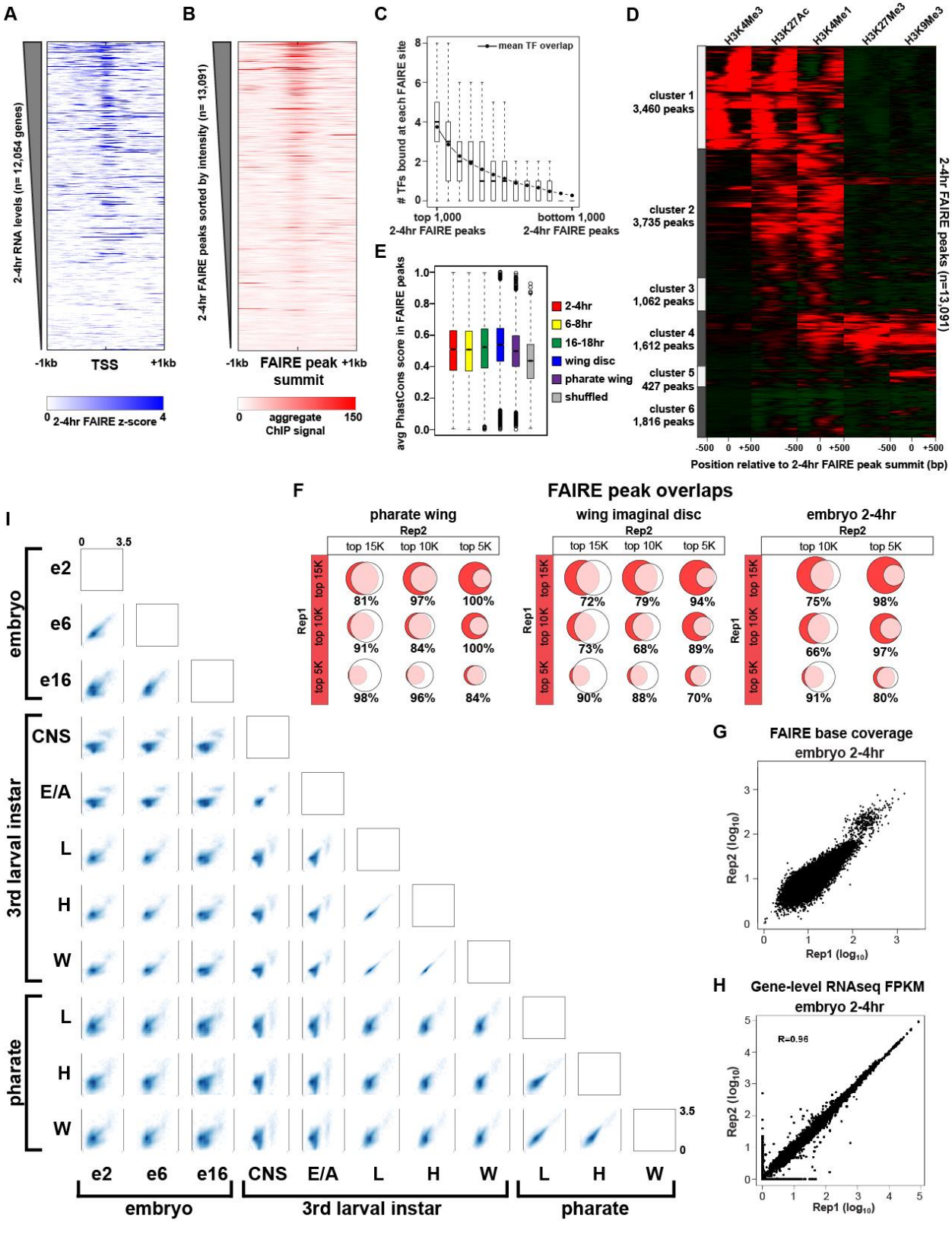

A
2-4hr RNA levels (n= 12,054 genes)
-1kb
TSS
+1kb
0 2-4hr FAIRE z-score 4
B
2-4hr FAIRE peaks sorted by intensity (n= 13,091)
-1kb
FAIRE peak summit
+1kb
0 aggregate ChIP signal 150
C
mean TF overlap
TFs bound at each FAIRE site
top 1,000 2-4hr FAIRE peaks
bottom 1,000 2-4hr FAIRE peaks
D
H3K4Me3
H3K27Ac
H3K4Me1
H3K27Me3
H3K9Me3
cluster 1 3,460 peaks
cluster 2 3,735 peaks
cluster 3 1,062 peaks
cluster 4 1,612 peaks
cluster 5 427 peaks
cluster 6 1,816 peaks
2-4hr FAIRE peaks (n=13,091)
-500 0 +500
Position relative to 2-4hr FAIRE peak summit (bp)
E
avg PhastCons score in FAIRE peaks
2-4hr
6-8hr
16-18hr
wing disc
pharate wing
shuffled
F
FAIRE peak overlaps
pharate wing
wing imaginal disc
embryo 2-4hr
Rep2
Rep1
top 15K
top 10K
top 5K
81% 97% 100%
91% 84% 100%
98% 96% 84%
72% 79% 94%
73% 68% 89%
90% 88% 70%
75% 98%
66% 97%
91% 80%
G
FAIRE base coverage
embryo 2-4hr
Rep2 (log10)
Rep1 (log10)
H
Gene-level RNAseq FPKM
embryo 2-4hr
R=0.96
Rep2 (log10)
Rep1 (log10)
I
embryo
e2
e6
e16
3rd larval instar
CNS
E/A
L
H
W
pharate
L
H
W
0 3.5
e2 e6 e16 CNS E/A L H W L H W
embryo
3rd larval instar
pharate

Figure S1, related to Figure 1. (**A**) 2-4 hr FAIRE signal (10bp windows) surrounding transcription start sites (TSS) for all genes, ordered by RNA levels. (**B**) Heatmap of the aggregate blastoderm transcription factor ChIP signal (Bradley et al., 2010) (10bp windows) surrounding 2-4 hr FAIRE peak maxima. FAIRE peaks are ranked according to MACS q-value (Zhang et al., 2008). (**C**) Boxplots of the number of 2-3hr transcription factors bound at 2-4 hr FAIRE sites. FAIRE peaks are divided into bins of 1,000 peaks, ranked by MACS q-value. The box represents the inner quartile range (IQR), whiskers represent 1.5-times IQR, and the line represents the median. For clarity purposes, outliers are not plotted. Filled circles connected by the solid line represent the mean number of transcription factor peaks overlapped in a given bin. (**D**) Heap map of histone modification signals across all 2-4 hr FAIRE peaks. Each column plots the average histone modification signal in 10bp bins, spanning the region 500 bp up and down from each FAIRE peak summit. The y-axis represents all 2-4 hr FAIRE peaks, which have been hierarchically clustered. (**E**) Boxplots of the average PhastCons (Siepel et al., 2005) score in FAIRE peaks at different stages of development, as compared to peaks randomly distributed across the genome (shuffled). The box represents the IQR, whiskers represent 1.5-times IQR, and the line represents the median. Outliers are indicated by circles. **(F)** Venn diagrams of pairwise peak overlaps for top 5,000, top 10,000, and top 15,000 FAIRE peaks ranked by q-values determined by MACS (Zhang et al., 2008). (**G**) Scatterplot of the average FAIRE signal (base coverage in counts per million, $\log_{10}$) in 500-bp windows that overlap the union set of peaks for 2-4 hr replicates. Windows with exceptionally high FAIRE signal ($>10^2$) exclusively map to the 4^{th} chromosome, possibly due to the unique chromatin structure of this chromosome (Riddle et al., 2012). (**H**) Scatterplot of FPKM values for all *Drosophila* genes (obtained with Cufflinks (Trapnell et al., 2010)) for 2-4 hr replicates. Pearson correlation (R) is shown in the top left of each scatterplot. (**I**) Pairwise scatterplots of FAIRE signal for all samples included in the manuscript (data processed as in **G**).

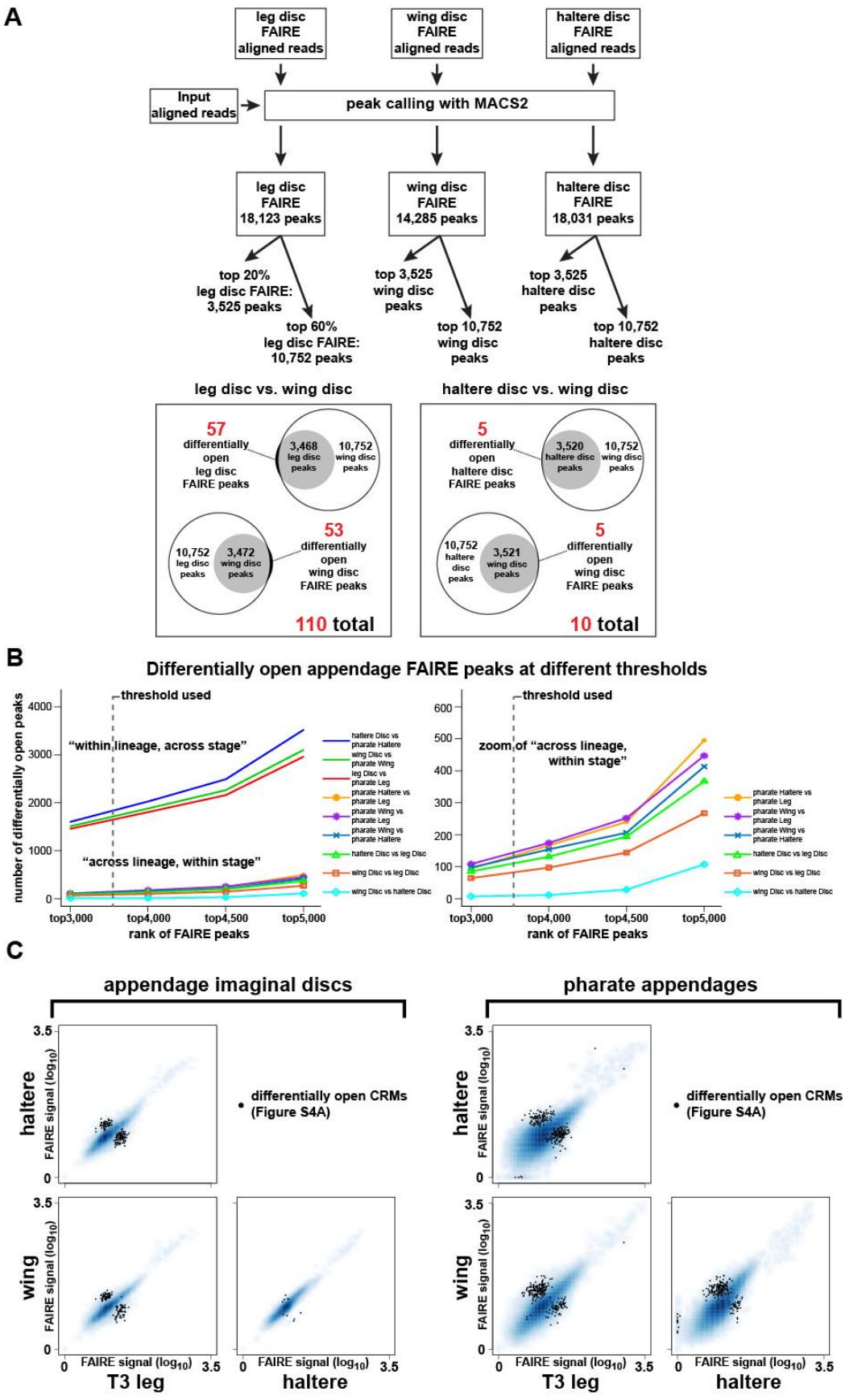
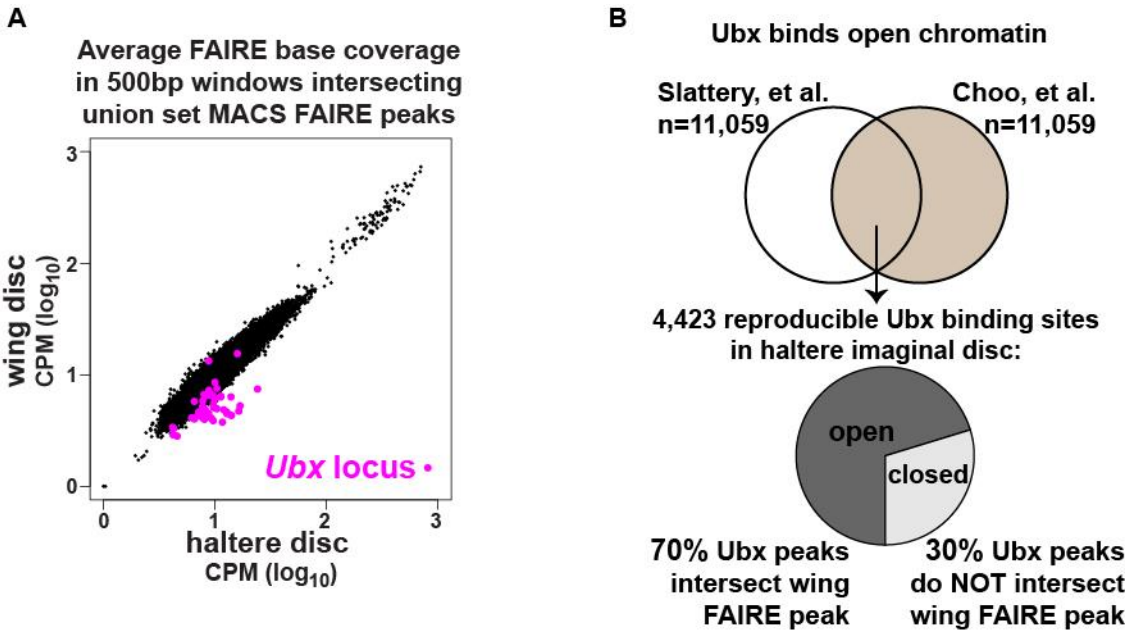

A
leg disc FAIRE aligned reads
wing disc FAIRE aligned reads
haltere disc FAIRE aligned reads
Input aligned reads
peak calling with MACS2
leg disc FAIRE 18,123 peaks
wing disc FAIRE 14,285 peaks
haltere disc FAIRE 18,031 peaks
top 20% leg disc FAIRE: 3,525 peaks
top 60% leg disc FAIRE: 10,752 peaks
top 3,525 wing disc peaks
top 10,752 wing disc peaks
top 3,525 haltere disc peaks
top 10,752 haltere disc peaks
leg disc vs. wing disc
57 differentially open leg disc FAIRE peaks
3,468 leg disc peaks
10,752 wing disc peaks
10,752 leg disc peaks
3,472 wing disc peaks
53 differentially open wing disc FAIRE peaks
110 total
haltere disc vs. wing disc
5 differentially open haltere disc FAIRE peaks
3,520 haltere disc peaks
10,752 wing disc peaks
10,752 haltere disc peaks
3,521 wing disc peaks
5 differentially open wing disc FAIRE peaks
10 total
B
Differentially open appendage FAIRE peaks at different thresholds
threshold used
"within lineage, across stage"
"across lineage, within stage"
number of differentially open peaks
rank of FAIRE peaks
top3,000
top4,000
top4,500
top5,000
zoom of "across lineage, within stage"
C
appendage imaginal discs
pharate appendages
haltere
wing
FAIRE signal (log10)
differentially open CRMs (Figure S4A)
T3 leg
haltere

Figure S2, related to Figure 3. Defining differentially open chromatin. (**A**) FAIRE-seq and Input reads were aligned to the *Drosophila* dm3 reference genome with Bowtie (Langmead et al., 2009). Aligned reads were used for peak calling with MACS (Zhang et al., 2008), using Input reads as the control dataset. Peaks were ranked by q-value. The top 20% (3,525 peaks) and 60% (10,752 peaks) were selected from the leg disc dataset. An equivalent number of peaks were selected from the wing and haltere disc datasets. We defined a peak as differentially open if it was within the top 20% of FAIRE peaks from the first sample, and did not intersect with a peak in the top 60% from the second sample. By these criteria, there are 110 differentially open peaks between leg discs and wing discs, and 10 differentially open peaks between wing discs and halter discs. See **Experimental Procedures** for more details. (**B**) Line graphs displaying the number of differentially open chromatin regions across a range of indicated thresholds. The vertical dotted line marks the threshold (see **A**) used in the paper. (**C**) Scatterplots of the average FAIRE signal (base coverage in counts per million, $\log_{10}$) in 500-bp windows that intersect with the union set of peaks for each sample. Each black point represents the average FAIRE signal in a FAIRE peak that was called differentially accessible between the two plotted samples.



C

Ubx sites at Ubx-responsive genes are more likely to be open

Source of Ubx-responsive gene list	Ubx peak category	Ubx peaks at Ubx-responsive genes	Ubx peaks at Ubx non-responsive genes	p-value*
Hersh, et al, 2007	OPEN CLOSED	458 128	2,648 1,189	4.9×10^{-6}
Pavlopoulos, et al, 2011	OPEN CLOSED	57 11	3,088 1,320	1.5×10^{-2}

*Fisher's Exact test

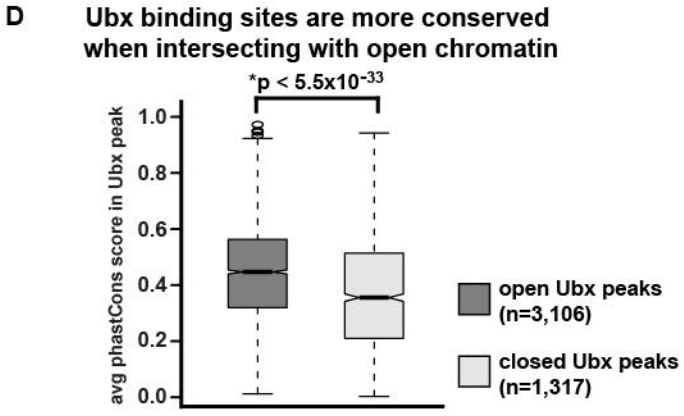


Figure S3, related to Figure 3. Open chromatin at the *Ubx* locus, and at sites of Ubx binding. (**A**) Scatterplot comparing FAIRE signal (base coverage in counts per million, $\log_{10}$) between haltere and wing discs. Each point represents the average FAIRE signal in 500-bp windows for all windows intersecting the union set of FAIRE peaks from the two plotted samples. Windows originating from the *Ubx* locus are designated in magenta. Many of the *Ubx*

windows are more accessible in haltere discs. (**B**) Venn diagram depicting the overlap between Ubx binding sites in the haltere disc from two independent ChIP-chip experiments (Choo et al., 2011; Slattery et al., 2011). The MAT FDR5 peaks were used from (Slattery et al., 2011), and an equal number of peaks were used from (Choo et al., 2011) (called using ChIPOTle (Buck et al., 2005), and ranked by p-value). The 4,423 reproducible Ubx binding sites were used for subsequent analysis. 70% of reproducible Ubx haltere peaks lie within an open chromatin region in wing discs (and in haltere discs). (**C**) A greater proportion of reproducible Ubx peaks at *Ubx*-responsive genes reside within open chromatin sites. (**D**) Boxplots of the average PhastCons (Siepel et al., 2005) scores for the DNA sequences located within reproducible Ubx peaks, categorized according to their intersection with wing FAIRE peaks. The box represents the inner quartile range (IQR), the whiskers represent 1.5-times IQR, and the line represents the median. Outliers are plotted as circles. Ubx peaks located within open chromatin sites are significantly more conserved (Student's t-test).

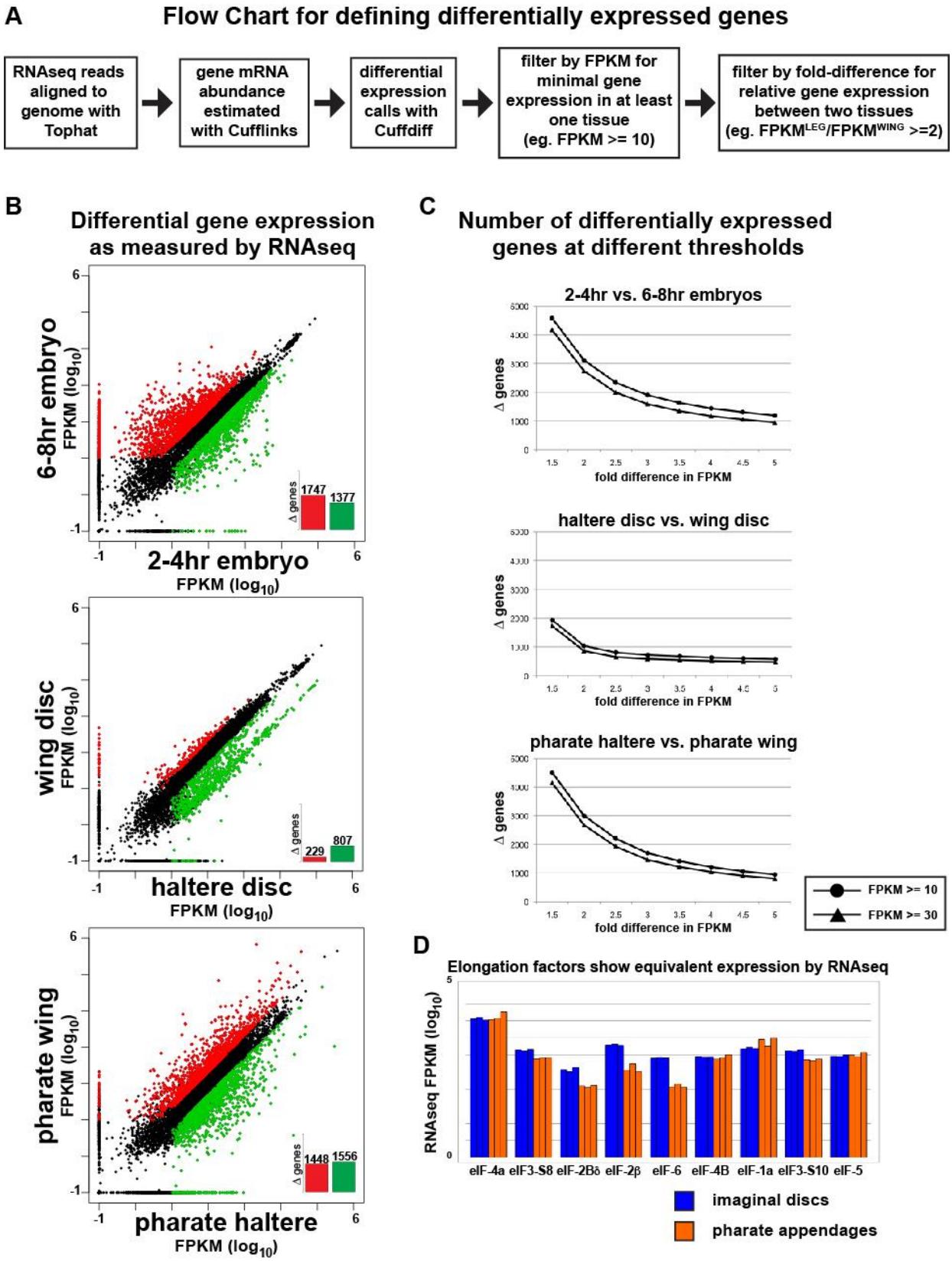
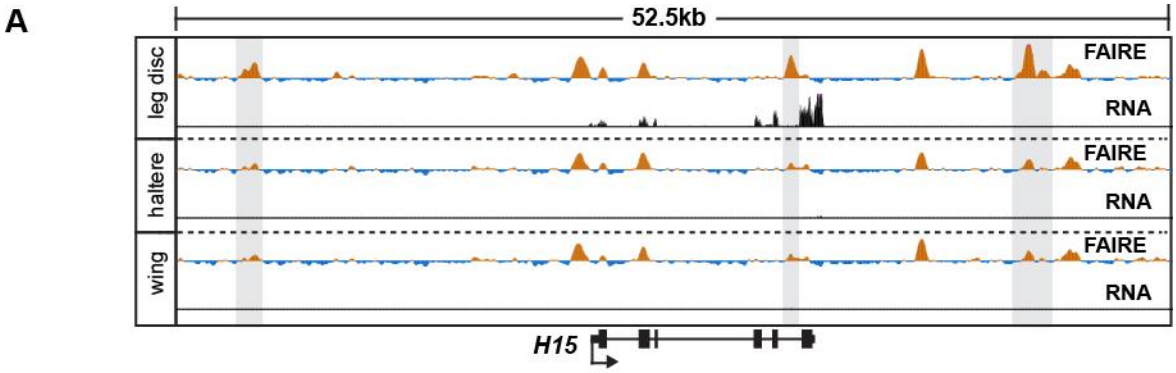
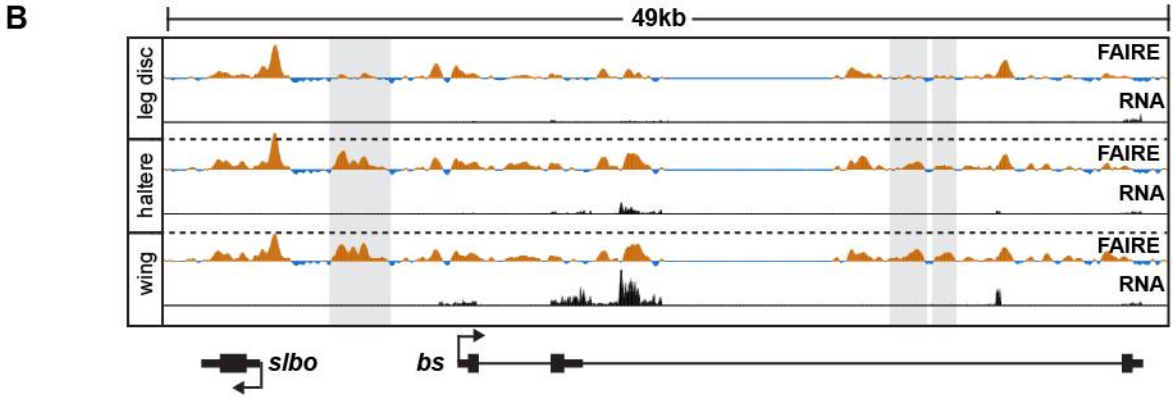
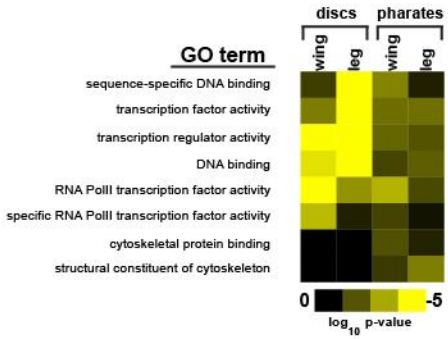

A
Flow Chart for defining differentially expressed genes
RNAseq reads aligned to genome with Tophat
gene mRNA abundance estimated with Cufflinks
differential expression calls with Cuffdiff
filter by FPKM for minimal gene expression in at least one tissue (eg. FPKM >= 10)
filter by fold-difference for relative gene expression between two tissues (eg. FPKM^LEG/FPKM^WING >=2)
B
Differential gene expression as measured by RNAseq
6-8hr embryo FPKM (log10)
2-4hr embryo FPKM (log10)
Δ genes
1747
1377
wing disc FPKM (log10)
haltere disc FPKM (log10)
229
807
pharate wing FPKM (log10)
pharate haltere FPKM (log10)
1448
1556
C
Number of differentially expressed genes at different thresholds
2-4hr vs. 6-8hr embryos
haltere disc vs. wing disc
pharate haltere vs. pharate wing
Δ genes
fold difference in FPKM
FPKM >= 10
FPKM >= 30
D
Elongation factors show equivalent expression by RNAseq
RNAseq FPKM (log10)
eIF-4a
eIF3-S8
eIF-2Bδ
eIF-2β
eIF-6
eIF-4B
eIF-1a
eIF3-S10
eIF-5
imaginal discs
pharate appendages

Figure S4, related to Figure 3. Differential gene expression despite similar open chromatin profiles. (**A**) Flow chart outlining the steps taken to define differentially expressed genes. (**B**) Scatterplots of gene FPKM values from RNA-seq of embryos, imaginal discs, and pharate appendages show different gene expression profiles between different samples, including those of the same developmental stage. Colored points designate genes that were called differentially expressed by Cuffdiff, have a minimum gene expression (FPKM) value greater than or equal to 10 in at least one sample, and have a fold-difference greater than or equal to 2 between samples. The number of differentially expressed genes is depicted by the inset bar graphs. (**C**) Plots of the number of genes differentially expressed between samples at different minimum expression (FPKM) and fold-difference thresholds. (**D**) Histograms of gene FPKM values of selected translational elongation factors, which we presume to be equivalently expressed between samples, demonstrate RNA-seq quantification is accurate. FPKM values were obtained with Cufflinks.





C

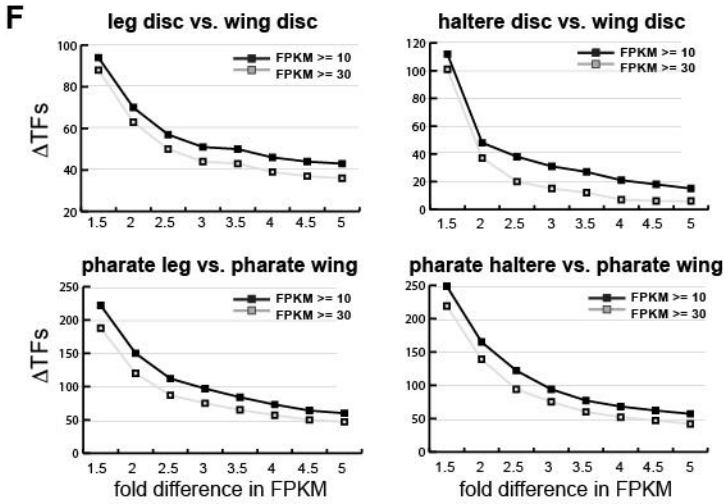


D

samples	#ΔCRMs	#ΔCRMs @TFs	p-value
leg disc vs. wing disc	110	42	$1.995e^{-5}$
pharate leg vs. pharate wing	220	38	$1.000e^{1}$

E

samples	#genes ΔCRMs	#TFs ΔCRMs	p-value
leg disc vs. wing disc	89	31	$1.919e^{-6}$
pharate leg vs. pharate wing	200	31	0.04779



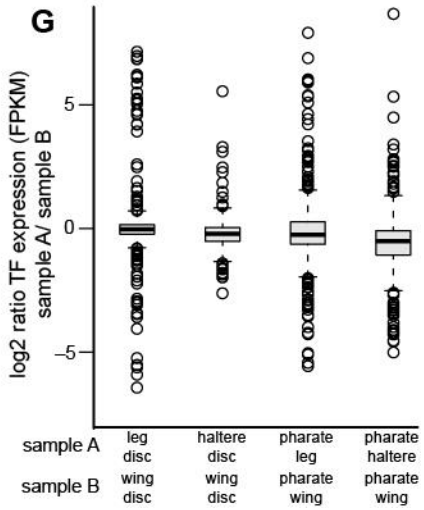


Figure S5, related to Figure 4. Transcription factors have differentially open chromatin, and are differentially expressed between the developing appendages. (**A**) Browser representations of FAIRE and RNA signals in imaginal discs at the *H15* (**A**) and *blistered* (*bs*) loci (**B**), key transcription factors involved in leg (H15) and wing/haltere (Bs) development. (**C**) Heat map for the top five Gene Ontology (GO) terms in Molecular Function from the DAVID (Huang da et al., 2009) Functional Annotation Chart for the nearest expressed gene to differentially open FAIRE peaks in wing versus leg imaginal discs and wing versus leg pharate appendages. (**D**) Table showing the total number of differentially open FAIRE peaks (ΔCRMs), the number of ΔCRMs whose closest gene is a transcription factor (ΔCRMs @TF), and the p-value for Fisher's exact test examining whether ΔCRMs are more likely to be located at a TF than similarly-ranked FAIRE peaks that are shared between wings and legs. (**E**) Table showing the number of expressed genes nearest to ΔCRMs, and of those genes, the number that are TFs. Also shown is the p-value for Fisher's exact test examining whether genes with ΔCRMs are more likely to be TFs than genes with similarly-ranked FAIRE peaks that are shared between wings and legs. The likelihood for ΔCRMs to be located at TFs is greater in imaginal discs than in pharate appendages. (**F**) Plots of the number of transcription factors (TFs) that are differentially expressed between the leg, wing, and haltere imaginal discs and the pharate appendages at different minimum expression (FPKM) and fold-difference thresholds. The list of putative transcription factors was obtained from www.FlyTF.org. (**G**) Boxplots of the ratio in expression (FPKM) for transcription factors expressed (FPKM >= 30) in the developing appendages. The box represents the inner quartile range (IQR), the whiskers represent 1.5-times IQR, and the line represents the median. Outliers are plotted as circles. Although many key transcription factors are differentially expressed as shown in **F**, most transcription factors are expressed at similar levels between appendages at a given developmental stage.

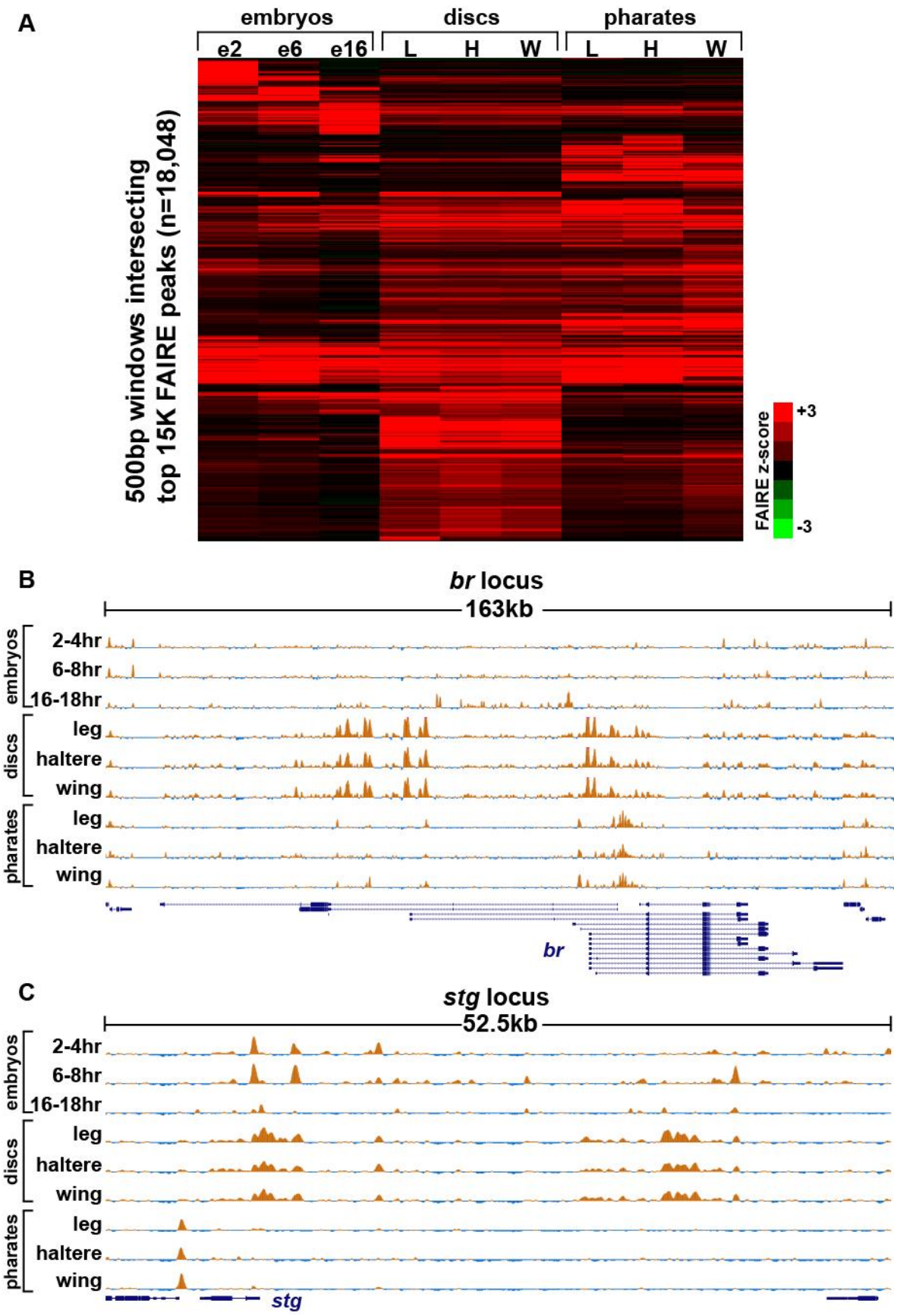
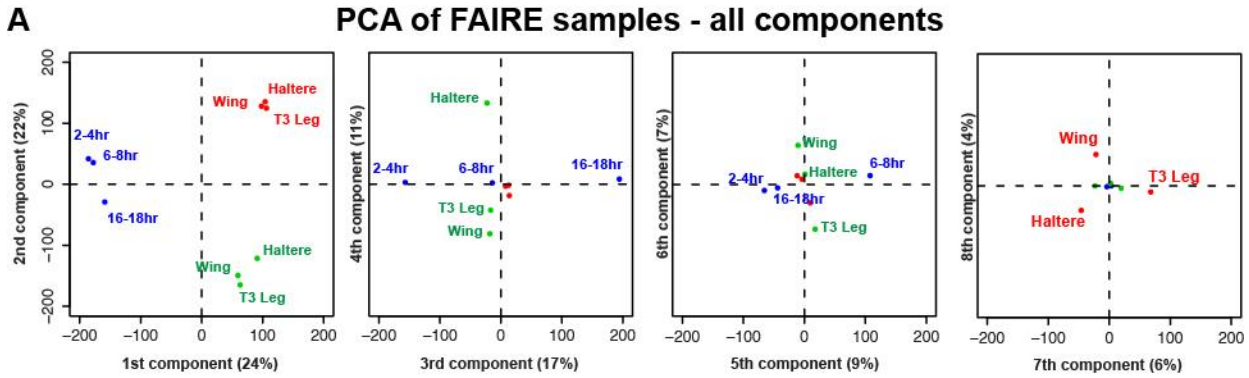
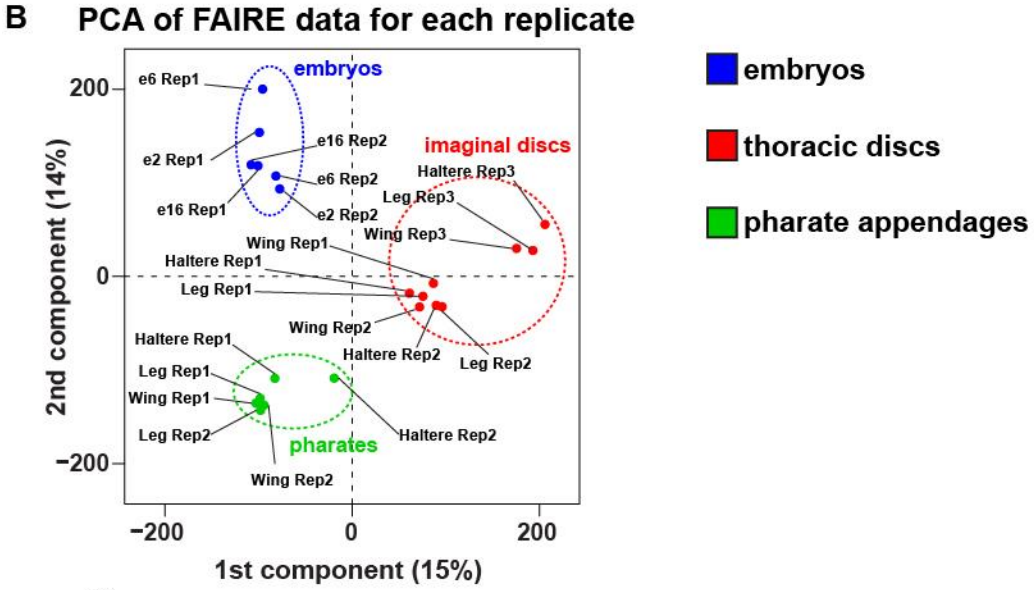
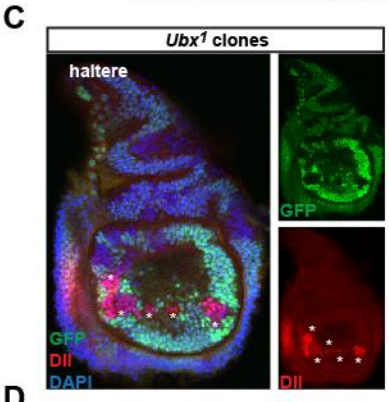
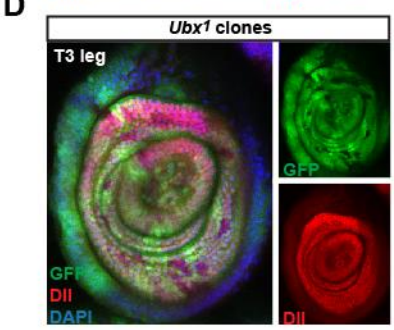

A
embryos
discs
pharates
e2
e6
e16
L
H
W
L
H
W
500bp windows intersecting
top 15K FAIRE peaks (n=18,048)
FAIRE z-score
+3
-3
B
br locus
163kb
embryos
2-4hr
6-8hr
16-18hr
discs
leg
haltere
wing
pharates
leg
haltere
wing
br
C
stg locus
52.5kb
embryos
2-4hr
6-8hr
16-18hr
discs
leg
haltere
wing
pharates
leg
haltere
wing
stg

Figure S6, related to Figure 5. Open chromatin profiles change over time. (**A**) Hierarchical clustering of FAIRE data from the three embryo stages, the three appendage imaginal discs, and the three pharate appendages. Open chromatin profiles are dynamic between stages; however, they are very similar to each other within a given developmental stage. FAIRE signal is plotted for 500-bp windows that overlap (by at least 90%) a region from the union set of FAIRE peaks for the nine samples. While some of the differences in open chromatin profiles between the imaginal disc and pharate appendage samples may be ascribed to loss of body wall tissue in the dissected pharate appendage samples, these data demonstrate the presence of many novel open chromatin regions in the pharate appendage samples that are not open in the imaginal discs, supporting the dynamic change in open chromatin over time. (**B**) Browser representation of FAIRE signal from all three embryonic stages, the three imaginal discs, and the three pharate appendages at the *broad* (*br*) and (**C**) *string* (*stg*) loci. Open chromatin profiles are dynamic between stages, but are very similar within a given developmental stage.


A
PCA of FAIRE samples - all components
2nd component (22%)
1st component (24%)
4th component (11%)
3rd component (17%)
6th component (7%)
5th component (9%)
8th component (4%)
7th component (6%)
2-4hr
6-8hr
16-18hr
Haltere
Wing
T3 Leg


B
PCA of FAIRE data for each replicate
embryos
imaginal discs
pharates
2nd component (14%)
1st component (15%)
e6 Rep1
e2 Rep1
e16 Rep1
e16 Rep2
e6 Rep2
e2 Rep2
Haltere Rep3
Leg Rep3
Wing Rep3
Wing Rep1
Haltere Rep1
Leg Rep1
Wing Rep2
Haltere Rep2
Leg Rep2
Haltere Rep1
Leg Rep1
Wing Rep1
Leg Rep2
Haltere Rep2
Wing Rep2
embryos
thoracic discs
pharate appendages


C
Ubx1 clones
haltere
GFP
Dll
DAPI


D
Ubx1 clones
T3 leg
GFP
Dll
DAPI

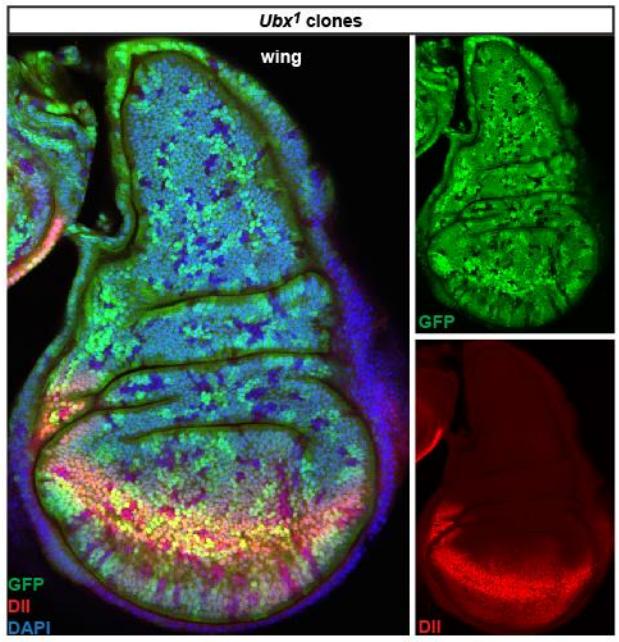

E
Ubx1 clones
wing
GFP
Dll
DAPI

Figure S7, related to Figure 7. (A, B). Greater similarity by developmental stage than by lineage. (**A**) Analysis of individual replicates. PCA scores of the first two components separate FAIRE samples into groups according to developmental stage, rather than by lineage. Several subsequent components, accounting for smaller percentages of the variance present within the data, stratify the samples by developmental stage. The percentage of the total variance represented by each component is shown in parentheses. (**B**) PCA scores of individual replicates again separates FAIRE samples into groups by developmental stage rather than by lineage, despite the fact that each replicate was performed on different days with different sample collections. The three imaginal disc points located to the right of the remaining points correspond to the third imaginal disc FAIRE replicate. (**C, D, E**) ***Ubx* represses *Dll* in the pouch of haltere imaginal discs**. (**C**) *Ubx* null clones, generated by mitotic recombination and marked by the absence of GFP, express *Dll* (red) along the dorsal-ventral axis of the haltere. *Dll* is normally expressed in the analogous population of cells in wing imaginal discs, demonstrating that *Ubx* represses *Dll* expression along the dorsal-ventral axis in the haltere. (**D**) Loss of *Ubx* function has no effect on *Dll* expression in T3 leg discs, where *Ubx* is expressed, but *Dll* is controlled by a different set of DNA regulatory elements (Estella et al., 2008; McKay et al., 2009). (**E**) Loss of *Ubx* function has no effect on *Dll* expression in wing discs, where *Ubx* is not expressed in wild-type animals. See **Experimental Procedures** for experimental details.

construct	gene	chr	left	right	length	loc.	distance to TSS	FAIRE timing	CRM activity?	CRM timing	CRM tissue
HB01	*hb*	3R	4506238	4506842	604	3'	17kb	2-8hAEL	yes	4hAEL	neuroblasts
HB02	*hb*	3R	4507462	4508512	1050	3'	15kb	2-8hAEL	yes	5hAEL	tracheal/mesoderm, CNS midline
HB03	*hb*	3R	4509972	4510848	876	3'	12.5kb	6-8hAEL	yes	5hAEL	tracheal/meso-derm
HB04	*hb*	3R	4512727	4513672	945	3'	10kb	2-18hAEL	yes	6hAEL	CNS, mesoderm
HB05	*hb*	3R	4515175	4516425	1250	3'	7kb	2-8hAEL	yes	5hAEL	CNS
HB06	*hb*	3R	4524744	4526029	1285	5'	1.2kb	2-4hAEL	yes	3hAEL	blastoderm
GSB01	*gsb*	2R	20952684	20953772	1088	3'	3kb	2-8hAEL	yes	5hAEL	ectoderm, mesoderm
GSBN01	*gsb-n*	2R	20938379	20938959	580	intron	1kb	2-4hAEL	yes	5hAEL	ectoderm, CNS
PRD01	*prd*	2L	12079764	12080433	669	3'	5.5kb	2-4hAEL	yes	5hAEL	ectoderm
STAT01	*Stat92E*	3R	16368439	16369389	950	intron	8.5kb	2-8hAEL, discs	yes	4hAEL	mesoderm, CNS
BRK01	*brk*	X	7193623	7195486	1863	5'	6.5kb	all stages	yes	5hAEL	ectoderm*, CNS
BRK02	*brk*	X	7200037	7201140	1103	5'	1kb	all stages	yes	6hAEL	amg, CNS, sg*
CHINMO_01	*chinmo*	2L	1636251	1639869	3618	5'	11kb	16-18hAEL	yes	10hAEL	CNS, sg*
CHINMO_02	*chinmo*	2L	1642980	1644174	1194	5'	7kb	6-18hAEL	yes	10hAEL	CNS, sg*
CHINMO_03	*chinmo*	2L	1653996	1654804	808	intron	3kb	16-18hAEL	yes	10hAEL	CNS, sg*
CHINMO_04	*chinmo*	2L	1656477	1658529	2052	intron	5kb	16-18hAEL, pharates	yes	10hAEL	CNS, sg*
EX01	*ex*	2L	441225	442874	1649	intron	10kb	embryos, discs	yes	8hAEL	ectoderm
DLL01	*Dll*	2R	20734855	20736112	1257	3'	32.5kb	leg discs	yes	3rd Instar	neurons
DLL02	*Dll*	2R	20749315	20750462	1147	3'	47kb	leg discs	yes	3rd Instar	proximal leg*
DLL03	*Dll*	2R	20755974	20757557	1583	3'	53.5kb	leg discs	yes	3rd Instar	distal leg
DLL04	*Dll*	2R	20724449	20725912	1463	3'	22kb	dorsal discs	yes	3rd Instar	wing, haltere#
VG01	*vg*	2R	8795879	8798121	2242	3'	24kb	dorsal discs	yes	3rd Instar	wing pouch
RHO_01	*rho*	3L	1450977	1453548	2571	5'	10kb	leg discs	yes	3rd Instar	distal leg
KN01	*kn*	2R	10669720	10671599	1879	intron	23kb	dorsal discs	yes	3rd Instar	wing, haltere#

Table S1, related to Figure 2. Enhancer activity chart. Table listing cloned open chromatin regions used for transgenic reporter analyses, including construct name, gene name, location of enhancer relative to transcription start site (TSS), distance of enhancer to TSS, timing of enhancer opening in FAIRE assays, whether construct exhibits reporter activity, timing of enhancer reporter activity, and tissue in which enhancer is active. * indicates the enhancer is active in cells in which the gene is not normally expressed; frequent ectopic enhancer activity in the salivary gland suggests a predisposition of the reporter vector or of the genomic DNA surrounding the *attP* site to be expressed in this tissue. # indicates differential enhancer activity

between wing and haltere. abbreviations: CNS, central nervous system; amg, anterior midgut; sg, salivary gland.

Table S2, related to Figure 4. Differentially accessible peaks in imaginal discs. Table listing the differentially open chromatin regions for each pairwise comparison between appendage imaginal discs, including each peak's chromosome (chrom_peak), left and right coordinates (left_peak, right_peak), name, MACS2 $-LOG_{10}$(qvalue), and the nearest gene.

Table S3, related to Figure 5. Differentially accessible peaks in pharate appendages. Table listing the differentially open chromatin regions for each pairwise comparison between pharate appendages, including each peak's chromosome (chrom_peak), left and right coordinates (left_peak, right_peak), name, MACS2 $-LOG_{10}$(qvalue), and the nearest gene.

Data File S1, related to Figure 2. Enhancer activity patterns. Confocal microscope images of cloned open chromatin regions used in transgenic reporter assays. Where indicated, *in situ* images were obtained from the BDGP *in situ* database (Tomancak et al., 2002; Tomancak et al., 2007).

Supplementary Experimental Procedures

Drosophila strains and sample collections

For embryo experiments, the OregonR strain was used (BDSC #25211). For imaginal disc and pharate appendage experiments, females of the reference genome strain *y;cn,bw,sp* (BDSC #2057) were used. Samples were grown and collected as previously reported (Agelopoulos et al., 2012; Estella et al., 2008). Embryo stages were confirmed with DAPI staining and light microscopy. Stage 13 pupae were identified by darkening of wing and bristle pigmentation.

RNA preparation

Staged samples were homogenized in 1mL of TRIzol (Invitrogen) in a Dounce homogenizer and frozen in liquid nitrogen. RNA was subsequently extracted following manufacturer's protocol, then treated with DNaseI and purified with Qiagen RNeasy columns. mRNA was purified with Sera-Mag oligo(dT) beads (ThermoScientific) and fragmented with heat and divalent cations (Ambion). First and second-strand cDNA synthesis was performed with Invitrogen's Superscript II kit.

FAIRE

Staged embryos and 3rd instar imaginal discs were crosslinked as previously described (Agelopoulos et al., 2012), (Estella et al., 2008), with the following exceptions. Crosslinking was performed for 10 minutes at room temperature. Fixed embryos were homogenized in Buffer A and filtered through 60uM Nitex membrane. Embryo nuclei were pelleted at 4,500Xg for 20 minutes. Nuclei/imaginal discs were resuspended in FAIRE Lysis Buffer, and subjected to four rounds of bead-beating with 2mm tungsten beads, as described previously (Simon et al., 2012). Chromatin was sonicated in a BioRuptor to a size range of 500bp-2kb. Soluble chromatin was recovered after centrifugation at 15,000Xg for 5-minutes. Pharate appendages were treated as described above, with the following exceptions. Crosslinking was performed for 20 minutes,

eight rounds of bead-beating were performed, and a Branson Sonifier was used for sonication. FAIRE was performed as previously described (Simon et al., 2012).

Preparation of High-Throughput Sequencing libraries

Library preparation was performed essentially as described (Simon et al., 2012), using 1-100ng ds-cDNA (RNA-seq), FAIRE-enriched DNA, or sonicated genomic DNA as Input material. Sequencing was performed on an Illumina GAII or Hi-Seq machine at the UNC High-Throughput Sequencing Facility. Libraries for Hi-Seq machines were prepared with internal index adapters and multiplexed.

Sequencing data analysis

FAIRE-seq data were processed essentially as previously described (Simon et al., 2012). In brief, reads were filtered with Tagdust (Lassmann et al., 2009), trimmed to a length of 36bp if necessary, and then aligned to the reference genome (dm3) using Bowtie (Langmead et al., 2009) (version 0.12.3) with up to four possible alignments, and two seed mismatches permitted. Since replicate data were highly correlated, reads from each replicate were pooled for further analysis. Reads resulting from PCR amplification errors and other artifacts were removed for downstream analyses (see below). FAIRE-seq signal files were generated by extending each read to a total of 110bp (the average size of FAIRE-enriched DNA fragments), and the number of reads overlapping each base in the genome was counted (base coverage). Since the different samples have different physiological properties, the signal to noise ratio varies between samples (for example, pharate appendages have more cuticle than imaginal discs). To minimize the impact of these differences when comparing between samples of different developmental stages, we converted FAIRE signal to z-scores: genomic DNA signal (normalized to read depth)

was subtracted from FAIRE signal (normalized to read depth) at each base, and z-scores were generated at each base by calculating the mean and standard deviation of the FAIRE base coverage signal for individual chromosome arms, subtracting the mean signal from the signal at each base on the given chromosome arm, and dividing by the standard deviation. FAIRE and DNaseI peaks were called with MACS2 (Zhang et al., 2008), using aligned read files as input, a shift size of 125bp, and a q-value cutoff of 1e-2. Hierarchical clustering (using average linkage) and principal component analysis was performed with Cluster 3.0 (de Hoon et al., 2004). Union sets of peaks were generated by concatenating individual sample peaks into a single file; overlapping windows were merged using BEDTools (Quinlan and Hall, 2010). For heat maps, the union set of peaks was intersected with FAIRE data that had been averaged into 500-bp bins across the genome, requiring an overlap of at least 90%. RNA-seq data were aligned to the reference genome (dm3) using TopHat (version 1.1.4), and assembled into transcripts with Cufflinks (Trapnell et al., 2009) (version 0.9.3), using the RefSeq genome annotation, upper-quartile normalization, and sequence bias correction. Differential gene expression calls were made with Cuffdiff (version 0.9.3), using the RefSeq genome annotation, upper-quartile normalization, and sequence bias correction, and additional filtering, as outlined in **Fig. S4**. The UCSC genome browser was used to visualize data (Kent et al., 2002) (http://genome.ucsc.edu).

Sequencing artifacts

In three instances, reads that we deemed to be technical artifacts were manually removed. The first instance involved a large read pileup at a single genomic locus on the X-chromosome (chrX 21834650-21835230) in multiple replicates across many samples. The cause of this artifact is likely due to a repetitive element. The second instance involved a square-shaped pileup of reads in the coding sequence at a single genomic locus (Iswi) that appeared in a single replicate. The cause of this artifact was unknown, but likely involved aberrant PCR amplification

(which occurs as part of the library preparation). The third instance involved five square-shaped read pileups in two samples from the same replicate. The cause of these artifacts was due to trace amounts of plasmid DNA from our reporter constructs contaminating the samples prior to the PCR amplification step of high-throughput sequencing library preparation. Plasmid DNA was also found in the raw reads. In all three instances, since the other replicates showed no signs of these signals, and the signals were recognizably distinct in shape and amplitude, we felt comfortable removing the associated reads from downstream analysis.

Enhancer cloning and immunofluorescence experiments

To test the sufficiency of individual open chromatin sites to control transcription, we cloned twenty-four open chromatin regions and placed them upstream of the yeast transcription factor GAL4 for transgenesis (**Table S1**). We chose regions that are differentially accessible in our three *Drosophila* embryo timepoints from developmentally important genes known to be expressed at these stages, including: *hunchback*, *gooseberry*, *gooseberry-neuro*, *paired*, *Stat92E*, *brinker*, *chronologically-inappropriate morphogenesis*, *expanded*, *Distalless*, *vestigial*, *rhomboid*, and *knot*. We used only FAIRE data, without consulting any other datasets (for example, ChIP, evolutionary conservation) to identify target regions for cloning. PCR primers were designed according to the boundaries of open chromatin regions, and amplified products were subcloned into pCR8/GW (Invitrogen), moved into the attB-containing germline transformation vector pθUGG or pMintgate (Jiang et al.), and injected into embryos containing the attP2 site at 68A4 (BestGene, Chino Hills, CA). Immunofluorescence experiments were performed as previously described (Estella et al., 2008). Antibodies to Hunchback were a gift from Chris Doe. Antibodies to Vestigial were a gift from Gary Struhl. Antibodies to Distalless were a gift from Richard Mann. Antibodies to Knot (Collier) were a gift from Michele Croatzier. The *ex-lacZ* strain was a gift from Allen Laughon. The *hs-flp; CyO/IF; Frt82B, Ubx*^{1}*/TM6B* line

was a gift from Richard Mann. *Ubx* mitotic clones were generated by heat-shock for 1-hour at 37C at 60-72hrs.

Supplementary References

Agelopoulos, M., McKay, D.J., and Mann, R.S. (2012). Developmental regulation of chromatin conformation by Hox proteins in Drosophila. Cell Rep *1*, 350-359.

Bradley, R.K., Li, X.Y., Trapnell, C., Davidson, S., Pachter, L., Chu, H.C., Tonkin, L.A., Biggin, M.D., and Eisen, M.B. (2010). Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related Drosophila species. PLoS Biol *8*, e1000343.

Buck, M.J., Nobel, A.B., and Lieb, J.D. (2005). ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. Genome Biol *6*, R97.

Choo, S.W., White, R., and Russell, S. (2011). Genome-wide analysis of the binding of the Hox protein Ultrabithorax and the Hox cofactor Homothorax in Drosophila. PLoS ONE *6*, e14778.

de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.

Estella, C., McKay, D.J., and Mann, R.S. (2008). Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into Distalless during Drosophila leg development. Dev Cell *14*, 86-96.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Jiang, L., Pearson, J.C., and Crews, S.T. (2010). Diverse modes of Drosophila tracheal fusion cell transcriptional regulation. Mech Dev *127*, 265-280.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lassmann, T., Hayashizaki, Y., and Daub, C.O. (2009). TagDust--a program to eliminate artifacts from next generation sequencing data. Bioinformatics *25*, 2839-2840.

McKay, D.J., Estella, C., and Mann, R.S. (2009). The origins of the Drosophila leg revealed by the cis-regulatory architecture of the Distalless gene. Development *136*, 61-71.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Riddle, N.C., Jung, Y.L., Gu, T., Alekseyenko, A.A., Asker, D., Gui, H., Kharchenko, P.V., Minoda, A., Plachetka, A., Schwartz, Y.B.*, et al.* (2012). Enrichment of HP1a on Drosophila chromosome 4 genes creates an alternate chromatin structure critical for regulation in this heterochromatic domain. PLoS Genet *8*, e1002954.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S.*, et al.* (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res *15*, 1034-1050.

Simon, J.M., Giresi, P.G., Davis, I.J., and Lieb, J.D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat Protoc *7*, 256-267.

Slattery, M., Ma, L., Negre, N., White, K.P., and Mann, R.S. (2011). Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in Drosophila. PLoS ONE *6*, e14686.

Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S.E.*, et al.* (2002). Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol *3*, RESEARCH0088.

Tomancak, P., Berman, B.P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S.E., and Rubin, G.M. (2007). Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biol *8*, R145.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.
Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.