

# Gene expression in early *Drosophila* embryos is highly conserved despite extensive divergence of transcription factor binding.

Mathilde Paris<sup>1</sup>, Tommy Kaplan<sup>1,2</sup>, Xiao Yong Li<sup>1,3</sup>, Jacqueline E. Villalta<sup>2</sup>, Susan E. Lott<sup>1,4</sup>, Michael B. Eisen<sup>1,2</sup>

<sup>1</sup> Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, California, United States of America, <sup>2</sup> School of Computer Science and Engineering, The Hebrew University, Jerusalem, Israel, <sup>3</sup> Howard Hughes Medical Institute, University of California Berkeley, Berkeley, California, United States of America, <sup>4</sup> Department of Evolution and Ecology, University of California, Davis, California, United States of America.

Corresponding authors: MP and MBE

## **ABSTRACT**

To better characterize how variation in regulatory sequences drives divergence in gene expression, we undertook a systematic study of transcription factor binding and gene expression in the blastoderm embryos of four species that sample much of the diversity in the 60 million-year old genus *Drosophila*: *D. melanogaster*, *D. yakuba*, *D. pseudoobscura* and *D. virilis*. We compared gene expression, as measured by mRNA-seq to the genome-wide binding of four transcription factors involved in early development, as measured by ChIP-seq (Bicoid, Giant, Hunchback and Krüppel). Surprisingly, we found that mRNA levels are much better conserved than individual binding events. We looked at binding characteristics that may explain such evolutionary disparity. As expected, we found that binding divergence increases with phylogenetic distance. Interestingly, binding events in non-coding regions that were bound strongly by single factors, or bound by multiple factors, were more likely to be conserved. As this class of sites are most likely to be involved in gene regulation, the divergence of other bound regions may simply reflect their lack of function. We used a model of quantitative trait evolution to compare the changes of gene expression with nearby regulatory TF binding. We found that changes in gene expression were poorly explained by changes in associated TF binding. These results suggest that some of the differences in sequence and binding have limited effect on gene expression or act in a compensatory manner to maintain the overall expression levels of regulated genes.

## **INTRODUCTION**

In their pursuit of the genetic basis of phenotypic evolution, researchers have favored divergence of gene expression as a major source of diversity [1,2]. This divergence is thought to find its origins in the divergence of transcriptional regulation, notably the divergent binding of transcription factors to specific DNA sequences. These DNA sequences often undergo rapid changes during evolution [3-6]. In agreement, genome-wide binding of transcription factors has often been shown to be extensively divergent in related species, where, based on functional conservation, little divergence was expected [7-15]. While many studies have compared gene expression between species at the genomic scale and reported

various degrees of divergence (e.g. [16-18]), the downstream effects of TF binding divergence on gene expression has received very little attention and it remains unclear how differences in DNA lead to changes in gene expression.

In this study, we have explored the relationship between the divergence of transcription factor binding and gene expression. We measured gene expression along with genome-wide binding profiles of four TFs that together regulate the early steps of segmentation along the anterior-posterior (A-P) axis in four *Drosophila* species that sample much of the diversity in the 60 million-year old genus *Drosophila*. These species, *D. melanogaster*, *D. yakuba*, *D. pseudoobscura* and *D. virilis* diverged between 5 and 40 million year ago [19] (Figure 1), spanning evolutionary distances up to 1.5 substitutions per neutral site, equivalent to the split among amniotes [20].

We focused on A-P patterning in *Drosophila* embryos for several reasons: it has been extensively studied in the model species *D. melanogaster*, and it occurs very early during development, at the onset of zygotic transcription, in a comparatively simple transcriptional context. In addition, segmentation along the A-P axis is phenotypically highly conserved, allowing us to sample extensive sequence divergence without the complications that arise from comparing species with significantly different morphologies.

## **RESULTS**

### **Chromatin immunoprecipitation (ChIP) from divergent *Drosophila* species**

We established large populations of *D. melanogaster* (Oregon R), *D. pseudoobscura* (MV2-25) and *D. virilis* (V46) for embryo collections. In our studies of transcription factor binding in *D. melanogaster* we generally use embryos from one hour long collections aged for an additional two hours to target the cellular blastoderm stage [14,21,22], during which many key events in patterning occur. However, because developmental timing varies between species, we optimized collection conditions for each species to obtain similar stage distributions (Table S1).

We fixed embryos in 1% formaldehyde to cross-link proteins to DNA, and purified chromatin. We immunoprecipitated cross-linked chromatin from all three species using rabbit polyclonal antibodies raised against *D. melanogaster* Bicoid (BCD), Hunchback (HB), Giant (GT) and Krüppel (KR) [22]. For each factor, we performed parallel ChIP using antibodies that were purified against recombinant versions of the corresponding *D. melanogaster* or *D. virilis* proteins (we were unable to affinity purify GT antisera against *D. virilis* GT). We used *D. virilis* proteins to avoid biases due to the recognition of a greater number of epitopes in *D. melanogaster* than in the more distantly related *D. pseudoobscura* and *D. virilis*. Following immunoprecipitation (IP), we sequenced the recovered DNA as well as input controls.

We generated a total of 20 ChIP datasets to which we added 8 datasets from our previously published comparison of the binding of these factors in *D. melanogaster*

and the closely related *D. yakuba* ([14], see Figure S1). We mapped reads to the corresponding genome sequences with Bowtie [23], and identified genomic regions significantly bound by each factor in any of the four species using two peak callers, Grizzly Peak [24,25] and MACS [26]. Roughly similar numbers of peaks were found per ChIP (Table S2). Replicates showed good reproducibility, especially for GT and HB IPs (Figure S2).

### **Extensive divergence of transcription factor binding**

To compare binding across species, we first aligned the genomes of all four species using Mercator [27] and PECAN [28] and identified orthologous regions present in all four genomes. We then projected the normalized and log transformed binding intensities of bound regions from all IPs onto the coordinates of the whole-genome alignment and compared occupancy, as illustrated for the *even-skipped* locus (Figure 2). We obtained 2061, 4191, 4986 and 5309 sets of orthologous regions bound by BCD, GT, HB and KR, respectively. We also collapsed the 16,547 regions identified as bound by individual transcription factors, and merged them into a common set of 10,137 merged regions, bound by at least one factor in at least one species (see Figure 2). We then computed the occupancy of every factor in each species along each of these regions and compared binding intensities (Figure 3A and Figure S3). As expected, the extent of conservation of sites for each factor decreased with increasing phylogenetic distance from *D. melanogaster* (Figure 3B). Relatively few regions were bound by any factor in all four species (Figure 3C), demonstrating that transcription factor binding has diverged considerably.

### **What is driving binding divergence/conservation?**

Given the high degree of divergence of TF binding, we evaluated the characteristics of bound regions that were associated with differing degrees of divergence. We used a Brownian motion model for the evolution of quantitative traits [29-31], commonly used in molecular evolutionary studies [32-34], to impute binding values at the three internal nodes on the tree (Figure 1), including the root, and computed changes in binding along each of the six branches of the tree. It takes into account the inherent phylogenetic inertia in our dataset (the historical dependency between species), in contrast to pairwise comparisons. We then compared these estimates of TF binding divergence to changes in the sequence of the corresponding bound regions. We found a small, although significant, correlation between substitution rate and binding divergence (Figure 4A and S4).

We next asked whether changes in DNA binding properties may explain binding divergence. We first verified that each TF exhibited similar DNA binding properties in the different species (Figure 4B and Figure S5), and used these conserved binding profiles to predict TF binding in all bound regions (Figures 4C and S6). We again applied a Brownian motion model to impute predicted binding values at the three internal nodes on the tree, and computed changes in predicted binding along each of the six branches of the tree. We found that the gains and losses of known motifs for each factor were correlated with corresponding changes in its binding along the

phylogenetic tree (Figures 4D-4E and S7). Overall, these results suggest that changes in binding locations of BCD, GT, HB and KR are caused in part by alterations of specific DNA-binding sites.

We then examined the relationship between peak conservation and three additional factors: peak height, location of peak relatively to genes, and binding of other factors to the same region. All three were correlated with peak conservation (Figure 5A-D). Of note, these covariates were not independent of each other – higher peaks are more likely to be clustered with peaks for other factors in non-coding regions.

We previously observed that the occupancy of these four factors (and others) is highly correlated in *D. melanogaster* [22,35], and *D. yakuba* [14]. We analyzed correlations amongst the binding of all four factors in each species independently or in the union of bound regions in all species using principal components analysis, and found that the same correlation is present in each species, including in *D. pseudoobscura* and *D. virilis*, as well as in the combined dataset from the four species (Figures 5C and S8). In *D. melanogaster*, this axis is highly correlated with the DNA-binding levels of the protein Zelda (zld) (Figures 6C and S9), in agreement with previous studies showing that the correlation among the binding of different factors is driven by the early binding and chromatin-shaping activities of Zelda [22,25]. Zelda is present in all *Drosophila* species, its CAGGTAG binding site is enriched under the peaks from all species (Figure 6A), and the Zelda binding site predicts PC1 in all four species (Figure S9B), suggesting that its effect on chromatin and transcription factor binding is conserved.

In addition to correlations in binding of different factors within a single species, we also previously found that changes in binding of different AP factors were correlated between *D. melanogaster* and the closely related *D. yakuba* [14]. We repeated this analysis on the four species dataset and find that the relationship between binding changes that affect all four factors and changes in Zelda motifs extends to the entire tree (Figure 6D-E). Although the effect is weaker, it is significant.

Overall, using all parameters described above (TF-specific and Zelda motif enrichments, proximity to genes, number of other TFs binding the same locus), a multiple linear regression allowed us to explain about 29-36% of the variance in binding within each species (Pearson correlations  $\sim 0.6$ , Figure S10 A-D). In comparison, 3-7% of the variance in TF binding divergence between species could be explained using the same parameters (7-21% if phylogenetic inertia is not taken into account, Figure S10 E-L). This divergence occurs mostly through changes in TF-specific binding motifs, as well as potentially coordinated changes between TFs and with Zelda.

### **mRNA levels are better conserved than TF binding overall**

To investigate how this level of divergence of transcription factor binding affects gene expression, we sequenced mRNA from embryos from each species harvested at the end of cellularization using high-throughput mRNA sequencing (mRNAseq). We

first refined the annotations for the non-*melanogaster* species using cufflinks [36] and pools of ~50 embryos spanning the same stages used in the ChIP experiments. This method builds on an existing annotation and mRNAseq data to discover novel transcripts. The annotated transcriptomes size increased by 17%, 19% and 24% for *D. yakuba*, *D. pseudoobscura* and *D. virilis*, respectively (Table S3). We used these improved annotations plus the *D. melanogaster* reference annotation and restricted our analyses to the 8,555 protein-coding genes for which we could identify orthologs present exactly once in all four species. We then sequenced mRNA levels in several individuals from each species and found that mRNA levels are highly correlated between species (Figure 7A), with the correlation decaying with phylogenetic distance (Figure 7B). In addition, overall divergence of mRNA levels was significantly lower than divergence of any TF binding, as measured by the variance of normalized data along the *Drosophila* tree (Figure 7C).

### **TF binding divergence is poorly correlated with divergence of mRNA levels**

To directly compare variation in transcription factor binding to variation in gene expression, we first matched each gene to the largest and closest bound region, recognizing that some of these associations were likely to be incorrect. We focused on the 4846 genes that were annotated in all four species and expressed in at least one of them. Of these, 3024 could be associated with nearby binding of at least one transcription factor in one species. We partitioned genes depending on whether transcripts in *D. melanogaster* blastoderm embryos were deposited by the mother into the egg, or were a product of zygotic transcription, as defined by [37]: 2056 genes were categorized as “maternal”, 388 as “zygotic” and 394 as “both maternal and zygotic” (these are both deposited by the mother and transcribed by the zygote; the 186 remaining genes were not categorized).

We first analyzed for each gene the relationship between mRNA levels and nearby TF binding. TF binding, and especially strong and clustered binding, was preferentially localized near zygotic genes (Figure 7D). Within a species, 17% to 23% of variance in mRNA levels of zygotic genes could be attributed, by multiple regression, to variance of levels of associated TF binding. Accordingly, we found a significant correlation between measured mRNA levels and mRNA levels predicted from a multiple linear regression of associated nearby TF binding (Pearson correlation 0.46, p.value <  $10^{-16}$ , Figure 7E). Association between mRNA levels and nearby TF binding for maternally deposited genes was comparatively much weaker and intermediate for “maternal/zygotic” genes (Figure S11).

Despite this relatively good relationship between TF binding and gene expression within a species, we found a weak relationship between the variations in TF binding and gene expression along the tree (Figure 7F). Using multiple linear regression we found that trans-species variation in TF binding is positively correlated with trans-species variation in mRNA levels, but only ~2% of variance in mRNA levels could be attributed to variation in TF binding for zygotic genes using a linear regression (3% if phylogenetic inertia is not taken into account). Divergence of mRNA levels was

also weakly but significantly correlated with changes of associated TF binding near zygotic genes (Figure S12).

## **DISCUSSION**

This study addresses both the genomic causes and the consequences on gene expression of TF binding divergence. We found that binding of anterior-posterior (AP) factors has changed extensively along the *Drosophila* tree, with little effect on downstream gene expression. We identified two potential genetic sources of binding divergence. We found that alterations of BCD, HB, GT and KR-specific sites drive binding divergence, as has been repeatedly found for other TFs [7,11,13,14,38], although our study is among the first ones to quantify this effect. This suggests that the binding differences seem to be due to an appreciable degree to differences in the *cis*-regulatory regions more than functional differences of the TFs themselves [39]. Interestingly, we found that the AP factors are undergoing correlated evolution among themselves (Figures 5D and 6D) and potentially with the protein Zelda (figure 6E). This suggests that, in addition to TF specific changes in binding, Zelda may be orchestrating more global changes that influence several TFs at the same time.

The weak correlation of changes in transcription factor binding and gene expression may seem unexpected. Given the high degree of morphological conservation amongst these four species, we predicted and observed highly conserved embryonic gene expression – both maternal and zygotic. Yet conservation in gene expression was not coupled with conservation in transcription factor binding. However, careful dissection of single enhancers has previously demonstrated conserved activity of *cis*-regulatory sequences or conserved gene expression despite extensive sequence divergence and TF binding motif turnover [40-42]. Our results, based on measured TF binding at a large scale, are thus consistent with these studies.

One relatively simple explanation is that our measurements of transcription factor binding are more subject to systematic biases and noise, since ChIP experiments are less robust than mRNA-seq as they involve species-specific reagents. However, the recovery of motif signals for BCD, GT, HB and KR argues against this, as does the enrichment of the Zelda binding motif in these regions in all species. In addition, ChIP-seq and mRNA-seq data were highly correlated within each species. Finally, the divergence of TF binding we observe is in line with most genome-wide comparisons of TF binding in other models [7-9,13,15,43], although the divergence was smaller in this study, consistent with other observations [11,14,44]. Thus it is unlikely that the incongruence between TF binding divergence and gene expression divergence stems entirely from overestimated binding divergence.

The differential evolution of TF binding may partly explain how fast evolving TF binding leads to conserved mRNA levels. We found that clustered and high peaks located outside coding regions are more constrained than isolated smaller peaks (Figure 5). The former, better conserved, are more likely to be functional and to regulate transcription, compared to the latter, suspected to represent a more

opportunistic binding [22,45]. Accordingly, TF binding is higher and better conserved when associated with zygotic genes than with maternal genes (Figure S13). This differential conservation of TF binding reconciles previous studies demonstrating signatures of selection on CTCF binding evolution [11], or neutral evolution of CEBPA or HNF4A binding [7], when Twist binding associated with important genes was shown to be highly conserved [44]

The limited divergence of mRNA levels is weakly explained by differences in TF binding (Figure 7F). Several hypotheses may explain this result: compensatory changes in TF binding may have evolved to maintain conserved mRNA levels, although we failed to detect such events at the genomic scale. ChIP signal is a weak predictor of transcriptional regulation since it only captures where a TF binds on a genome [46]. Our study illustrates the current limits of our understanding of the functional links between a transcription factor binding a locus on one hand, and a gene being transcribed on the other hand [47]. Taking into account other transcriptomic marks like insulator binding [48], or changes in chromatin state [49], should improve our predictive power of gene expression divergence. Overall, our results highlight the need for integrated approaches to study the evolution of gene expression, taking into account the transcriptional context of each gene.

## **METHODS**

### **Antibodies used for ChIP.**

We used rabbit polyclonal antibodies raised against the *D.melanogaster* versions of the key A-P regulators Bicoid (BCD), Hunchback (HB), Giant (GT) and Krüppel (KR) [22]. They were affinity purified either against the *D.melanogaster* version of the proteins (recognizing the largest set of epitopes), or against the most distantly related version from *D.virilis* (recognizing the most highly conserved epitopes). The different proteins are highly conserved throughout the *Drosophila* gender (69%, 87%, 76% and 70% aminoacid identity between *D.melanogaster* and *D.virilis* BCD, HB, GT and KR epitopes, respectively), allowing an excellent cross-reactivity of the serum against proteins from all species.

### **Embryo Collections for ChIP.**

We collected embryos spanning early cellularization process (end of stage 4 to mid stage 5), during which the regulatory events that initiate segmentation along the A-P axis take place. Because the developmental speeds and optimal growth temperatures vary between species, the collection conditions were optimized for each species (table S1). Embryos from all species were fixed and processed either for RNAseq, as described above, or for ChIP (except *D.yakuba*), according to the protocol described below. In addition, we collected single embryos for each species, at the end of cellularization, just prior to gastrulation, based on morphological criteria, homogenized the embryos in TRIzol (Life Technologies) and processed the samples for DNA and total RNA extraction, as described below.

### **In Vivo Formaldehyde Cross-Linking of *D.melanogaster*, *D.pseudoobscura* and *D.virilis* embryos, followed by chromatin immunoprecipitation.**

Embryos were collected as described above and fixed with formaldehyde. The chromatin was isolated through CsCl gradient ultracentrifugation as previously described [22].

The chromatin used for immunoprecipitation was fragmented through sonication using a Bioruptor to an average fragment size of 180bp. After sonication ChIP was carried out using affinity purified rabbit polyclonal antibodies directed against large parts of BCD, HB, GT or KR, and prepared as described above: antibodies affinity-purified against the *D.melanogaster* epitope [22] or against the *D.virilis* orthologous parts of the *D.melanogaster* epitopes (for BCD, HB and KR). Both sets of antibodies essentially give similar results (GT and HB giving the most reproducible results), although the *D.virilis*-specific antibodies were less efficient, as expected since they recognize the smaller set of conserved epitopes compared to the other ones (lower signal over noise ratio) (see Figure S2 for qualitative and quantitative comparison of replicates). Some samples from *D.melanogaster*, *D.pseudoobscura* and *D.virilis* (labeled with an asterisk in Figure S1) were pooled prior to immunoprecipitation, in



order to minimize variation due to sample handling and protocol adherence. Pooling had very limited effect on downstream read mapping because these species are distantly related, and only a handful of regions were excluded for these pooled samples. However, we used only reads that could be unambiguously assigned to a single species. To this end, reads were separately mapped to each genome sequence and the reads that mapped to several genomes were discarded. Reads that mapped better to a genome sequence in particular (with at least 2 less mismatches), were recovered in a second time.

The DNA libraries for sequencing were prepared from the ChIP reaction and from Input DNA following the Illumina protocol for preparing samples for ChIP sequencing of DNA. All library amplifications were carried out by 15 cycles of PCR. After the amplification step, we size-selected DNA fragments of 190-290 bp. Library quality, fragment size, and concentration were measured as described for mRNAseq. Libraries were sequenced on a GAllx or HiSeq 2000 Illumina sequencer.

### **mRNAseq library preparation.**

Individual embryos from late stage 5 were chosen for RNA extraction based on morphology (having completed cellularization), which allows sampling of homologous stages across species. Embryo sampling was performed as described in [37]. Libraries were then made from total RNA of 6 individuals of each species, for a total of 24 libraries, using the mRNA TruSeq kit from Illumina, following the manufacturer's instructions. Size distribution of the library fragments was checked on a Bioanalyzer (Agilent) using the high sensitivity kit, and library concentration was measured by QPCR using Kapa Biosystems PCR kits for Illumina sequencing libraries, according to the manufacturer's instruction. Libraries were sequenced in 2 lanes (12 libraries per lane) using an Illumina HiSeq sequencer. mRNA levels were later averaged over the 6 individuals from the same species.

In parallel, pools of embryos spanning the end of stage 4 to mid stage 5 were collected (see Table S1). After RNA extraction, samples were processed for library preparation, as described above and paired-end libraries were sequenced on an Illumina GAllx.

### **Reference annotation based transcript assembly and estimation of mRNA levels using *D.melanogaster*, *D.yakuba*, *D.pseudoobscura* and *D.virilis* mRNAseq data.**

Reads were mapped to the genomes using tophat [50].

Refined annotations of *D. yakuba*, *D. pseudoobscura* and *D. virilis* were based on the reference annotations (Flybase releases 1.3, 2.22 and 1.2 respectively) and mRNAseq data obtained from pools of embryos. The reference annotations and mapped reads were given as inputs to cufflinks (version 1.0.3) [36] with the options `-g` and `'--3-overhang-tolerance 300'`. Statistics on the new annotations are given

Table S3. Then mRNA levels from individual embryos were then estimated using cufflinks [51] and the above refined annotations (except *D. melanogaster* samples for which the reference annotation was used).

### **Mapping Sequenced Tags to Genomes**

We used the Apr. 2006 assembly (Flybase Release 5) of the *D. melanogaster* genome, the February 2006 assembly (Flybase release 2) of the *D.pseudoobscura* genome and the February 2006 assembly (Flybase release 1) of the *D.virilis* genome. *D.yakuba* as well as some of *D.melanogaster* dataset were previously published [14].

We trimmed all sequenced tags so that their average quality was above 30 and mapped the tags to the genomes using Bowtie v0.12.7 [23] with command-line options '-v 1 -m 1' for small reads (length below 35bp) and '-v 1 -m 3' for long reads (length above 70bp), thereby keeping only tags that mapped uniquely to the genome with at most one or three mismatch. For ChIP experiments done on pooled chromatin, reads were mapped to each species separately and reads that ambiguously mapped to several genomes were tossed. A very small quantity of mapped reads (less than 1%) were discarded, and the pooling had virtually no effect on read mapping.

### **Peak calling.**

ChIP data was parsed independently for each experiment using two separate peak callers. First, we used MACS (version 1.4) [26], with the following parameter "-g dm --off-auto --nomodel --pvalue=1e-2" and "--shiftsize=110 --mfold=10,10000 --slocal=2000 --llocal=20000" or "--shiftsize=60 --mfold=4,10000" depending on the length distribution for DNA fragment sizes prior to sequencing. We also called peaks using Grizzly Peak fitting program [24,25] with estimated DNA fragment length of 150 or 250.

We then intersected the two sets of peaks, and filtered out all peaks not supported by both methods. To account for low complexity peaks and possible PCR artifacts, we further removed peaks with negative correlation (<-0.1) among the Forward and Reverse reads, peaks where 60% of the reads mapped to less than 1% of the positions, and peaks whose height was less than three times the height of Input reads in the same locus.

We took as an initial dataset the union of all bound regions in the different replicates. Roughly similar numbers of peaks were identified in each species (table S2), except in the case of *D.pseudoobscura* that displayed less BCD, GT and KR peaks, potentially partly due to coverage differences on the Müller element E, fused to the X chromosome in the *Pseudoobscura* group [5,52].

## **Genome versions, whole-Genome Alignment and Orthology Comparisons**

We used Flybase annotations r5.39, r1.3, 2.22 and 1.2 for *D. melanogaster*, *D.yakuba*, *D.pseudoobscura* and *D.virilis*, respectively. To produce the 4-species whole genome alignment, we used a large-scale orthology mapping created by PECAN with the option to identify syntenic regions of the genomes. Each region was then aligned with pecan [28] with the default options.

## **Establishing the set of bound regions from a 4-way genome alignment**

We assigned orthology between regions using the following pair-wise rule: bound regions in 2 different species are considered orthologous if they share a non-null intersection on the alignment (see Figure 2). Regions were removed from the analysis based on quality check: (i) Clusters that displayed a genomic length variability above five folds between any two species were considered to have an unreliable alignment, (ii) clusters for which less than 80% of genomic length from any species could be unambiguously mapped on the sequence were considered to have too ambiguous occupancy values and (iii) sets of orthologous peaks comprised of a small peak (below 0.7) in only one of the seven replicates were considered to be too uncertain.

We adopted a conservative approach when comparing binding between species: we used a peak height threshold only to establish our orthologous sets of regions, but did use it when calling peaks. This means that a high threshold is used to call a set of regions, but a lower threshold is used to determine if binding is conserved.

For each set of bound regions, we assigned a quantitative value per species, corresponding to the average of the maximum occupancies between replicates in each species-specific ChIP.

## **Association between genes and regions bound by TFs.**

We associated TF binding and genes according to the following rule: we associated with each gene the module (defined as the union of regions bound by any of the 4 TFs) located within 10,000bp from the gene and bound by the biggest number of TF. So any gene was associated with a nearby regions bound by 0 (if no TF bound within 10,000bp from the gene) to all four TFs.

## **Orthology assignment between genes of different species**

Orthology assignment between genes was established based on the whole-genome alignment: genes were considered orthologous if the coordinates of their exons intersect more than 40% of their total length and if their orientation is the same (or unknown). Because this method is genome-alignment based, it takes into account both sequence similarity and synteny, thus favoring ortholog over paralog association. We removed from the analysis genes that showed orthology inconsistencies (e.g. genes in a species with several orthologs in another species).

## **Clustering of Bound Regions**

The peaks called for every ChIP were further clustered, based on their mapping to the four-way alignment of the genomes. Overlapping peaks for the same TF over two or more species were merged.

## **Reconstruction of ancestral values of TF binding and mRNA levels**

Ancestral values were reconstructed according to a Brownian motion model [29] using the R package *ape* (function “*ace*”). Under a Brownian motion model, continuous characters evolve randomly following a random walk. The values at the three internal nodes, including the root, as well as values of binding variance were obtained for each cluster.

Quantitative divergence along the tree was estimated from the changes occurring along the six branches of the tree. It was calculated as the differences between the seven nodes of the tree (four leaves and three internal nodes). This estimation explicitly takes phylogenetic inertia into account. We also made simpler pairwise comparisons between the four leaves, which do not take phylogenetic inertia into account, but do not make any assertion about phylogenetic processes.

Before comparing the variance of TF binding and mRNA levels according to a Brownian motion model (Figure 7C), we scaled the normalized values with their standard deviation. We used these scaled normalized values to estimate maximum likelihood parameters of the model.

## **Statistical analyses**

The occupancies of the different TFs in each cluster per species were computed as the maximum occupancy value within the cluster coordinates. Values were normalized (median).

The Principal component analyses were conducted on the concatenate of TF occupancies in the 4 species, so that the results from the different components would be comparable between species. Note that no axis seemed to be correlated with species-specific differences. As a control, we also performed a PCA on species-specific data and axes were very similar to the axes from the trans-species PCA, especially for the first axis. Correlation of TF occupancy values projected on PC1 from the global PCA or from species specific PCAs are above 0.99 for each species and are above 0.77 for the 3 other axes. The measurements of the contribution of each of the PC axes to within or between species TF variation were obtained from a multiple linear regression of TF binding on the different axes of the PCA.

We used multiple linear regression to model TF binding and TF binding divergence from associated factors such as TF-specific and Zelda motif enrichments, proximity to genes and the number of TFs binding this locus. mRNA levels were modeled from associated BCD, GT, HB and KR binding, and divergence of mRNA levels from divergence of associated TF binding.

## **Funding**

This work was funded by a Howard Hughes Medical Institute investigator award to MBE and by NIH grant HG002779 to MBE. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

MP was funded by a long-term post-doctoral fellowship awarded by the Human Frontier Science Program organization (HFSP) and by a Genentech Foundation fellowship. TK was funded by a Long-Term Post-Doctoral Fellowship by the European Molecular Biology Organization (EMBO), and is a member of the Israeli Center of Excellence (I-CORE) for Gene Regulation in Complex Human Diseases, and the Israeli Center of Excellence (I-CORE) for Chromatin and RNA in Gene Regulation. SL was funded by NIH/NIGMS grant K99/R00-GM098448.

## **REFERENCES**

1. King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188: 107–116.
2. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134: 25–36. doi:10.1016/j.cell.2008.06.030.
3. Ludwig MZ, Kreitman M (1995) Evolutionary dynamics of the enhancer region of even-skipped in *Drosophila*. *Mol Biol Evol* 12: 1002–1011.
4. Moses AM, Pollard DA, Nix DA, Iyer VN, Li X-Y, et al. (2006) Large-scale turnover of functional transcription factor binding sites in *Drosophila*. *PLoS Comput Biol* 2: e130. doi:10.1371/journal.pcbi.0020130.
5. Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky S, et al. (2005) Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome research* 15: 1–18. doi:10.1101/gr.3059305.
6. Balhoff JP, Wray GA (2005) Evolutionary analysis of the well characterized endo16 promoter reveals substantial variation within functional sites. *Proc Natl Acad Sci USA* 102: 8591–8596. doi:10.1073/pnas.0409638102.
7. Schmidt D, Wilson MD, Ballester B, Schwalie PC, Brown GD, et al. (2010) Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* 328: 1036–1040. doi:10.1126/science.1186176.
8. Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, et al. (2007) Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39: 730–732. doi:10.1038/ng2047.
9. Borneman AR, Gianoulis TA, Zhang ZD, Yu H, Rozowsky J, et al. (2007) Divergence of transcription factor binding sites across related yeast species. *Science* 317: 815–819. doi:10.1126/science.1140748.
10. Zeitlinger J, Zinzen RP, Stark A, Kellis M, Zhang H, et al. (2007) Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev* 21: 385–390. doi:10.1101/gad.1509607.
11. Ni X, Zhang YE, Nègre N, Chen S, Long M, et al. (2012) Adaptive evolution and the birth of CTCF binding sites in the *Drosophila* genome. *PLoS Biol* 10: e1001420. doi:10.1371/journal.pbio.1001420.

12. Zheng W, Zhao H, Mancera E, Steinmetz LM, Snyder M (2010) Genetic analysis of variation in transcription factor binding in yeast. *Nature* 464: 1187–1191. doi:10.1038/nature08934.
13. Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A, et al. (2010) Variation in transcription factor binding among humans. *Science* 328: 232–235. doi:10.1126/science.1183621.
14. Bradley RK, Li X-Y, Trapnell C, Davidson S, Pachter L, et al. (2010) Binding Site Turnover Produces Pervasive Quantitative Changes in Transcription Factor Binding between Closely Related *Drosophila* Species. *PLoS Biol* 8: e1000343. doi:10.1371/journal.pbio.1000343.
15. Wilson MD, Barbosa-Morais NL, Schmidt D, Conboy CM, Vanes L, et al. (2008) Species-specific transcription in mice carrying human chromosome 21. *Science* 322: 434–438. doi:10.1126/science.1160930.
16. Mcmanus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, et al. (2010) Regulatory divergence in *Drosophila* revealed by mRNA-seq. *Genome research* 20: 816–825. doi:10.1101/gr.102491.109.
17. Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, et al. (2011) The evolution of gene expression levels in mammalian organs. *Nature* 478: 343–348. doi:10.1038/nature10532.
18. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, et al. (2010) Gene expression divergence recapitulates the developmental hourglass model. *Nature* 468: 811–814. doi:10.1038/nature09634.
19. Russo CA, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of drosophilid species. *Mol Biol Evol* 12: 391–404.
20. Lin MF, Deoras AN, Rasmussen MD, Kellis M (2008) Performance and scalability of discriminative metrics for comparative gene identification in 12 *Drosophila* genomes. *PLoS Comput Biol* 4: e1000067. doi:10.1371/journal.pcbi.1000067.
21. Li X-Y, Thomas S, Sabo PJ, Eisen MB, Stamatoyannopoulos JA, et al. (2011) The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding. *Genome Biol* 12: R34. doi:10.1186/gb-2011-12-4-r34.
22. Li X-Y, MacArthur S, Bourgon R, Nix D, Pollard DA, et al. (2008) Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol* 6: e27. doi:10.1371/journal.pbio.0060027.
23. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-

efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25. doi:10.1186/gb-2009-10-3-r25.

24. Capaldi AP, Kaplan T, Liu Y, Habib N, Regev A, et al. (2008) Structure and function of a transcriptional network activated by the MAPK Hog1. *Nat Genet* 40: 1300–1306. doi:10.1038/ng.235.
25. Harrison MM, Li X-Y, Kaplan T, Botchan MR, Eisen MB (2011) Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS Genet* 7: e1002266. doi:10.1371/journal.pgen.1002266.
26. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9: R137. doi:10.1186/gb-2008-9-9-r137.
27. Dewey CN (2006) Whole-genome alignments and polytopes for comparative genomics.
28. Paten B, Herrero J, Beal K, Fitzgerald S, Birney E (2008) Enredo and Pecan: genome-wide mammalian consistency-based multiple alignment with paralogs. *Genome research* 18: 1814–1828. doi:10.1101/gr.076554.108.
29. Felsenstein J (1973) Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet* 25: 471–492.
30. Felsenstein J (1985) Phylogenies and the comparative method. *American Naturalist*.
31. Martins EP, Garland T Jr (1991) Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. *Evolution* 45: 534–557.
32. Mattila TM, Bokma F (2008) Extant mammal body masses suggest punctuated equilibrium. *Proc Biol Sci* 275: 2195–2199. doi:10.1098/rspb.2008.0354.
33. Chaix R, Somel M, Kreil DP, Khaitovich P, Lunter GA (2008) Evolution of primate gene expression: drift and corrective sweeps? *Genetics* 180: 1379–1389. doi:10.1534/genetics.108.089623.
34. Bedford T, Hartl DL (2009) Optimization of gene expression by natural selection. *Proc Natl Acad Sci USA* 106: 1133–1138. doi:10.1073/pnas.0812009106.
35. MacArthur S, Li X-Y, Li J, Brown JB, Chu HC, et al. (2009) Developmental roles of 21 *Drosophila* transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions.



Genome Biol 10: R80. doi:10.1186/gb-2009-10-7-r80.

36. Roberts A, Pimentel H, Trapnell C, Pachter L (2011) Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics*. doi:10.1093/bioinformatics/btr355.
37. Lott SE, villalta JE, Schroth GP, Luo S, Tonkin LA, et al. (2011) Noncanonical Compensation of Zygotic X Transcription in Early *Drosophila melanogaster* Development Revealed through Single-Embryo RNA-Seq. *PLoS Biol* 9: e1000590. doi:10.1371/journal.pbio.1000590.
38. He Q, Bardet AF, Patton B, Purvis J, Johnston J, et al. (2011) High conservation of transcription factor binding and evidence for combinatorial regulation across six *Drosophila* species. *Nat Genet*. doi:doi:10.1038/ng.808.
39. Fowlkes CC, Eckenrode KB, Bragdon MD, Meyer M, Wunderlich Z, et al. (2011) A conserved developmental patterning network produces quantitatively different output in multiple species of *Drosophila*. *PLoS Genet* 7: e1002346. doi:10.1371/journal.pgen.1002346.
40. Ludwig MZ, Palsson A, Alekseeva E, Bergman CM, Nathan J, et al. (2005) Functional evolution of a cis-regulatory module. *PLoS Biol* 3: e93. doi:10.1371/journal.pbio.0030093.
41. Swanson CI, Schwimmer DB, Barolo S (2011) Rapid evolutionary rewiring of a structurally constrained eye enhancer. *Curr Biol* 21: 1186–1196. doi:10.1016/j.cub.2011.05.056.
42. Hare EE, Peterson BK, Eisen MB (2008) A careful look at binding site reorganization in the even-skipped enhancers of *Drosophila* and sepsids. *PLoS Genet* 4: e1000268. doi:10.1371/journal.pgen.1000268.
43. Conboy CM, Spyrou C, Thorne NP, Wade EJ, Barbosa-Morais NL, et al. (2007) Cell cycle genes are the evolutionarily conserved targets of the E2F4 transcription factor. *PLoS ONE* 2: e1061. doi:10.1371/journal.pone.0001061.
44. He Q, Bardet AF, Patton B, Purvis J, Johnston J, et al. (2011) High conservation of transcription factor binding and evidence for combinatorial regulation across six *Drosophila* species. *Nat Genet* 43: 414–420. doi:10.1038/ng.808.
45. Fisher WW, Li JJ, Hammonds AS, Brown JB, Pfeiffer BD, et al. (2012) DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*. *Proc Natl Acad Sci USA* 109: 21330–21335. doi:10.1073/pnas.1209589110.
46. Pepke S, Wold B, Mortazavi A (2009) Computation for ChIP-seq and RNA-seq studies. *Nat Methods* 6: S22–S32. doi:10.1038/nmeth.1371.

47. Weirauch MT, Hughes TR (2010) Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. *Trends in Genetics* 26: 66–74. doi:10.1016/j.tig.2009.12.002.
48. Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, et al. (2010) A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* 6: e1000814. doi:10.1371/journal.pgen.1000814.
49. Filion GJ, van Bemmell JG, Braunschweig U, Talhout W, Kind J, et al. (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143: 212–224. doi:10.1016/j.cell.2010.09.009.
50. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105–1111. doi:10.1093/bioinformatics/btp120.
51. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28: 511–515. doi:10.1038/nbt.1621.
52. Schaeffer SW, Bhutkar A, McAllister BF, Matsuda M, Matzkin LM, et al. (2008) Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics* 179: 1601–1655. doi:10.1534/genetics.107.086074.
53. Kaplan T, Li X-Y, Sabo PJ, Thomas S, Stamatoyannopoulos JA, et al. (2011) Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding during early *Drosophila* development. *PLoS Genet* 7: e1001290. doi:10.1371/journal.pgen.1001290.

## **FIGURE CAPTIONS**

**Figure 1:** Phylogenetic tree of the *Drosophila* genus. The 4 species studied here (*D.melanogaster*, *D.yakuba*, *D.pseudoobscura*, *D.virilis*) are highlighted and illustrated with a picture of an adult as well a blastoderm embryo. Scales for adults (Nicolas Gompel, Flybase) and embryos are indicated. The three internal nodes of the (*D.melanogaster*, *D.yakuba*, *D.pseudoobscura*, *D.virilis*) tree are highlighted. Divergence times are indicated under the tree [19].

**Figure 2:** Comparison of binding profiles of BCD, GT, HB and KR at the *even-skipped* locus in the four species *D.melanogaster*, *D.yakuba*, *D.pseudoobscura* and *D.virilis*. An illustration of the two types of comparisons made in this study are highlighted in grey: trans-species comparison for each single TF (right) or trans-TFs comparisons (left).

**Figure 3:** Comparison of BCD, GT, HB and KR binding in *D.melanogaster*, *D.yakuba*, *D.pseudoobscura* and *D.virilis*. **A.** Pair-wise comparisons of BCD, GT, HB or KR binding between *D. melanogaster* and *D. pseudoobscura*. [OR All pair-wise comparisons of GT binding between *D.melanogaster*, *D.yakuba*, *D.pseudoobscura* and *D.virilis*. Closely related species display high binding conservation than more distantly related species, as confirmed by Spearman correlations.]. See Figure S3 for all pair-wise comparisons. **B.** Neighbor-joining trees based on pairwise distance matrices of TF occupancy at bound loci (Spearman's correlation coefficient). **C.** Distribution of the number of species in which TF was detected per cluster, from a species-specific peak, to a peak conserved in all 4 species.

**Figure 4:** TF-specific motif turnover drives TF binding divergence. **A.** Comparison of quantitative variation of BCD binding divergence vs. underlying sequence divergence. Binding divergence was measured by the variance in a Brownian motion model of BCD binding divergence, and sequence divergence was measured by the total length of a PhyML phylogenetic tree based on sequence. **B.** TF-specific motifs are enriched in under bound regions by any of the four TFs in all species. Overall TF-specific enrichment was calculated per CHIP. Summarized results for from 12, 4, 8 and 4 ChIPs are displayed for *D. melanogaster*, *D. yakuba*, *D. pseudoobscura* and *D. virilis*. **C.** Enrichment of BCD motifs in bound regions is quantitatively highly predictive of BCD binding. BCD binding was predicted from underlying BCD motifs only [53] **D.** Comparison of BCD divergence along any branch of the tree depending on divergence of BCD-specific motifs TAATCC. **E.** Values of BCD divergence (same as **D**) were partitioned into three categories, depending on predicted changes of BCD binding along a branch, based on BCD binding motif turnover (thresholds indicated by vertical lines in **D**). \*\*\*: p.value < 0.001. Similar plots for GT, HB and KR can be found in Figures S6 and S7.

**Figure 5:** BCD, GT, HB and KR binding is differentially conserved **A.** Comparison of qualitative conservation of TF binding in the different species. A conservation score, corresponding to the average number of species in which binding was detected (1-4), was calculated for each set of orthologous regions and ranked according to

ancestral mean, as estimated using a Brownian motion model. **B.** Comparison of quantitative variation of trans-species binding variation, represented by the variance in a Brownian motion model of TF binding evolution, and depending on ancestral mean. **C.** Average conservation score (1-4) depending on peak location in *D. melanogaster*. **D.** Average conservation score depending on the number of other A-P factor binding the same locus. To correct as much as possible for TF binding differences linked to different wiring sizes, clusters were binned into 10 bins, depending on the estimated ancestral values, and the conservation was estimated independently in each bin. The average conservation is displayed.

Figure 6: Zelda divergence may drive TF binding divergence. **A.** Zelda CAGGTAG motif is enriched in all ChIPs in all species. The enrichment variability between species is mostly due to differences in ChIPs qualities. **B/D** Principal Component Analysis of binding of all factors. PCA of **(B)** the binding strength in the 4 species together and **(D)** the relative change in binding strength along each branch of the tree across all peaks. Each row represents a factor, and each column is a principal component of the relevant data. The color represents the sign (red positive, green negative) and magnitude (color intensity) of each value in each principal component vector. In each case the sign of the first principal component is the same for all four factors, indicating that the dominant driver of both interspecies divergence and quantitative variation within single species is a coordinated change in binding strength of all factors. This effect explained 40% of the variation between species, and 58% of the variation within species. **C.** Zelda occupancy [25] and PC1 coordinates are highly correlated ( $r \sim 0.79$ ). **D.** Changes in PC1 along the branches of the tree correlate with changes in Zelda binding predicted by the enrichment of the Zelda motif. **F.** Reconstructed changes in PC1 values along each branch of the tree were partitioned in three categories, depending on the intensity and direction of predicted changes, based on differences in Zelda motifs. \*\*\* p.value < 0.001

Figure 7: mRNA levels are highly conserved despite high divergence of BCD, GT, HB and KR binding. **A.** Pairwise comparison of mRNA levels in *D. melanogaster*, *D. yakuba*, *D. pseudoobscura* and *D. virilis* blastoderm embryos. Maternal genes are highlighted in grey, zygotic genes are in orange, among which known A-P targets are highlighted in black. **B.** Neighbor-joining tree based on pair-wise distance matrices of mRNA levels (Spearman's correlation coefficient). **C.** Phylogenetic variance of mRNA levels is significantly lower than variance of BCD, GT, HB or KR binding (Wilcoxon test p-value <  $10^{-16}$ ), In order to compare variance, quantitative values were normalized by dividing each dataset by its standard deviation, on which parameters of the Brownian motion model were reestimated. **D.** Proportion of bound regions associated with maternal, zygotic or maternal-zygotic genes, depending on the number of TFs binding the region. **E.** mRNA levels of zygotic genes are well predicted by associated TF binding in all species. mRNA levels were predicted from a multiple linear regression of associated nearby TF binding. **F.** Changes of mRNA levels along each branch of the tree are modestly but significantly correlated with predicted changes based on quantitative changes of associated TF binding.

## SUPPLEMENTARY FIGURES

Figure S1: Summary table of the dataset presented in this study. 1 to 3 ChIPs were performed for each of the TFs BCD, GT, HB and KR in each species among *D.melanogaster*, *D.yakuba*, *D.pseudoobscura* and *D.virilis*. Two types of antibodies were used (red: produced and purified using *D.melanogaster* epitopes; grey: produced using *D.melanogaster* epitopes and purified using *D.virilis* epitopes). In addition, blastoderm embryos were processed for mRNAseq experiments. \*: experiments for which chromatin from *D. melanogaster*, *D. pseudoobscura* and *D. virilis* were pooled before ChIP.

Figure S2: Qualitative and quantitative comparison of TF occupancy between replicates in **(A)** *D.melanogaster* (three replicates per TF) and **(B)** *D.pseudoobscura* (two replicates per TF), which are the 2 species for which replicates are available). **A.** For clarity, occupancy on the z axis was color-coded.

Figure S3: Pairwise comparison of raw BCD, GT, HB and KR binding measurements between species.

Figure S4: Enrichment of TF-specific motifs under called peaks for each ChIP. A 1kb window centered on each peak summit was screened for the presence of TF-specific motifs. The enrichment over the window is displayed per position. The red line corresponds to the same analysis using a randomized PWM.

Figure S5: TF-specific motifs are predictive of TF binding. Binding intensity was predicted in each cluster and each species based only from the presence of TF-specific motifs<sub>[53]</sub>.

Figure S6: Overall sequence divergence is poorly correlated with binding divergence. The plots are similar to Figure 3A. Comparison of quantitative variation of BCD, GT, HB and KR binding divergence vs. underlying sequence divergence. Binding divergence was measured by the variance in a Brownian motion model of binding divergence, and sequence divergence was measured by the total length of a PhyML phylogenetic tree based on sequence.

Figure S7: Motif turnover is predictive of TF binding divergence. **A-D.** Binding divergence of BCD, GT, HB and KR along each branch of the tree is correlated with divergence of predicted TF binding, based only on TF-specific **E-H.** Same as **A-D.** Values were partitioned into three categories, depending on predicted changes of binding along a branch, based on binding motif turnover (thresholds indicated by vertical lines in **A-D.**). \*\*\*: p-value < 0.001.

Figure S8: Principal Component Analysis of binding of all factors in each species. Each row represents a factor, and each column is a principal component of the relevant data. The color represents the sign (red positive, green negative) and magnitude (color intensity) of each value in each principal component vector. In each case the sign of the first principal component is the same for all four factors, indicating that the dominant driver of both interspecies divergence and quantitative

variation within single species is a coordinated change in binding strength of all factors.

Figure S9: Zelda binding may drive BCD, GT, HB and KR binding in all four species. **A.** Zelda motif enrichment is highly predictive of Zelda binding in *D. melanogaster*. Zelda binding was predicted based only from the presence of TF-specific motifs [53] and compared to measured Zelda binding in *D. melanogaster* blastoderm embryos [25]. **B.** Zelda binding predicted from motif enrichment is highly correlated with TF binding coordinates projected on PC1 in all four species.

Figure S10: TF binding (**A**) and binding divergence are (**B**) better predicted by an integrative model of binding, rather than just motif enrichment. BCD, GT, HB and KR binding and binding divergence are well predicted by a multiple linear regression that takes into account motif enrichment, predicted Zelda binding, the nature of a nearby gene (if any) as well as the number of other TFs binding the same locus.

Figure S11: mRNA levels are better predicted by associated nearby binding for zygotic than for maternal genes. **A-C.** Comparison of mRNA levels depending on the number of TFs associated with the gene. **D-F.** Comparison of mRNA levels between measured values and values predicted only on associated nearby TF binding.

Figure S12: Divergence mRNA levels along the *Drosophila* tree are better predicted by associated divergence of nearby binding for zygotic than for maternal genes. **A-C.** Comparison of (**A-F**) branch-wise and (**G-I**) pairwise quantitative changes in mRNA levels depending on quantitative changes of associated TF binding. Changes in mRNA levels were predicted using a multiple linear regression. (**D-F**) Same as **A-C**. Values were partitioned into three categories, depending on predicted changes of binding along a branch.

Figure S13: TF binding associated with zygotic genes is higher and better conserved than TF binding associated with maternal genes. TF binding divergence (**A**) and ancestral TF binding (**B**) for clusters associated with zygotic or maternal genes were calculated as in Figure 6. Binding divergence and ancestral values were estimated using a Brownian motion model of TF binding divergence. P-values of mean comparison (Wilcoxon test) are displayed above each graph.

## TABLES AND SUPPLEMENTARY TABLES

Table S1: Collection conditions.

Species	Temperature	Collection	Incubation
<i>D.melanogaster</i>	25°C	1h	2h
<i>D.yakuba</i> [14]	25°C	1h	1h45
<i>D.pseudoobscura</i>	25°C	1h	2h15
<i>D.virilis</i>	20°C	2h45	4h45

Table S2: Number of called peaks per TF per species.

Species	BCD	GT	HB	KR
<i>D.melanogaster</i>	728	2544	2950	3321
<i>D.yakuba</i>	849	1989	2412	2864
<i>D.pseudoobscura</i>	474	2284	1476	816
<i>D.virilis</i>	926	1179	2060	1556

Table S3: Statistics on improvement of gene annotation using mRNAseq.

Species	# Bases covered	kept genes	Modified genes	New Genes	kept isoforms	Modified isoforms	new isoforms
<i>D.melanogaster</i>	31,202,700 (+2.4%)	12865	1496	168	23186	531	1917
<i>D.yakuba</i>	26,688,208 (+17.8%)	13739	2563	544	14824	2081	1793
<i>D.pseudoobscura</i>	27,999,754 (+19.3%)	13274	2811	706	14886	2247	2140
<i>D.virilis</i>	27,183,880 (+24.7%)	11321	3113	680	12725	2474	2341