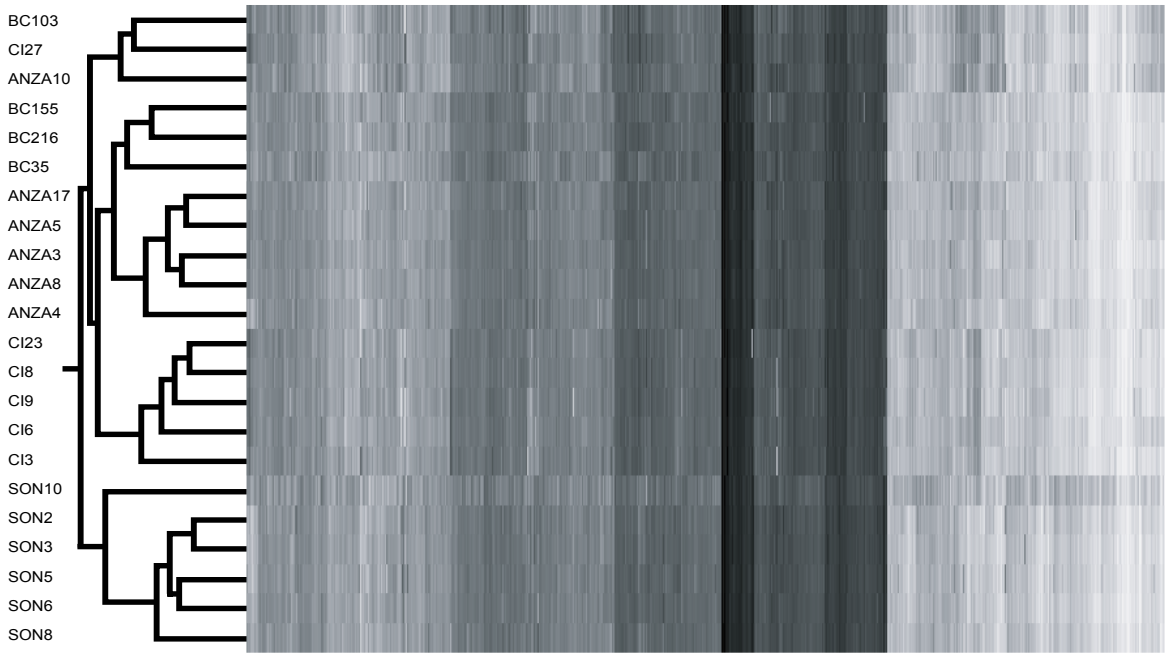

Figure 16. Two-level hierarchical clustering of expression intensities for the 14,519 genes examined. Expression intensities range from black (highest expression) to white (lowest expression). The tree indicates the transcriptional clustering across the 22 isofemale lines studied.

Table 5. Number of significant genes with significant subspecies and within subspecies variation [Line(Subspecies)] using different FDR cutoff values

	FDR		
	0.05	0.01	0.001
Subspecies	6343	4705	3092
Line(Subspecies)	2788	1971	1315

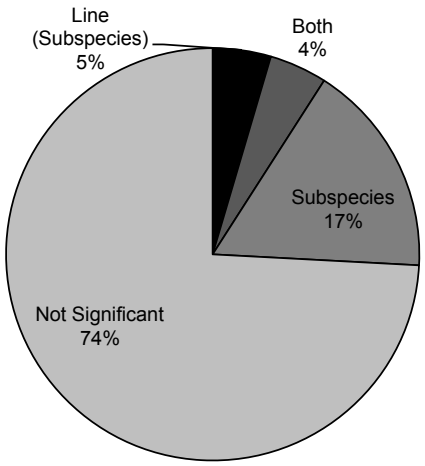

Figure 17. Proportion of genes with a significant within subspecies effect [Line(Subspecies)], subspecies effect, both within and between subspecies effects and not significant using and FDR cutoff value of 0.001.

Table 6. Genes with significant between subspecies differences belonging to Gustatory Receptor (GR), Odorant Receptor (OR), Odorant Binding Protein (OBP), Glutathione S-Transferase (GST), Cytochrome P450 (P450) and UDP-Glycosyltransferase (UGT) gene families

D. moj Symbol	*D. moj* name	*D. mel* Ortholog	Gene Family
Dmoj\GI11424		Gr61a	GR
Dmoj\GI11499		Gr43b	GR
Dmoj\GI12801		Gr63a	GR
Dmoj\GI15700		Gr2a	GR
Dmoj\GI17782		Gr28a	GR
Dmoj\GI20824		Gr43a	GR
Dmoj\GI22003		Gr94a	GR
Dmoj\GI23037		Gr98b	GR
Dmoj\GI11973		CG6776	GST
Dmoj\GI11974		CG6781	GST
Dmoj\GI14608		CG1681	GST
Dmoj\GI16623		GstE2	GST
Dmoj\GI19388		GstE10	GST
Dmoj\GI19515		CG16936	GST
Dmoj\GI20072		CG4688	GST
Dmoj\GI20122		GstE6	GST
Dmoj\GI20123	Dmoj\GstE6b		GST
Dmoj\GI20124		GstE9	GST
Dmoj\GI22354	Dmoj\GstD1d	GstD1d	GST
Dmoj\GI22356		GstD10	GST
Dmoj\GI23193	Dmoj\GstD1b		GST
Dmoj\GI23194	Dmoj\GstD1e		GST

Table 6. (Continued)

D. moj Symbol	*D. moj* name	*D. mel* Ortholog	Gene Family
Dmoj\GI23195		GstD2	GST
Dmoj\GI23196		CG17639	GST
Dmoj\GI23596		CG9363	GST
Dmoj\GI17488	Dmoj\Obp28a	Pbprp5	OBP
Dmoj\GI15510		Obp8a	OBP
Dmoj\GI19754		Obp47a	OBP
Dmoj\GI19915		Obp49a	OBP
Dmoj\GI20270		Obp44a	OBP
Dmoj\GI21087		Obp56h	OBP
Dmoj\GI23726		Obp93a	OBP
Dmoj\GI14836		Or1a	OR
Dmoj\GI16944		Or13a	OR
Dmoj\GI17592	Dmoj\Or67a-1		OR
Dmoj\GI17593	Dmoj\Or67a-2		OR
Dmoj\GI19019		Or59b	OR
Dmoj\GI19311		Or49b	OR
Dmoj\GI19887		Or45b	OR
Dmoj\GI23263	Dmoj\Or85a-1		OR
Dmoj\GI23327	Dmoj\OrN2-2		OR
Dmoj\GI23643		Orco	OR
Dmoj\GI23646		Or83a	OR
Dmoj\GI23916		Or85d	OR
Dmoj\GI24760		Or33c	OR
Dmoj\GI10234		Cyp313b1	P450
Dmoj\GI11220		Cyp318a1	P450
Dmoj\GI12456		Cyp4d8	P450
Dmoj\GI12535		Cyp4d20	P450
Dmoj\GI13002		Cyp12d1-d	P450
Dmoj\GI15489		Cyp6v1	P450
Dmoj\GI16990		Cyp309a1	P450
Dmoj\GI17558		Cyp28a5	P450
Dmoj\GI18694		Cyp4e2	P450
Dmoj\GI18702		Cyp6a17	P450
Dmoj\GI18705		Cyp317a1	P450
Dmoj\GI20052			P450
Dmoj\GI20196		Cyp49a1	P450
Dmoj\GI20222		Cyp9h1	P450
Dmoj\GI20230		Cyp6a18	P450
Dmoj\GI20372		Cyp4p1	P450
Dmoj\GI20893	Dmoj\Cyp6a21b		P450
Dmoj\GI20894		Cyp6a21	P450
Dmoj\GI21924		Cyp4ac1	P450
Dmoj\GI22127		Cyp4c3	P450
Dmoj\GI23350		Cyp304a1	P450
Dmoj\GI24047		Cyp12e1	P450
D. moj Symbol	*D. moj* name	*D. mel* Ortholog	Gene Family
Dmoj\GI24725		Cyp9f2	P450
Dmoj\GI10120		Ugt86Da	UGT
Dmoj\GI17057		Ugt37c1	UGT
Dmoj\GI17058		Ugt36Bb	UGT
Dmoj\GI17522		Ugt37a1	UGT
Dmoj\GI17523		Ugt37b1	UGT
Dmoj\GI19212		CG15661	UGT
Dmoj\GI19214		CG4302	UGT
Dmoj\GI22626		Ugt86Dj	UGT
Dmoj\GI22627		Ugt35a	UGT
Dmoj\GI22630		Ugt86Dd	UGT

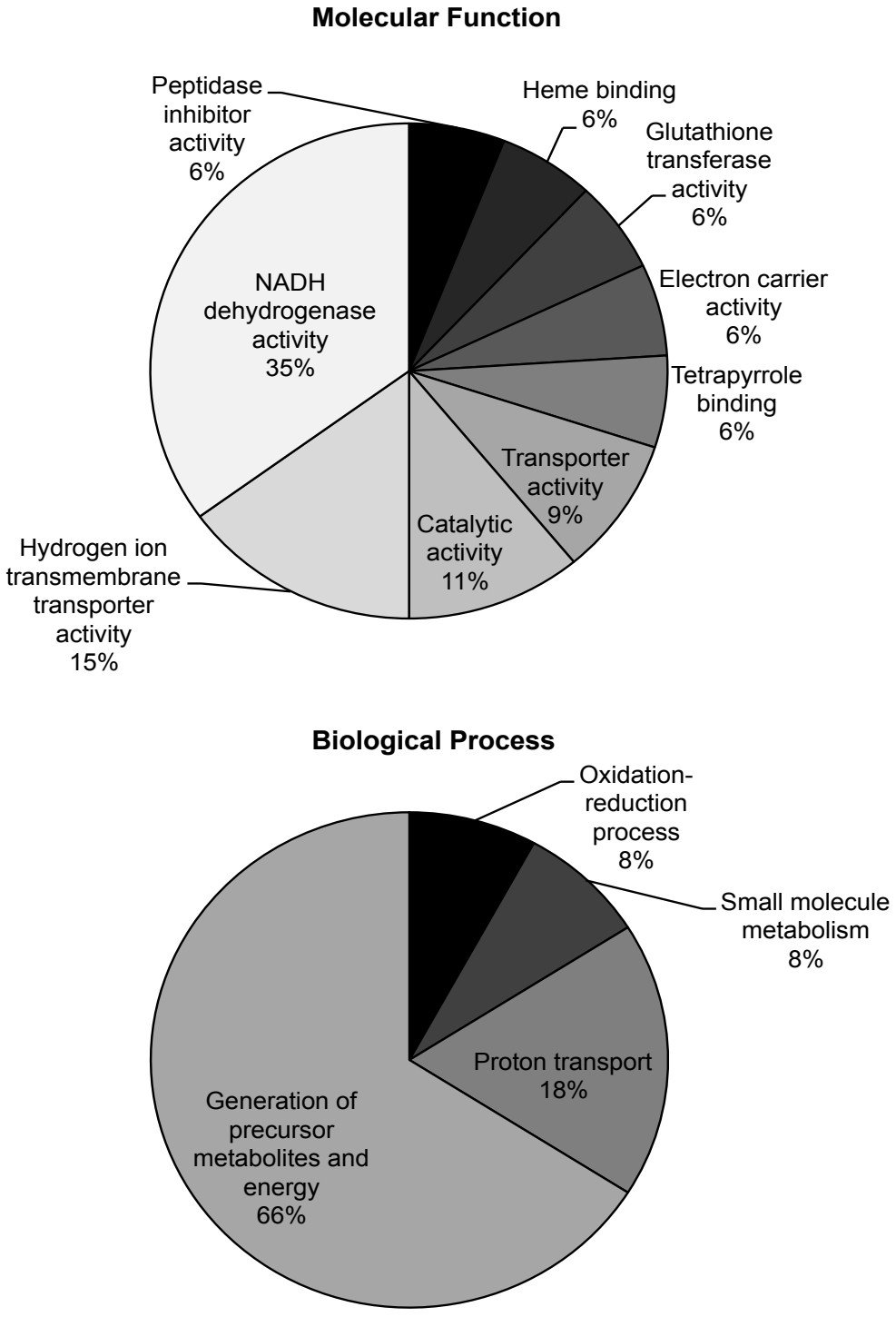

Figure 18. Composition of the overrepresented Molecular Function and Biological Process GO terms for those genes who exhibited a significant between subspecies effect.

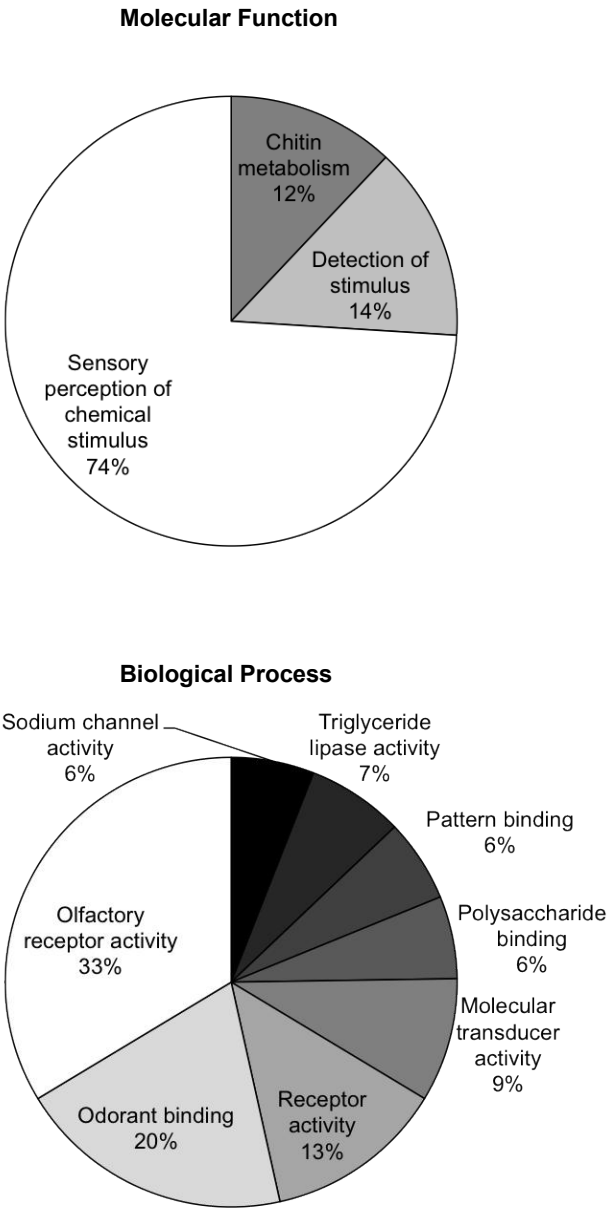

Figure 19. Composition of the overrepresented Molecular Function and Biological Process GO terms for those genes who exhibited a significant within subspecies effect.

Table 7. Number of significant transcriptional differences (FDR = 0.1%) between pairs of the *D. mojavensis* subspecies

	Catalina Island	Mojave	Baja California	Sonora
Catalina Island	-			
Mojave	861	-		
Baja California	396	580	-	
Sonora	2,493	1,263	837	-

Table 8. Chromosomal location of the significant differentially expressed between subspecies

Chromosome	Muller Element	Observed	Expected [1]
1	A	434	485.6
2	E	765	702.0
3	B	567	537.9
4	D	540	561.6
5	C	555	573.9

[1] Based upon the total number of genes in the respective chromosomes in the *D. mojavensis* genome.

Ecological Divergence in the *D. mojavensis* Transcriptome

The four host subspecies of *D. mojavensis* are presented by distinct ecological conditions, both biotic and abiotic in nature. These factors, among others (*e.g.* geographic isolation and sexual selection), have contributed to the morphological/structural (Etges et al. 2009; Pfeiler, Castrezana, and Reed 2009), life history (Etges and Heed 1987; Etges 1990), behavioral (Krebs and Markow 1989; Markow 1991), molecular (Matzkin 2004; Machado et al. 2007; Matzkin 2008; Smith et al. 2012), biochemical (Matzkin 2005) and transcriptional (Matzkin et al. 2006; Matzkin 2012) variation in this species. Here we have identified the inherent modifications in the transcriptional profiles of the four *D. mojavensis* subspecies, independent of their diet.

Transcriptional regulatory differences accumulate as a function of isolation, having arisen by *cis* changes followed by *cis* and *trans* coevolution (Wittkopp, Haerum, and Clark 2004; Gordon and Ruvinsky 2012). As is clear from the clustering of the four subspecies by the entire transcriptome (Fig. 15), these subspecies have undergone a significant degree of transcriptional evolution. While we cannot specify the particular evolutionary force(s) that shaped the observed transcriptional divergence, the geographic isolation between the subspecies and historical bottlenecks (Smith et al. 2012) suggest that genetic drift could have influenced the transcriptional divergence (Stone and Wray 2001; Wray et al. 2003; Lynch, Scofield, and Hong 2005). At the same time, the known ecological differences between the subspecies ultimately can point to those genes whose expression difference may have been shaped by local adaptation. It is this adaptation to local conditions, partly via transcriptional changes, that contributes to the reduced success of migrant individuals, therefore aiding in the genetic isolation of the subspecies.

The necrotic cactus habitats utilized by each of the four host subspecies have marked chemical differences (Kircher 1982). In addition to the chemical composition of the different cactus species, the microfloral communities are critical in creating the chemical profile of what larvae and adult flies ingest (Starmer 1982b; Starmer 1982a; Fogleman and Starmer 1985; Starmer et al. 1990). Among these chemical differences are nutritional compounds, such as carbohydrates and lipids (Vacek 1979; Kircher 1982; Fogleman and Abril 1990). Of the 3,092 overall transcriptional differences across the subspecies, the significant majority had some role in metabolism (Fig. 17). Observed differences in metabolism-related genes

observed thus may directly reflect adaptation to the different nutritional compositions of necrotic cactus hosts. As well as differences in nutritional compounds, several compounds present in cactus necroses can be toxic. Previous studies of transcriptional changes induced in response to necrotic cactus compounds revealed that many detoxification genes were affected (Matzkin et al. 2006; Matzkin 2012). In the present study we also observed expression differences in detoxification genes such as Glutathione S-transferases, Cytochrome P450 and UDP-Glycosyltransferases. Unlike in the prior studies, however, these differences are fixed and do not involve cactus compounds for their induction. Of course it is feasible that some of the significant gene expression differences across the subspecies could have been a result of a subspecies-specific response to the banana media. We reason that given the composition of the banana media (banana, yeast, molasses, corn syrup, antifungal and agar), many detoxification genes would not be influenced. Given that all the differences observed in the present study were seen in the absence of cactus compounds, many of the metabolism and detoxification genes underlying local adaptation appear to have been canalized with respect to their transcriptional profile.

Transcriptional Divergence and Speciation

The number of transcriptional differences differs drastically depending upon the particular subspecies analyzed (Table 7). Currently, several molecular studies (Machado et al. 2007; Matzkin 2008; Smith et al. 2012) support the earlier idea by Ruiz *et al.* (1990) that the center of genetic diversity for *D. mojavensis* resides in Baja California. *Drosophila mojavensis* colonized mainland Sonora, Mojave and Catalina Island following its allopatric divergence from its sister species, *D. arizonae,* in Baja California. This historical model of the origin of the *D. mojavensis* subspecies predicts that, in the absence of selection, the fewest number of fixed transcriptional differences should involve comparisons involving Baja California and the other three subspecies, while the largest difference should occur between non-Baja California comparisons. Our observed pair-wise transcriptional differences (Table 7) support this prediction. Interestingly we observed that comparisons using the mainland Sonora subspecies exhibited the largest number of transcriptional differences. One major difference between the Sonora subspecies and the other three is the fact that it is the only one that is sympatric with its sister species *D. arizonae* (Fellows and Heed 1972; Ruiz, Heed, and Wasserman 1990). Reinforcement has markedly shaped the behavior of the sympatric Sonoran subspecies (Markow 1981) and this in turn could have potentially shaped its transcriptome. Given the expected multifaceted nature of reinforcement, it would be expected that several genomic regions would be affected, which would include transcriptional differences.

Chromosomal inversions have been proposed to play a role in speciation (Noor et al. 2001; Rieseberg 2001; Machado, Haselkorn, and Noor 2007). Inversions facilitate the persistence of linkage blocks that contain sterility factors and as a result higher levels of divergence should be expected around the inverted regions. We observed a greater number of expression differences in chromosomes 2 and 3 and lower number in chromosomes 1, 4 and 5 than expected by chance (Table 8). In *D. mojavensis* only chromosomes 2 and 3 are polymorphic and several of these inversions are fixed between the subspecies (Wasserman

1962; Mettler 1963; Johnson 1980; Ruiz, Heed, and Wasserman 1990). According to the model, elevated levels of fixed sequence differences should be observed around the inversion breakpoints, it is probable that these sequence differences could also affect transcriptional levels as it was observed here. Although we have not yet mapped the breakpoint sequences of the inversions in *D. mojavensis*, the increased level of transcriptional divergence in chromosomes 2 and 3 as a whole, provides us with some initial tantalizing evidence on the role of inversions in local adaptation and isolation in these subspecies. Current genome sequencing of the three remaining *D. mojavensis* subspecies by Matzkin will provide the information needed to begin to answer this question.

Although the largest number of transcriptional differences observed were between subspecies, approximately 9% the transcriptome exhibited significant within subspecies variation of which roughly half also differed between subspecies (Fig. 16). A disproportionate large number were involved with chemosensory pathways mostly dealing with odorant behavior (Fig. 18). Thirty Odorant Receptors had significant within subspecies variation, 11 of which also had significant between subspecies differences. Significant within subspecies variation could possibly be a result of relaxation of selection or balancing selection. This would suggest that at least directional selection has not played a role in shaping the pattern transcriptional variation in these chemosensory genes. Alternatively, the significant within subspecies variation could be result of the lack of stimuli (*i.e.* cactus derived compound) present in the rearing medium (banana-molasses food). Several chemosensory genes have been previously shown to respond to cactus rearing (Matzkin et al. 2006; Matzkin 2012). Given variation in host chemical composition, it is possible that for certain genes it would be advantageous to have a more plastic transcriptional profile. Additional studies comparing cactus vs. non-cactus rearing are necessary to determine roles, if any, of chemosensory genes in the local adaptation of the different host subspecies.

Conclusion

Adaptation to local ecological conditions can be a potent driver of divergence between isolated populations (Funk, Nosil, and Etges 2006; Nosil 2007; Schluter and Conte 2009; Nosil 2012). *Drosophila mojavensis* offers a powerful system to investigate the role of local adaptation in the process of speciation. These subspecies already have accumulated a series of behavioral, morphological, life history and molecular differences (Etges and Heed 1987; Markow 1991; Matzkin 2005; Matzkin et al. 2006; Machado et al. 2007; Etges et al. 2009; Smith et al. 2012). Here we presented evidence that local adaptation might have been responsible in shaping the transcriptome of these ecologically different cactus host subspecies. Studies aiming to link genome level differences with the observed transcriptional differences as well as to functional, physiological and behavior differences are ongoing in our laboratories.

Acknowledgments

We thank Georgina J. Matzkin (a NSF Research Experience for Teachers intern) and Jasmin Uribe for their assistance in this project. We also would like to thank the UCSD Drosophila Stock Center. This project was supported by a grant from the Science Foundation Arizona to LMM and TAM and a NSF award (DEB-1219387) to LMM.

References

Bolstad, B. M., R. A. Irizarry, M. Astrand, and T. P. Speed. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185-193.

Bono, J. M., L. M. Matzkin, E. S. Kelleher, and T. A. Markow. 2011. Postmating transcriptional changes in reproductive tracts of con- and heterospecifically mated Drosophila mojavensis females. *Proceedings of the National Academy of Sciences* 108:7878-7883.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Inc., Sunderland.

Dobzhansky, T. 1937. *Genetics and the Origin of Species.* Columbia University, New York.

Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* 450:203-218.

Eden, E., D. Lipson, S. Yogev, and Z. Yakhini. 2007. Discovering Motifs in Ranked Lists of DNA Sequences. *PLoS Computational Biology* 3:e39.

Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48.

Etges, W. J. 1990. Direction of life history evolution in Drosophila mojavensis. Pp. 37-56 *in* J. S. F. Barker, W. T. Starmer, and R. J. MacIntyre, eds. *Ecological and evolutionary genetics of Drosophila.* Plenum Press, New York.

Etges, W. J., C. C. De Oliveira, M. G. Ritchie, and M. A. F. Noor. 2009. Genetics of Incipient Speciation in Drosophila mojavensis: II. Host Plants and Mating Status Influence Cuticular Hydrocarbon QTL Expression and GxE Interactions. *Evolution* 63:1712-1730.

Etges, W. J., and W. B. Heed. 1987. Sensitivity to larval density in populations of Drosophila mojavensis: Influences of host plant variation on components fitness. *Oecologia* 71:375-381.

Fellows, D. F., and W. B. Heed. 1972. Factors affecting host plant selection in desert-adapted cactiphilic Drosophila. *Ecology* 53:850-858.

Fogleman, J. C., and J. R. Abril. 1990. Ecological and evolutionary importance of host plant chemistry. Pp. 121-143 *in* J. S. F. Barker, W. T. Starmer, and R. J. MacIntyre, eds. *Ecological and evolutionary genetics of Drosophila*. Plenum Press, New York.

Fogleman, J. C., and W. T. Starmer. 1985. Analysis of the community structure of yeasts associated with the decaying stems of cactus. III. Stenocereus thurberi. *Microbial Ecology* 11:165-173.

Funk, D., P. Nosil, and W. Etges. 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Science* 103:3209-3213.

Gordon, K. L., and I. Ruvinsky. 2012. Tempo and Mode in Evolution of Transcriptional Regulation. *Plos Genetics* 8:e1002432.

Heed, W. B. 1982. The origin of Drosophila in the sonoran desert. Pp. 65-80 *in* J. S. F. Barker, and W. T. Starmer, eds. *Ecological genetics and evolution: The cactus-yeast-Drosophila model system.* Academic Press, New York.

Heed, W. B. 1978. Ecology and genetics of Sonoran desert *Drosophila*. Pp. 109-126 *in* P. F. Brussard, ed. *Ecological genetics:* The interface. Springer-Verlag.

Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat* 4:249-264.

Johnson, W. R. 1980. *Chromosomal polymorphism in natural populations of the desert adapted species, Drosophila mojavensis*. Pp. 75. Department of Ecology and Evolutionary Biology. University of Arizona, Tucson, AZ.

Kircher, H. W. 1982. Chemical composition of cacti and its relationship to Sonoran Desert Drosophila. Pp. 143-158 in J. S. F. Barker, and W. T. Starmer, eds. *Ecological genetics and evolution: The cactus-yeast-Drosophila model system.* Academic Press, New York.

Krebs, R. A., and T. A. Markow. 1989. Courtship behavior and control of reproductive isolation in Drosophila mojavensis. *Evolution* 43:908-913.

Lynch, M., D. G. Scofield, and X. Hong. 2005. The Evolution of Transcription-Initiation Sites. *Molecular Biology and Evolution* 22:1137-1146.

Machado, C. A., T. S. Haselkorn, and M. A. F. Noor. 2007. Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in Drosophila pseudoobscura and D. persimilis. *Genetics* 175:1289-1306.

Machado, C. A., L. M. Matzkin, L. K. Reed, and T. A. Markow. 2007. Multilocus nuclear sequences reveal intra- and interspecific relationships among chromosomally polymorphic species of cactophilic Drosophila. *Molecular Ecology* 16:3009-3024.

Markow, T. A. 1981. Courtship behavior and control of reproductive isolation between Drosophila mojavensis and Drosophila arizonensis. *Evolution* 35:1022-1026.

Markow, T. A. 1991. Sexual isolation among populations of Drosophila mojavensis. *Evolution* 45:1525-1529.

Matzkin, L. M. 2008. The Molecular Basis of Host Adaptation in Cactophilic Drosophila: Molecular Evolution of a Glutathione S-Transferase Gene (GstD1) in Drosophila mojavensis. *Genetics* 178:1073-1083.

Matzkin, L. M. 2004. Population genetics and geographic variation of alcohol dehydrogenase (Adh) paralogs and glucose-6-phosphate dehydrogenase (G6pd) in Drosophila mojavensis. *Molecular Biology and Evolution* 21:276-285.

Matzkin, L. M. 2005. Activity variation in alcohol dehydrogenase paralogs is associated with adaptation to cactus host use in cactophilic Drosophila. *Molecular Ecology* 14:2223-2231.

Matzkin, L. M. 2012. Population transcriptomics of cactus host shifts in Drosophila mojavensis. *Molecular Ecology* 21:2428-2439.

Matzkin, L. M., T. D. Watts, B. G. Bitler, C. A. Machado, and T. A. Markow. 2006. Functional genomics of cactus host shifts in Drosophila mojavensis. *Molecular Ecology* 15:4635-4643.

Mettler, L. E. 1963. Drosophila mojavensis baja, a new form in the mulleri complex. *Drosophila Information Service* 38:57-58.

Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. *Biological Symposia* 6:71-125.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences* 98:12084-12088.

Nosil, P. 2012. *Ecological Speciation.* Oxford University Press, Oxford.

Nosil, P. 2007. Divergent Host Plant Adaptation and Reproductive Isolation between Ecotypes of Timema cristinae Walking Sticks. *The American Naturalist* 169:151-162.

Pfeiler, E., S. Castrezana, and L. K. Reed. 2009. Genetic, ecological and morphological differences among populations of the cactophilic Drosophila mojavensis from southwestern USA and northwestern Mexico, with descriptions of two new subspecies. *Journal of Natural History* 43:923-938.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution* 16:351-358.

Ruiz, A., W. B. Heed, and M. Wasserman. 1990. Evolution of the Mojavensis cluster of cactophilic Drosophila with descriptions of two new species. *Journal of Heredity* 81:30-42.

Schluter, D., and G. L. Conte. 2009. Genetics and ecological speciation. *Proceedings of the National Academy of Sciences* 106:9955-9962.

Smith, G., K. Lohse, W. J. Etges, and M. G. Ritchie. 2012. Model-based comparisons of phylogeographic scenarios resolve the intraspecific divergence of cactophilic Drosophila mojavensis. *Molecular Ecology* 21:3293-3307.

Starmer, W. T. 1982a. Analysis of the Community Structure of Yeasts Associated with the Decaying Stems of Cactus. I. Stenocereus gummosus. *Microbial Ecology* 8:71-81.

Starmer, W. T. 1982b. Associations and Interactions Among Yeasts, Drosophila and Their Habitats. Pp. 159-174 in J. S. F. Barker, and W. T. Starmer, eds. *Ecological genetics and evolution: The cactus-yeast-Drosophila model system.* Academic Press, New York.

Starmer, W. T., M. A. Lachance, H. J. Phaff, and W. B. Heed. 1990. The biogeography of yeasts Associated with decaying cactus tissue in North America, the Caribbean, and Northern Venezuela. Pp. 253-296 *in* M. K. Hecht, B. Wallace, and R. J. Macintyre, eds. *Evolutionary Biology.* Plenum Publishing Corporation.

Stone, J., and G. Wray. 2001. Rapid evolution of *cis*-regulatory sequences via local point mutations. *Molecular Biology and Evolution* 18:1764-1770.

Storey, J., and R. Tibshirani. 2003. Statistical significance for genomewide studies. *Proccedings of the National Academy of Sciences, USA* 100:9440-9445.

Vacek, D. C. 1979. The microbial ecology of the host plants of Drosophila mojavensis. Pp. 112. *Department of Ecology and Evolutionary Biology.* University of Arizona, Tucson, AZ.

Wasserman, M. 1962. *Cytological studies of the repleta group of the genus Drosophila.* V. The mulleri subgroup. University of Texas Publication 6205:85-118.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in cis and trans gene regulation. *Nature* 430:85-88.

Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules. 2001. Assessing gene significance from cDNA microarray expression data via mixed models. *Journal of Computational Biology* 8:625-637.

Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman, and L. A. Romano. 2003. The Evolution of Transcriptional Regulation in Eukaryotes. *Molecular Biology and Evolution* 20:1377-1419.