REVIEW

Genome clashes in hybrids: insights from gene expression

CR Landry[1], DL Hartl[1] and JM Ranz[2]

[1]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA and *[2]*Department of Genetics, University of Cambridge, Cambridge, UK*

In interspecific hybrids, novel phenotypes often emerge from the interaction of two divergent genomes. Interactions between the two transcriptional networks are assumed to contribute to these unpredicted new phenotypes by inducing novel patterns of gene expression. Here we provide a review of the recent literature on the accumulation of regulatory incompatibilities. We review specific examples of regulatory incompatibilities reported at particular loci as well as genome-scale surveys of gene expression in interspecific hybrids. Finally, we consider and preview novel technologies that could help decipher how divergent transcriptional networks interact in hybrids between species.

Heredity (2007) **99,** 483–493; doi:10.1038/sj.hdy.6801045; published online 8 August 2007

Introduction

Complex biological systems can exhibit multiple solutions for a particular evolutionary problem (Wagner, 2005). As a consequence, separate evolutionary lineages can accumulate genetic differences at orthologous genes yet maintain similar phenotypes. During this divergence, the molecular coevolution of genes ensures that their functions are maintained despite the accumulation of differences in regulatory and coding sequences (Dover and Flavell, 1984). This coevolution has an immense impact on the evolutionary process because the potential for change at the molecular level is much greater than at the phenotypic level.

Crosses between species or populations can reveal such coevolution among genes. When two species come together and form hybrids, alleles that have not previously occurred together may interact and produce novel phenotypes. Understanding the molecular bases of these newly generated non-additive, epistatic interactions is of paramount importance in evolutionary biology as it is central to the evolution of hybrid incompatibilities (the so-called Dobzhansky–Muller incompatibilities) (Dobzhansky, 1937; Muller, 1940), migrational load (Lenormand, 2002) and the appearance of key adaptive phenotypes observed in cases of hybrid speciation (Rieseberg *et al.*, 1996).

In this review, regulatory incompatibilities (RIs) are defined as interactions among elements of transcriptional networks that lead to novel expression-phenotypes in interspecific hybrids. We first discuss why RIs are

expected to accumulate between diverging species. Second, we critically review recent studies that reported large-scale patterns of gene expression in interspecific hybrids or for particular genes that are relevant to this model. We summarize lines of research that would improve our description and understanding of regulatory interactions. We also offer a brief survey of genomic techniques that open new avenues of research by enabling the measurements of the transcriptional activity of the two divergent genomes in interspecific hybrids.

The genetic bases of the evolution of gene-expression and transcription networks: an overview

The genetic bases of gene-expression variation have been reviewed elsewhere (Gibson and Weir, 2005). Here we limit ourselves to introduce some basic concepts that are pertinent to the accumulation of RIs. Regulation of gene expression is complex and intricate due to the large number of elements involved (proteins, RNA and DNA molecules, chromatin structure) as well as the number of interactions among them. These interactions are spatially and temporally dynamic such that genes display a diversity of expression patterns across the life cycle of an organism, across tissues and as a response to particular environmental cues. This complexity is epitomized by transcription initiation, the best-studied level of gene regulation (Box 1), in which a variety of transcription factors orchestrate the regulation of downstream genes that are required in the cell to accommodate changes in physiological conditions. These molecular inputs modulate the activity of the basal transcriptional machinery on the effector genes through transient interactions of the *trans*-acting factors with particular sequences of the gene regulatory elements

Correspondence. *Current address: Dr CR Landry, Département de Biochimie, Université de Montréal, Québec, Canada H3C 3J7.*
E-mail: clandry@post.harvard.edu or Dr JM Ranz, Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK.
E-mail: jmr68@mole.bio.cam.ac.uk

Box 1 Complexity of gene expression

Regulation of gene expression begins with transcription initiation and largely depends on a set of regulatory sequences, the basal promoter, various protein complexes, and epigenetic mechanisms such as DNA methylation and histone modification (Arnone and Davidson, 1997; Carroll *et al.*, 2001; Arnosti, 2003; Jaenisch and Bird, 2003; Levine and Tjian, 2003; Wray *et al.*, 2003; Davidson, 2006). Initiation of transcription is accomplished at the basal promoter located immediately upstream from the transcription unit. Promoters typically include consensus sequences (TATA sequence -TATA-, initiator sequence -INR- and/or downstream promoter element -DPE-), which are used as a docking point of the RNA polymerase II holoenzyme complex and general transcription factors. Tissue- and gene-specific transcription factors that interact with the basal machinery to fine tune expression level usually bind to scattered clusters of DNA-binding sites (5–15 bp long). These clusters are called regulatory modules or enhancers and are located in the vicinity of the transcription unit, usually upstream, but also downstream such as in introns. Importantly, each enhancer drives one particular spatial-temporal component of the overall expression profile of a gene. Enhancers, unlike basal promoters, are usually located at distances of few to several kilobases from the gene and exert their role in an orientation-independent manner, which occurs by facilitating or interfering with the binding of the general transcriptional machinery to the basal promoter and thus modifying the rate of transcription. The effect of enhancers is exerted mainly through remodeling of chromatin structure and through protein–protein interactions with the general transcriptional machinery or with other transcription factors; in most cases, chromatin remodeling/modifying complexes and cofactors are essential in the fine-tuning of the established interactions. A typical gene encompasses from 5 to 15 enhancers separated from each other by spacer regions. Enhancer sequences have a variable length but are at least 300 bp long and usually less than 10 kb in *Drosophila* and 100 kb in mammals. Many enhancers encompass 10–15 binding sites that can be occupied by 4–15 different transcription factors. The description above however corresponds to regulatory regions of developmental genes, which have been preferentially studied. Some genes also exist in which there are highly structured regulatory regions under strong organizational constraints (Thanos and Maniatis, 1995) and extensive regulatory regions with no obvious sign of clustering of transcription-binding sites (Klingler *et al.*, 1996).

Other mechanisms that regulate gene activity downstream of transcription are becoming more appreciated. Those mechanisms act at the post-transcriptional level through the processing, turnover, transport and localization of the transcripts and at the post-translational level through modifications of the encoded proteins (see Mata *et al.* (2005) for a thorough review). These levels of regulation can be important but are usually overlooked in interpreting, for example, microarray data. Most technologies measure the steady-state levels of mRNA, which are the result of both the rate of transcription and of RNA turnover. Thus, different events and pathways are involved in stabilizing the transcript, in eliminating faulting transcripts (Fasken and Corbett, 2005) and in adjusting the level of mRNA to the physiological needs of the organism (Khodursky and Bernstein, 2003). Many of these regulatory events involve the presence of specific sequences in the 3′UTR, which are bound by different ribosomal-binding proteins (Parker and Song, 2004). The relevance of the gene regulation at this level is underscored by the fact that approximately 50% of the affected transcripts in cultures of human cells under stress conditions became modified in abundance more as a result of their stability than because of altered transcription (Fan *et al.*, 2002). In mouse, proper brain development seems to be largely dependent on the translational control of the *Otx2* gene, which is modulated by a 140 bp sequence in the 3′UTR (Boyl *et al.*, 2001). Also, polymorphisms in the 3′UTR of the gene *Gpc3*, which correlates with small but reproducible changes in gene expression, account for most of the variation in body mass in mice (Oliver *et al.*, 2005). Finally, microRNAs, which recognize 3′UTR sequence motifs, have shown their remarkable importance in modulating the abundance of a significant fraction of the transcripts in eukaryotes either by direct cleavage of the transcript or by repressing their translation (Zamore and Haley, 2005).

In general, regulatory elements and molecules are classified as acting in *cis* or in *trans*, literally meaning on the same side or on the opposite side, respectively. *cis*-regulatory elements, such as transcription-factor-binding sites in enhancer sequences, are found on the same DNA molecule and in the vicinity of the gene being regulated and are therefore said to act in *cis*. Molecules that diffuse in the cell are said to act in *trans* because their activity is not specific to one copy of a gene: both alleles at one locus are equally likely to interact with a factor acting in *trans*. A transcription factor is thus an example of molecule acting in *trans*. In genetic terms, genetic variation for gene expression acting in *cis* would segregate with the gene, whereas genetic variation segregating independently from the gene would act in *trans*. It is however important to note that this distinction is an oversimplification as recent findings show than an enhancer of an odorant receptor gene is able to modulate the transcriptional activity of many other odorant receptor genes scattered across mouse chromosomes (Lomvardas *et al.*, 2006).

---

acting in *cis* (Box 1). Expression profiling experiments have uncovered changes in mRNA levels on a genomic scale, both within and between species (Ranz and Machado, 2006; Whitehead and Crawford, 2006). This variation in mRNA abundance is the result of changes that affect the genes in *cis*, including those associated with their regulatory sequences, changes in the abundance or activity of *trans*-acting factors acting upstream of the gene in the transcriptional network, or a combination of both (Gibson and Weir, 2005).

Transcriptional variation associated with *cis*- and *trans*-regulatory changes has consequences on organismal phenotypes and contributes to biological diversity in the wild. Evolution of *cis*-regulatory DNA sequences is particularly well studied, and it is known that changes in regulatory regions containing transcription-factor-binding sites have played a critical role in the morphological diversification of animal and plant lineages (Carroll, 2005a). Early studies included those on *Caenorhabditis* and vertebrates, in which it was noticed that interspecific differences in expression profiles can be the result of changes in the regulatory sequences upstream of particular genes critical during development (Belting *et al.*, 1998; Wang and Chamberlin, 2002). These changes in gene expression can be due to *de novo* evolution of regulatory sequences (Gonzalez *et al.*, 1995; Papaceit *et al.*, 2004; Wratten *et al.*, 2006), although a large number of studies indicate that their evolution often involves the rearrangement and modification of sequences with known regulatory roles (Gompel *et al.*, 2005; Prud'-homme *et al.*, 2006). In several cases, mutations in *cis*-regulatory sequences imply gains or losses of binding sites for key selector genes in a lineage-specific manner leading to complex changes in the patterns of gene expression. As a consequence, the gene's transcript abundance, the timing of expression and/or the spatial pattern of expression can be modified, consequently remodeling the architecture of transcriptional networks. The origin of the male wing spot in *Drosophila biarmipes* is a clear example of phenotypic change driven by *cis*-regulatory changes that affect transcriptional networks. This phenotypic feature is absent from its closest relative

and is the result of a *de novo* regulation of the gene *yellow* by the selector gene *engrailed*, among others (Gompel *et al.*, 2005). Similarly, male-specific abdomen pigmentation has occurred recurrently in the *D. melanogaster* species group, sometimes clearly associated with the gain of abdominal-B-binding sites also in the gene *yellow* (Jeong *et al.*, 2006). Another example of rapid phenotypic divergence driven by *cis*-regulatory mutations comes from vertebrates. In the stickleback fish, changes in the *cis*-regulatory region of the gene *paired-like homeodomain transcription factor1* (*pitx1*) are involved in the modification of its expression profile, which leads to the loss of pelvic armor in several populations (Shapiro *et al.*, 2004). The Ascomycete fungi provide further evidence (Gasch *et al.*, 2004). For instance, it has been suggested that the peculiar mode of cell division of budding yeast has led to the recruitment of genes involved in budding to come under the control of E2F-like transcription factors, which are involved in cell-cycle regulation in eukaryotes. Likewise, the regulation of genes associated with proteasome function in *Saccharomyces cerevisiae* and *Candida albicans* is controlled by the transcription factor Rpn4p, but these genes are controlled by another transcription factor in the fungus *Neurospora crassa*. Evolution of gene expression driven by changes in *trans*, such as those in transcription factors, is also important in generating phenotypic diversity. Thus, the evolution of transcriptional regulators encoded by some *Hox* genes has played a central role in the diversification of the body plan of modern insects (Galant and Carroll 2002). Finally, changes in gene expression can also lead to deleterious phenotypes. For instance, it has been shown that *cis*-regulatory variation associated with changes in gene expression predisposes to complex diseases in humans such as heart diseases (reviewed in Rockman and Wray, 2002; Wray 2007).

Some properties of *cis*-regulatory sequences may contribute to their remarkable ability to accommodate changes both within and between species (Dermitzakis and Clark, 2002; Rockman and Wray, 2002; Costas *et al.*, 2003). The low complexity of regulatory sequences is thought to enhance the rate at which rearrangements take place (McGregor *et al.*, 2001a; Costas *et al.*, 2003). Further, the large evolutionary potential of key nodes in gene-expression networks is enhanced by the functional autonomy of regulatory modules (Box 1), which limits detrimental pleiotropic effects (Carroll, 2005b), and by other important properties of the regulatory code (Wilkins, 2002): first, transcription factors usually interact with a collection of similar sequences rather than with only one; second, transcription-factor-binding sites often appear in multiple copies within the same regulatory region so that variation in copy number does not necessarily have large effects on the phenotype; and third, there are several features of the organization of enhancers that are independent of sequence motifs that also affect gene regulation, such as periodicity (Erives and Levine, 2004) or threshold distance (Chiang *et al.*, 2006) between adjacent transcription-factor-binding sites or enhancers. Accordingly, it could be foreseen that in humans the estimated heterozygosity at functional *cis*-regulatory sites surpasses the heterozygosity at amino-acid sites (Rockman and Wray, 2002).

In summary, the 'degeneracy' of the regulatory code provides ample opportunity for neutral variation to segregate in natural populations, in a manner analogous to the synonymous variation that is a consequence of the redundancy of the genetic code. A significant part of this regulatory variation however does lead to change in gene expression and thus to modifications of the interactions in the transcription network and in organismal phenotypes. This evolutionary lability might also account for cases of rapid divergence of phenotypic characters even among closely related species.

Molecular coevolution in transcriptional networks

Evolutionary changes among interacting genes must accommodate each other in order to maintain the functionality of the molecular interactions. Dover and Flavell (1984) first proposed that changes in transcription-factor-coding genes had to evolve to cope with changes in regulatory sequences, which are prone to rapid turnover. Coevolution between transcription factors and their binding sites is probably one of the simplest scenarios of regulatory coevolution that could lead to RIs in hybrids between species. For example, a change in the number of binding sites may be compensated for by a change in the activity or abundance of the transcription factors. Another possibility is provided by the fact that transcription factors often act within complexes of molecules and therefore coevolution could take place among them. Furthermore, compensatory changes do not need to involve multiple gene products but can accumulate among peptides or protein domains. One such case comes from a study of compensatory domain evolution in the transcription activator NifA of the bacterium *Sinorhizobium meliloti* (Juarez *et al.*, 2000). Compared to other homologs, this *trans*-activator protein interacts only weakly with its enhancers due to the lack of a glycine residue at a specific position in its DNA-binding domain. However, a functional analysis of the activation domain showed that the loss of affinity for the enhancers was compensated by amino-acid changes in the activator domain that enhances its activity. Coevolution within any one molecule is not likely to contribute to RIs in hybrids between species, but this study nicely illustrates how functional changes can accumulate among regulatory elements such that, despite molecular evolution, the phenotype remains conserved.

One important aspect is that this type of molecular coevolution often involves upstream *trans*-regulatory factors, which were once considered infrequent. This preconception was based on the assumption that, because transcription factors regulate tens of genes in the genome, their function must be highly constrained and thus regulatory evolution is more likely to accumulate at the level of individual downstream genes, rather than at these key nodes in the transcriptional networks (Tautz, 2000). This view is changing with the accumulation of comparative genomic studies showing that transcription factors can in fact evolve rapidly compared with other classes of genes. For instance, a comparative analysis of the rate of protein evolution across functional biological classes in different animal phyla (Castillo-Davis *et al.*, 2004) showed that genes coding for regulatory proteins (regulation of transcription) are among the fastest evolving genes between the closely

related nematode species *Caenorhabditis elegans* and *C. briggsae*. Furthermore, a thorough analysis of the effects of amino-acid substitutions on the activation potential of transcription factors has shown that the differential activation by mutant proteins can be highly dependent on promoter target sequence (Inga *et al.*, 2002), suggesting that transcription factors may coevolve with specific subsets of targets while maintaining proper interaction with others. Increasing evidence of the importance of the evolution of transcription factors is also accumulating in developmental biology (Hsia and McGinnis, 2003).

Molecular coevolution and its population dynamics

RIs can arise when independent mutations that affect gene expression become part of different gene pools, either by genetic drift or positive selection. Although these mutations might act as functionally equivalent within each gene pool, it need not follow that they will be functionally compatible when present in the same genetic background. Several population genetic models have been proposed to account for the accumulation of coevolved genes and loci within species. The goal here is not to review all models of molecular coevolution but to give a brief overview because the organization of transcriptional networks may affect how coevolution among genes, and therefore RIs, may accumulate in diverging lineages. For two genes or sites in a genome to coevolve, at least one substitution at each locus must fix along a given lineage. In the model proposed by Dover and Flavell (1984), and further developed by others (McGregor *et al.*, 2001b; Shaw *et al.*, 2002), the driving mutations occur in *cis*-regulatory-binding sites, which bring about deleterious changes in gene expression and thus create a selection pressure on genes that act in *trans* to coevolve changes that compensate for the decreased or increased binding affinity. In this model of compensatory evolution, mutations that are deleterious when present individually are neutral or nearly neutral when combined (Kimura, 1985). The conditions under which this scenario is plausible are rather restrictive because the first mutation, which is deleterious, has to fix in the population before the second mutation occurs, and this event is unlikely to occur in large populations due to the high efficiency of natural selection. Recent models of compensatory evolution (Carter and Wagner, 2002; Weinreich and Chao, 2005) that take into account different demographic scenarios have however shown that even in large populations certain recombination rates render fixation of the first mutation unnecessary, and thus compensatory coevolution may occur more often than initially expected.

The model envisioned by Dobzhansky (1937) and Muller (1940) is drastically different, as it does not require the action of natural selection other than stabilizing selection on the trait. In this case, the first substitution in a lineage is neutral in its original background but its effect is to change the genetic background for future substitutions that will be neutral on the new genetic background but deleterious in the ancestral one. This scenario appears *a priori* more likely because no decrease in fitness is necessary. However, its importance depends on the architecture of the transcrip-

tional network, that is whether mutations are so highly epistatic that single mutations in the genetic background can change the effect of future mutations. The examination of quantitative models of transcriptional regulation reveals that parameters affecting the level of transcription, such as the number of binding sites upstream of a gene, the abundance of transcription factors and their affinity for the binding sites, and the cooperativity leading to non-linear interactions among the sites, give ample opportunity for the accumulation of compensatory changes whose activity yields equivalent phenotypes between species but whose interactions are deleterious in the hybrid (Gibson, 1996; Veitia, 2003; Landry *et al.*, 2005). Large-scale analysis of interactions among loci in shaping the architecture of gene expression in yeast revealed that interactions among pairs of loci contribute to the pattern of inheritance of 57% of the transcripts, revealing that epistasis pervades the transcriptional networks (Brem *et al.*, 2005).

Evidence for RIs from global patterns of gene expression in interspecific hybrids

RIs can be studied in genetic crosses as species specificity in the activity of regulatory molecules or by DNA-mediated transfer of genetic elements from one species into the other. Different methodologies can be used to detect the consequences of RIs, most commonly through methods that measure the level of mRNA abundance (microarray technology and real-time quantitative polymerase chain reaction), that help to visualize the spatial pattern of expression of mRNA species and proteins (*in situ* hybridization), and that scrutinize directly the interactions between effector proteins and regulatory modules (*in vivo* and *in vitro* binding assays). Specifically, expression profiling of interspecific hybrids suggests that a significant fraction of the divergence in gene expression is cryptic, with more divergence present at the regulatory level than would be predicted from examination of parental phenotypes only (True and Haag, 2001). Recent studies of expression profiling on hybrids of several species pairs are presented in Table 1. In most of those studies, the abundance for many mRNA species in interspecific hybrids is often not intermediate between the parents but higher (overexpression) or lower (underexpression), as expected in a model where RIs accumulate along diverging lineages, which leads to extreme phenotypes in hybrids. The novel patterns of expression that emerge in hybrid individuals might correspond to perturbed interactions such as those between the transcription factors of one species and the regulatory sequences of the other, or other factors discussed below.

At first sight, genome-wide studies of gene expression in hybrids seem to suggest numerous RIs between the parental genomes whose non-additive interactions result in the misexpression of tens to hundreds of genes. However, the number of incompatibilities is not necessarily large because the numerous changes could actually be due to only a few major regulatory genes whose effects cascade through the network affecting the expression of many genes. Thus, even if the cause of gene misregulation is among actual molecules of the transcriptional network to which belongs the misregu-

Table 1 Comparative studies of the expression profile of interspecific hybrids relative to parental species using microarray technology

Parental species		Samples	Genes surveyed	Interspecific differences	Type of misexpresion in the hybrid		Reference
					Underexpression	Overexpression	
Animals							
(Female)	*(Male)*						
D. melanogaster	*D. simulans*	Female adults	~4500	1946	1772	1311	(Ranz *et al.*, 2004)
		Female heads	~4500		61	84	
D. simulans	*D. mauritiana*	Male adults	~14 000	51	37	14	(Michalak and Noor, 2003)
D. simulans	*D. melanogaster*	Testes	10 500	1541	1063	475	(Haerty and Singh, 2006)
	D. mauritiana	Testes	10 500	200	89	73	
	D. sechellia	Testes	10 500	241	235	2	
D. simulans	*D. mauritiana*	Male adults	12 012	3420	505	63	(Moehring *et al.*, 2007)
	D. sechellia			2487	190	30	
D. simulans	*D. mauritiana*	Male adults	48 sperm genes	8	7	1	
	D. sechellia			13	11	2	
D. simulans	*D. mauritiana*	3rd instar male larvae	12 012	Negligible	—	—	
	D. sechellia						
D. simulans	*D. mauritiana*	3rd instar male larvae	48 sperm genes	Negligible	—	—	
	D. sechellia						
Plants							
H. annuus	*H. petiolaris*	Entire seedlings	2975	206	26	32	(Lai *et al.*, 2006); the hybrid species is named *Heliantus deserticola*.
A. thaliana	*A. arenosa*	Leaves[a]	11 199	1362	1038	320	(Wang *et al.*, 2006)
		Leaves[b]		1469	952	500	
Senecio vulgaris	*Senecio squalidus*	Mature flower buds[c]	6385	397	0	39	(Hegarty *et al.*, 2005)
		Mature flower buds[d]			0	6	

[a]Correspond to the hybrid lines Allo733.
[b]Correspond to the hybrid lines Allo738.
[c]Correspond to the hybrids *Senecio × baxteri*.
[d]Correspond to the hybrids *S. cambrensis*.

lated gene, it is not clear how many actual 'incompatibilities' there are. The expression level of a particular gene is a phenotype and therefore not necessarily a property of the gene itself; rather the genetic basis of variation in gene expression may be due to factors that act in *trans*. The high pleiotropy of transcriptional networks is substantiated by deletion experiments on some model organisms. For example, in budding yeast, it is not uncommon for a gene deletion to affect the expression of tens to hundreds of genes (Featherstone and Broadie, 2002). Furthermore, regulatory anomalies do not necessarily result from non-additive interaction within the transcriptional network. Since many of the studies are performed on whole animals or mixed tissues, anomalous transcription in the hybrid may be uninformative in regard to transcription within individual cells. For example, the atrophied gonads and a hypertrophied fat body may contribute to the abnormal transcription profile of female hybrids between *D. melanogaster* and *D. simulans* (Ranz et al., 2004). Finally, the presence of two divergent genomes in the same cells might induce a 'stress' signal that affects the transcription of many genes exclusively in the hybrid (Comai et al., 2003). Epigenetic changes such as abnormal methylation patterns seen in synthetic allotetraploids may also be important (Madlung et al., 2002). Finally, and most importantly, non-additivity in expression levels in F1 hybrids does not appear to be limited to crosses between species and appear to be important in crosses between inbred lines of *Drosophila* (Gibson et al., 2004). Novel expression patterns in F1 hybrids between species should therefore be interpreted with caution and further dissection of the expression profiles is warranted.

Different approaches have been used to evaluate the relative importance of actual RIs as opposed to downstream effects in contributing to novel gene expression. One of the first approaches was to restrict the study to specific tissue types so that the effects of developmental anomalies could be minimized. For example, in the *D. melanogaster × D. simulans* hybrids studied by Ranz et al. (2004), expression profiling of heads showed that, although the patterns were weaker, a high fraction of genes still showed expression levels outside the parental range. Others studies have not only focused on particular tissues or organs but also have focused on more closely related species, which have presumably accumulated fewer changes at the regulatory level. Thus, one study compared the expression profile in testes of the *Drosophila* species of the *simulans* clade and their hybrids (Haerty and Singh, 2006). The species used in the analysis shared an ancestor only 0.93 million years ago (Tamura et al., 2004), roughly one-fifth of the divergence time between the species analyzed by Ranz et al. (2004). Further, time-course experiments on developing hybrids before the onset of the novel phenotype can help to establish the temporal order of events. Thus, Barbash and Lorigan (2007) studied the pattern of abnormalities in gene expression associated with the lethality of hybrid male larvae between *D. melanogaster* and *D. simulans*. To this end, they used a strain that rescues the lethal phenotype to give rise to viable hybrid male larvae. The authors found that a relatively small fraction of genes were misexpressed and among them, those encoding proteosome subunits and those related to the immune system, were enriched. In the case of genes that encode proteosome subunits, the authors found no solid evidence for a causal role in the lethal phenotype, rather their misregulation may be a consequence of the hybrid lethality. Likewise, Moehring et al. (2007) compared the expression profiles of larvae and adults in order to determine when the phenotype of sterility of the hybrid males in the *simulans* clade appears. The authors found minimal misexpression in larvae as compared to adults, supporting the idea that disruptions in spermatogenesis occurs preferentially after the larval stage. Finally, an alternative approach to the use of interspecific hybrids that should help in the identification of genes experiencing RIs involves engineered strains. In those strains, part of the genome of one species has been replaced through genetic crosses with the orthologous fraction of the genome of a second species. The reduced number of exogenous genes (compared to 50% in interspecific hybrids) might be expected to result in fewer misregulated genes.

Some of the studies in Table 1 have been valuable for uncovering biologically coherent patterns of transcriptome divergence, for tackling the phylogenetic dynamics of misexpression, and for determining if regulatory incompatibilities accumulate at random within expression networks or whether there are fractions of the expression network space that are preferentially targeted. Genes that are sex-biased in expression, that is those functionally more closely related to sex and reproduction, diverge faster at the level of mRNA abundance than genes that are not sex-biased (Ranz et al., 2003). Accordingly, those genes that are sex-biased in expression should be the initial targets of RIs in hybrid individuals. Independent expression profiling of *Drosophila* hybrids agrees with this trend, which is largely accounted for by genes associated with the reproductive biology of the males (Michalak and Noor, 2003; Ranz et al., 2004). Recently, Haerty and Singh (2006) studied the expression profiles of hybrid males resulting from crossing *D. simulans* females with males from species in the *melanogaster* group. Consistent with previous results, the authors found that it was particularly genes with male-biased expression that experienced RI (and/or its consequences) when the cross was between species separated by short phylogenetic distances. However, these genes made up a smaller proportion of misexpressed genes at greater phylogenetic distances.

Haerty and Singh (2006) and Moehring et al. (2007) compared the overlap between the sets of genes that are affected by RIs in interspecific hybrids. The two studies used species pairs separated by virtually the same phylogenetic distance. Strikingly, Moehring et al. (2007) found a much more substantial overlap (128 genes of a total of 220 in the cross *D. simulans × D. sechellia* and of a total of 568 in the cross *D. simulans × D. mauritiana*) than Haerty and Singh (2006) (16 genes of a total of 237 in the cross *D. simulans × D. sechellia* and of a total of 162 in the cross *D. simulans × D. mauritiana*). This apparent contradiction might be largely explained by methodological differences between investigations, which included the experimental design, the type of microarray platform used and the strains used. Further studies will be necessary to determine if the dynamics of accumulation of RIs follows a similar pattern across lineages and the relationship between sex-biased gene expression and the rapid accumulation of RIs.

A different set of genome-wide surveys of gene expression has been performed in plants, which seem to accommodate the combination of divergent regulatory systems better than animals. The results have helped identify genes and pathways that underlie reproductive isolation and adaptation. For example, the hybrid sunflower *Helianthus deserticola*, unlike its parental species *H. annuus* and *H. petiolaris*, grows well in extreme desert floor habitat. Its resistance to desiccation may be related to the preferential over- and underexpression of genes encoding transporter proteins (Lai *et al*., 2006). Other genome-wide surveys have investigated changes in the transcriptome accompanying hybridization and polyploidization. Hegarty *et al*. (2005) and Hegarty *et al*. (2006) compared the expression profile of *Senecio cambrensis* (an allohexaploid hybrid), its parent species *S. squalidus* (diploid) and *S. vulgaris* (tetraploid), and their triploid sterile F$_1$ hybrid, *S. baxteri*. The authors found that a significantly larger fraction of genes are expressed outside the range of the parent species in *S. baxteri*, compared to *S. cambrensis*, which they interpreted as the result of a transition from the initial regulatory instabilities of the original F1 hybrid compared to a more stabilized situation after polyplodization. Genome doubling would help ameliorate the impact of hybridization on transcription. This interpretation was reinforced by a subsequent comparison of the expression profiles of synthetic *S. cambrensis* maintained across five generations and the naturally occurring *S. cambrensis* (Hegarty *et al*., 2005, 2006).

Evidence of RIs from studies on gene transfer

Although still scarce, increasing evidence indicates that compensatory changes in *cis* and in *trans* might be common. For instance, studies that have monitored the expression profile of a particular gene from one species inserted into the genetic background of another (Mitsialis and Kafatos 1985) suggest that expression profiles evolve as the result of the accumulation of coordinated changes both in *cis* and in *trans* (Schiff *et al*., 1992; Wittkopp *et al*., 2002). *cis*–*trans* coevolution has been explicitly invoked in *Drosophila* (Ludwig *et al*., 2005; Marcellini and Simpson, 2006), *Caenorhabditis* (Ruvinsky and Ruvkun, 2003) and in ascidians (Takahashi *et al*., 1999; Oda-Ishii *et al*., 2005). One of the best examples comes from the study of the gene *Endo16* in early sea urchin development (Romano and Wray, 2003). Despite a conserved pattern of expression between two distantly related species of sea urchins, *Strongylocentrotus purpuratus* and *Lytechnius variegatus*, the promoter sequences of the gene *Endo16* has been reshaped extensively. Reciprocal transformation with a reporter gene construct containing the promoter region of the two species shows novel patterns of expression in embryos, suggesting that changes in *trans* of the genes have also accumulated to compensate for the changes in *cis* to preserve the expression between the two species. Another example comes from the study of the Bicoid (Bcd)-binding sites in the upstream of the genes *hunchback* (*hb*) and *tailless* (*tll*) of *D. melanogaster* and the house fly *Musca domestica*. This study combined transgenics with *in vitro* binding assays to provide one of the first comparative measures of the extent of regulatory incompatibility between species (Shaw *et al*., 2002; Wratten *et al*., 2006). There is not only evidence for coevolution between the *trans*- and *cis*-regulatory elements, but also evidence that the incompatibility does not result in a reduced binding, but rather higher affinity. It was found that the *Drosophila* Bcd transcription factor binds the *tll* promoter region of *Musca* with greater affinity than that of *Drosophila*, while the *Musca* Bcd binds the *Musca* promoter with greater affinity than the *Drosophila*. Similar results have been found for the Cys2-His2 zinc-finger transcription factor Rpn4p in *S. cerevisiae* and *C. albicans* (Gasch *et al*., 2004). This protein controls the activity of proteasome genes. *In vitro* binding assays demonstrated that amino-acid differences in the DNA-binding domain of Rpn4p were responsible for altered specificity, which mirrors the divergence of the regulatory sequences with which they interact in the species. These gene-specific studies have so far been concerned mostly with distantly related species. Whether these findings apply to more closely related ones remains to be investigated.

Evidence of RIs from allele-specific expression assays

The conservation of function in the face of sequence divergence is strongly suggestive that compensatory changes also accumulate in *trans* to preserve the expression patterns of the gene. Divergence in *cis* and *trans* between species can interact in hybrids to produce novel patterns of expression. The advantage in studying these interactions is that *cis*-regulatory elements are usually closely linked to the gene and therefore are a component of the gene itself. Accordingly, studying allele-specific expression level can indicate the occurrence of changes in *cis* and *trans*.

A polymorphism in the promoter or other *cis*-regulatory region affects only the expression of the nearby gene and hence identical *cis*-regulatory elements should have equal effects on gene expression in a hybrid genetic background. This property of genetic regulatory systems can be used to detect *cis*-regulatory divergence between alleles of a gene, and was used as early as 1960s to study divergence in gene expression by assaying allozyme expression in human–mouse hybrid cells (Ohno, 1969) and later in *Drosophila* hybrids (Dickinson and Carson, 1979) and fishes (Parker *et al*., 1985). Since the two alleles and their *cis*-regulatory elements share the same pool of *trans* acting factors, unequal abundance of transcripts of the two alleles suggests the presence of genetic variation acting in *cis*. In the cases where crosses are performed between two inbred lines or closely related species, the divergence in gene-expression level between parental lines can be compared to the difference between alleles in the F$_1$ hybrid. The difference between the parental lines not assigned to *cis*-divergence is then assigned to *trans*-divergence (Wittkopp *et al*., 2004).

A simple example could be the loss of a transcription-factor-binding site in a regulatory region that can be compensated for by an increase in transcription factor activity. In this case, changes in *cis* and *trans* act in opposite directions, antagonistically, to maintain the gene-expression level (Figure 1), with changes in *cis* acting to lower transcript abundance and changes in *trans* acting to increase it. What would be expected were the two species to be crossed? The two alleles would
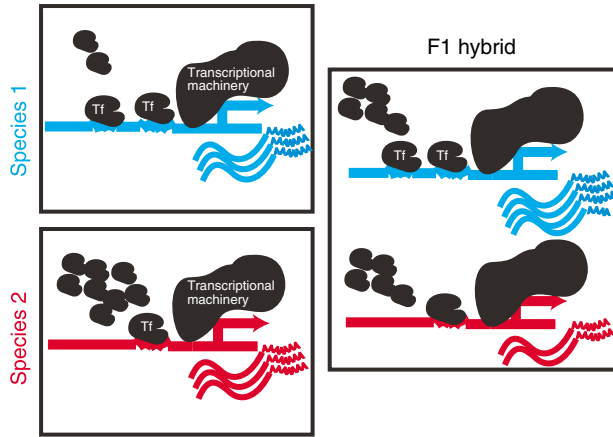
Figure 1 Schematic model of compensatory changes in *cis* and *trans* between species. Changes in *cis* and *trans* have accumulated between species but the expression level is conserved because these changes compensate one another. Here the differences are represented as a change in the number of DNA-binding sites and the abundance of transcription factor.

Figure 2 Divergence in gene expression between species and between alleles in F_1 hybrids of *Drosophila*. The *x* and *y* axes represent, respectively, the ratio of gene expression between species (on the log-2 scale) and the ratio of allelic expression in F_1 hybrid. The gray diagonal line represents the expected ratios for *cis*-regulatory variation that explains the between-species divergence. In red are genes that show *cis*-regulatory divergence only; in blue, *trans*-regulatory divergence only; in orange, *cis*- and *trans*- divergence and in green, *cis*–*trans* compensatory changes. The ratio of allelic expression in the hybrids is greater than between species or in the opposite direction. Original results from Landry *et al.* (2005).

now be in the same genetic background, but their *cis*-regulatory sequences are not equivalent. The two alleles would therefore be differentially expressed. One would consequently expect to observe a ratio of allelic expression in the F_1 hybrid that is greater than that observed between the species.

This rationale motivated a study of hybrids between *D. melanogaster* and *D. simulans*, which showed that 28 of 29 genes studied had divergent *cis*-regulatory elements (Wittkopp *et al.*, 2004). Several of the changes were compensatory changes between the species. Another study on the same species pair revealed that *cis*–*trans* compensatory changes may be abundant. Figure 2 presents the results from Landry *et al.* (2005). In this study, a large fraction of genes showed a difference in expression between alleles that was more extreme than that between the species. In many other cases, the allele that was more highly expressed in one parental species was the least expressed in the hybrid background. Interestingly, in several of these cases, the gene was also overexpressed or underexpressed in hybrids relative to the parental species, suggesting that these cases may well be the result of RIs leading to novel phenotypes in interspecific hybrids.

It is still unclear whether these *cis*–*trans* interactions are also present within species or whether they require a longer divergence time to accumulate. Few large-scale studies have been performed on allele-specific expression within species. In a recent study, allele-specific expression was estimated for 35 genes that were differentially expressed between two inbred lines of maize, B73 and Mo17 (Stupar and Springer, 2006). Most genes (31/35) showed differential allelic expression in the F_1 hybrid between the two lines, suggesting abundant *cis*-regulatory variation in gene expression. However, only four genes showed greater differences in allelic expression between alleles than between inbred lines and the differences were of small magnitude. Further studies within and between species of the same taxa will be necessary to ascertain whether the relative levels of *cis*- and *trans*-regulatory differences between species are also found within species. So far, this study
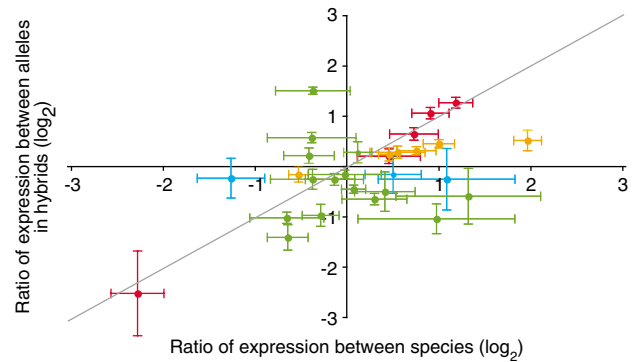
on maize suggests that *cis*–*trans* compensatory regulation and thus hidden coevolution may be less abundant within species than between species.

Future technological approaches

One big step forward in this field of research would be to be able to assess the regulatory activity of the two diverged genomes independently in F1 hybrids. Microarray technology can be applied to these questions by allowing simultaneous measurement of allelic expression in hybrids between species. In cases in which a large fraction of the genome of the two species has been sequenced, oligonucleotides specific to the alleles of the two species can be designed to independently estimate the expression of the two alleles. This approach has been applied to yeast with the use of Affymetrix arrays (Ronald *et al.*, 2005). Affymetrix arrays are designed such that expression level of each transcript is assayed with multiple probes distributed along the gene. Across two strains of yeast, the standard laboratory strain (S288c), from which the probes were designed, and a natural strain (RM) were performed. Several probe sets contained at least one probe that had a mismatch with the sequence of the RM gene. In these cases, the probe with a mismatch measures the expression of the S288c allele only. With only a few microarray measurements, 70 cases of differential allelic expression could be detected—when the expression level measured by the mismatched probe was not equal to half that measured by the other probes (as expected if the two alleles were equally expressed). This study elegantly illustrates how microarray technologies would allow one to interrogate the transcriptional activity of thousands of genes in hybrids between species. For species that have diverged sufficiently, it will be possible to design oligonucleotide arrays that will interrogate each allele specifically.

With the decreasing cost of DNA sequencing technologies and microarray platforms, it is likely that tools will soon be available to study the integration of divergent

transcriptional networks in interspecific hybrids. Novel sequencing technologies sequence DNA templates in a quantitative manner (Bentley, 2006). For instance, the 454 and Solexa sequencing technologies enable massively parallel sequencing of millions of DNA fragments with each sequence read on the order of a hundred nucleotides. If cDNA is sequenced instead of genomic DNA, the transcript abundance of the alleles from the two species can be compared and any discrepancy from equal allelic expression can be detected, allowing *cis*- and *trans*-regulatory divergence to be identified on a genomic scale. Furthermore, this approach does not require prior knowledge of the sequences of the two alleles.

Final remarks

The use of new methodologies that monitor different components of gene expression in interspecific hybrids or within the experimental framework of transformation experiments between species has provided ample evidence supporting the model under which interacting elements within the expression network coevolve and this leads to novel expression profiles when they are brought together in hybrids. Whether these novel expression profiles map to deleterious or innovative phenotypes typically observed in interspecific hybrids is currently unknown. However, current theoretical population genetics models suggest that the effects of accumulation of RIs and the architecture of transcriptional networks can directly influence the process of speciation (Porter and Johnson, 2002; Johnson and Porter, 2007).

Coevolution at the molecular level has largely been focused on the interaction between transcription factors and regulatory modules. Our increasing knowledge about the complexity of the regulation of gene expression forces us to explore other type of interactions such as those mediated by miRNAs, which may be critically important in explaining the novel expression phenotypes that are found in interspecific hybrids.

Finally, the population genetics of the accumulation of RIs has been mostly left untouched. Whether these accumulate through compensatory changes among interacting gene products or gene products and regulatory sequences or through a neutral process similar to the Dobzhansky–Muller neutral model (Dobzhansky, 1937; Muller, 1940) has been completely unexplored. The relative contribution of these two processes will depend not only on the classical population genetics parameters of the species studied but likely also on the genetic and molecular architecture of transcriptional networks, which we only begin to explore and understand.

Acknowledgements

References

Arnone MI, Davidson EH (1997). The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124**: 1851–1864.

Arnosti DN (2003). Analysis and function of transcriptional regulatory elements: insights from *Drosophila*. *Annu Rev Entomol* **48**: 579–602.

Barbash DA, Lorigan JG (2007). Lethality in *Drosophila melanogaster/Drosophila simulans* species hybrids is not associated with substantial transcriptional misregulation. *J Exp Zool B Mol Dev Evol* **308**: 74–84.

Belting HG, Shashikant CS, Ruddle FH (1998). Modification of expression and *cis*-regulation of Hoxc8 in the evolution of diverged axial morphology. *Proc Natl Acad Sci USA* **95**: 2355–2360.

Bentley DR (2006). Whole-genome re-sequencing. *Curr Opin Genet Dev* **16**: 545–552.

Boyl PP, Signore M, Acampora D, Martinez-Barbera JP, Ilengo C, Annino A *et al.* (2001). Forebrain and midbrain development requires epiblast-restricted Otx2 translational control mediated by its 3′ UTR. *Development* **128**: 2989–3000.

Brem RB, Storey JD, Whittle J, Kruglyak L (2005). Genetic interactions between polymorphisms that affect gene expression in yeast. *Nature* **436**: 701–703.

Carroll SB (2005a). *Endless Forms Most Beautiful: the New Science of Evo Devo and the Making of the Animal Kingdom*, 1st edn. W.W. Norton & Co.: New York.

Carroll SB (2005b). Evolution at two levels: on genes and form. *PLoS Biol* **3**: e245.

Carroll SB, Grenier JK, Weatherbee SD (2001). *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. Blackwell Science: Oxford.

Carter AJ, Wagner GP (2002). Evolution of functionally conserved enhancers can be accelerated in large populations: a population-genetic model. *Proc Biol Sci* **269**: 953–960.

Castillo-Davis CI, Kondrashov FA, Hartl DL, Kulathinal RJ (2004). The functional genomic distribution of protein divergence in two animal phyla: coevolution, genomic conflict, and constraint. *Genome Res* **14**: 802–811.

Chiang DY, Nix DA, Shultzaberger RK, Gasch AP, Eisen MB (2006). Flexible promoter architecture requirements for coactivator recruitment. *BMC Mol Biol* **7**: 16.

Comai L, Madlung A, Josefsson C, Tyagi A (2003). Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids? *Philos Trans R Soc Lond B Biol Sci* **358**: 1149–1155.

Costas J, Casares F, Vieira J (2003). Turnover of binding sites for transcription factors involved in early *Drosophila* development. *Gene* **310**: 215–220.

Davidson EH (2006). *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution*. Elsevier/Academic Press: Amsterdam.

Dermitzakis ET, Clark AG (2002). Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. *Mol Biol Evol* **19**: 1114–1121.

Dickinson WJ, Carson HL (1979). Regulation of the tissue specificity of an enzyme by a *cis*-acting genetic element: evidence from interspecific *Drosophila* hybrids. *Proc Natl Acad Sci USA* **76**: 4559–4562.

Dobzhansky T (1937). *Genetics and the Origin of Species*. Columbia University Press: New York.

Dover GA, Flavell RB (1984). Molecular coevolution: DNA divergence and the maintenance of function. *Cell* **38**: 622–623.

Erives A, Levine M (2004). Coordinate enhancers share common organizational features in the *Drosophila* genome. *Proc Natl Acad Sci USA* **101**: 3851–3856.

Fan J, Yang X, Wang W, Wood III WH, Becker KG, Gorospe M (2002). Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proc Natl Acad Sci USA* **99**: 10611–10616.

Fasken MB, Corbett AH (2005). Process or perish: quality control in mRNA biogenesis. *Nat Struct Mol Biol* **12**: 482–488.

Featherstone DE, Broadie K (2002). Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network. *BioEssays* **24**: 267–274.

Galant R, Carroll SB (2002). Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**: 910–913.

Gasch AP, Moses AM, Chiang DY, Fraser HB, Berardini M, Eisen MB (2004). Conservation and evolution of *cis*-regulatory systems in ascomycete fungi. *PLoS Biol* **2**: e398.

Gibson G (1996). Epistasis and pleiotropy as natural properties of transcriptional regulation. *Theor Popul Biol* **49**: 58–89.

Gibson G, Riley-Berger R, Harshman L, Kopp A, Vacha S, Nuzhdin S et al. (2004). Extensive sex-specific nonadditivity of gene expression in *Drosophila melanogaster*. *Genetics* **167**: 1791–1799.

Gibson G, Weir B (2005). The quantitative genetics of transcription. *Trends Genet* **21**: 616–623.

Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005). Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**: 481–487.

Gonzalez P, Rao PV, Nunez SB, Zigler Jr JS (1995). Evidence for independent recruitment of zeta-crystallin/quinone reductase (CRYZ) as a crystallin in camelids and hystricomorph rodents. *Mol Biol Evol* **12**: 773–781.

Haerty W, Singh RS (2006). Gene regulation divergence is a major contributor to the evolution of Dobzhansky–Muller incompatibilities between species of *Drosophila*. *Mol Biol Evol* **23**: 1707–1714.

Hegarty MJ, Barker GL, Wilson ID, Abbott RJ, Edwards KJ, Hiscock SJ (2006). Transcriptome shock after interspecific hybridization in senecio is ameliorated by genome duplication. *Curr Biol* **16**: 1652–1659.

Hegarty MJ, Jones JM, Wilson ID, Barker GL, Coghill JA, Sanchez-Baracaldo P et al. (2005). Development of anonymous cDNA microarrays to study changes to the Senecio floral transcriptome during hybrid speciation. *Mol Ecol* **14**: 2493–2510.

Hsia CC, McGinnis W (2003). Evolution of transcription factor function. *Curr Opin Genet Dev* **13**: 199–206.

Inga A, Storici F, Darden TA, Resnick MA (2002). Differential transactivation by the p53 transcription factor is highly dependent on p53 level and promoter target sequence. *Mol Cell Biol* **22**: 8612–8625.

Jaenisch R, Bird A (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33** (Suppl): 245–254.

Jeong S, Rokas A, Carroll SB (2006). Regulation of body pigmentation by the abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* **125**: 1387–1399.

Johnson NA, Porter AH (2007). Evolution of branched regulatory genetic pathways: directional selection on pleiotropic loci accelerates developmental system drift. *Genetica* **129**: 57–70.

Juarez K, Flores H, Davila S, Olvera L, Gonzalez V, Morett E (2000). Reciprocal domain evolution within a transactivator in a restricted sequence space. *Proc Natl Acad Sci USA* **97**: 3314–3318.

Khodursky AB, Bernstein JA (2003). Life after transcription—revisiting the fate of messenger RNA. *Trends Genet* **19**: 113–115.

Kimura M (1985). The role of compensatory mutation in molecular evolution. *J Genet* **64**: 7–19.

Klingler M, Soong J, Butler B, Gergen JP (1996). Disperse versus compact elements for the regulation of runt stripes in *Drosophila*. *Dev Biol* **177**: 73–84.

Lai Z, Gross BL, Zou Y, Andrews J, Rieseberg LH (2006). Microarray analysis reveals differential gene expression in hybrid sunflower species. *Mol Ecol* **15**: 1213–1227.

Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL (2005). Compensatory *cis*-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. *Genetics* **171**: 1813–1822.

Lenormand T (2002). Gene flow and the limits to natural selection. *Trends Ecol Evol* **17**: 183–189.

Levine M, Tjian R (2003). Transcription regulation and animal diversity. *Nature* **424**: 147–151.

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R (2006). Interchromosomal interactions and olfactory receptor choice. *Cell* **126**: 403–413.

Ludwig MZ, Palsson A, Alekseeva E, Bergman CM, Nathan J, Kreitman M (2005). Functional evolution of a *cis*-regulatory module. *PLoS Biol* **3**: e93.

Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J, Comai L (2002). Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic Arabidopsis allotetraploids. *Plant Physiol* **129**: 733–746.

Marcellini S, Simpson P (2006). Two or four bristles: functional evolution of an enhancer of scute in drosophilidae. *PLoS Biol* **4**: e386.

Mata J, Marguerat S, Bahler J (2005). Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci* **30**: 506–514.

McGregor AP, Shaw PJ, Dover GA (2001a). Sequence and expression of the hunchback gene in *Lucilia sericata*: a comparison with other Dipterans. *Dev Genes Evol* **211**: 315–318.

McGregor AP, Shaw PJ, Hancock JM, Bopp D, Hediger M, Wratten NS et al. (2001b). Rapid restructuring of bicoid-dependent hunchback promoters within and between Dipteran species: implications for molecular coevolution. *Evol Dev* **3**: 397–407.

Michalak P, Noor MA (2003). Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. *Mol Biol Evol* **20**: 1070–1076.

Mitsialis SA, Kafatos FC (1985). Regulatory elements controlling chorion gene expression are conserved between flies and moths. *Nature* **317**: 453–456.

Moehring AJ, Teeter KC, Noor MA (2007). Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. II. Examination of multiple-species hybridizations, platforms, and life cycle stages. *Mol Biol Evol* **24**: 137–145.

Muller H (1940). Bearing of the *Drosophila* work on systematics. In: Huxley J (ed). *The New Systematics*. Clarendon Press: Oxford.

Oda-Ishii I, Bertrand V, Matsuo I, Lemaire P, Saiga H (2005). Making very similar embryos with divergent genomes: conservation of regulatory mechanisms of Otx between the ascidians *Halocynthia roretzi* and *Ciona intestinalis*. *Development* **132**: 1663–1674.

Ohno S (1969). The preferential activation of maternally derived alleles in development of interspecific hybrids. In: Defendi V (ed). *Heterospecific Genome Interaction*. Wistar Institute Press: Philadelphia, pp 137–150.

Oliver F, Christians JK, Liu X, Rhind S, Verma V, Davison C et al. (2005). Regulatory variation at glypican-3 underlies a major growth QTL in mice. *PLoS Biol* **3**: e135.

Papaceit M, Orengo D, Juan E (2004). Sequences upstream of the homologous *cis*-elements of the Adh adult enhancer of *Drosophila* are required for maximal levels of Adh gene transcription in adults of *Scaptodrosophila lebanonensis*. *Genetics* **167**: 289–299.

Parker HR, Philipp DP, Whitt GS (1985). Gene regulatory divergence among species estimated by altered developmental patterns in interspecific hybrids. *Mol Biol Evol* **2**: 217–250.

Parker R, Song H (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol* **11**: 121–127.

Porter AH, Johnson NA (2002). Speciation despite gene flow when developmental pathways evolve. *Evol Int J Org Evol* **56**: 2103–2111.

Prud'homme B, Gompel N, Rokas A, Kassner VA, Williams TM, Yeh SD et al. (2006). Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**: 1050–1053.

Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL (2003). Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300**: 1742–1745.

Ranz JM, Machado CA (2006). Uncovering evolutionary patterns of gene expression using microarrays. *Trends Ecol Evol* **21**: 29–37.

Ranz JM, Namgyal K, Gibson G, Hartl DL (2004). Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaste*r and *Drosophila simulans*. *Genome Res* **14**: 373–379.

Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM (1996). Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* **272**: 741–745.

Rockman MV, Wray GA (2002). Abundant raw material for cis-regulatory evolution in humans. *Mol Biol Evol* **19**: 1991–2004.

Romano LA, Wray GA (2003). Conservation of Endo16 expression in sea urchins despite evolutionary divergence in both *cis* and *trans*-acting components of transcriptional regulation. *Development* **130**: 4187–4199.

Ronald J, Akey JM, Whittle J, Smith EN, Yvert G, Kruglyak L (2005). Simultaneous genotyping, gene-expression measurement, and detection of allele-specific expression with oligonucleotide arrays. *Genome Res* **15**: 284–291.

Ruvinsky I, Ruvkun G (2003). Functional tests of enhancer conservation between distantly related species. *Development* **130**: 5133–5142.

Schiff NM, Feng Y, Quine JA, Krasney PA, Cavener DR (1992). Evolution of the expression of the Gld gene in the reproductive tract of *Drosophila*. *Mol Biol Evol* **9**: 1029–1049.

Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jonsson B et al. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**: 717–723.

Shaw PJ, Wratten NS, McGregor AP, Dover GA (2002). Coevolution in bicoid-dependent promoters and the inception of regulatory incompatibilities among species of higher Diptera. *Evol Dev* **4**: 265–277.

Stupar RM, Springer NM (2006). *Cis*-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. *Genetics* **173**: 2199–2210.

Takahashi H, Mitani Y, Satoh G, Satoh N (1999). Evolutionary alterations of the minimal promoter for notochord-specific Brachyury expression in ascidian embryos. *Development* **126**: 3725–3734.

Tamura K, Subramanian S, Kumar S (2004). Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol Biol Evol* **21**: 36–44.

Tautz D (2000). Evolution of transcriptional regulation. *Curr Opin Genet Dev* **10**: 575–579.

Thanos D, Maniatis T (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell* **83**: 1091–1100.

True JR, Haag ES (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evol Dev* **3**: 109–119.

Veitia RA (2003). A sigmoidal transcriptional response: co-operativity, synergy and dosage effects. *Biol Rev Camb Philos Soc* **78**: 149–170.

Wagner A (2005). *Robustness and Evolvability in Living Systems*. Princeton University Press: Princeton.

Wang J, Tian L, Lee HS, Wei NE, Jiang H, Watson B et al. (2006). Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* **172**: 507–517.

Wang X, Chamberlin HM (2002). Multiple regulatory changes contribute to the evolution of the *Caenorhabditis* lin-48 ovo gene. *Genes Dev* **16**: 2345–2349.

Weinreich DM, Chao L (2005). Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evol Int J Org Evol* **59**: 1175–1182.

Whitehead A, Crawford D (2006). Variation within and among species in gene expression: raw material for evolution. *Mol Ecol* **15**: 1197–1211.

Wilkins AS (2002). *The Evolution of Developmental Pathways*. Sinauer Associates: Sunderland.

Wittkopp PJ, Haerum BK, Clark AG (2004). Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430**: 85–88.

Wittkopp PJ, Vaccaro K, Carroll SB (2002). Evolution of yellow gene regulation and pigmentation in *Drosophila*. *Curr Biol* **12**: 1547–1556.

Wratten NS, McGregor AP, Shaw PJ, Dover GA (2006). Evolutionary and functional analysis of the tailless enhancer in *Musca domestica* and *Drosophila melanogaster*. *Evol Dev* **8**: 6–15.

Wray GA (2007). The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet* **8**: 206–216.

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV et al. (2003). The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* **20**: 1377–1419.

Zamore PD, Haley B (2005). Ribo-gnome: the big world of small RNAs. *Science* **309**: 1519–1524.