Supplemental Material

ATAC-seq reveals regional differences in enhancer accessibility during the establishment of spatial coordinates in the *Drosophila* blastoderm

Marta Bozek, Roberto Cortini, Andrea Ennio Storti, Ulrich Unnerstall, Ulrike Gaul, Nicolas Gompel

Table legends

Supplemental Table S1. Expression constructs used for generation of transgenic strains. Name of the expression construct, name of the corresponding tagged domain, name of the basal promoter, name of the enhancer used as a driver of the nuclear tag (as in Segal et al. 2008), corresponding REDfly identifier of the enhancer (REDfly v.5.4.3, Gallo et al. 2011), reference to the original study that identified and characterized the enhancer, name of the enhancer's target gene, corresponding expression domain of the target gene.

Supplemental Table S2. Positions of tagged domains D1-D7 along 1-100% of the AP axis. 1%: anterior tip. 100%: posterior tip. 1: TRUE. 0: FALSE.

Supplemental Table S3. Summary of ATAC-seq samples. Name of the sample, corresponding tagged domain, name of the corresponding expression construct, corresponding homozygous transgenic line, scaling factor used for normalization of ATAC-seq signal: based on the total count of Tn5 transposase cuts, scaling factor used for normalization of ATAC-seq signal: based on the coverage of 1-100 bp ATAC-seq fragments.

Supplemental Table S4. Genomic coordinates of high-confidence ATAC-seq peaks. Chromosome name, start position, end position, peak ID, genomic annotation (as in Fig. 4A).

Supplemental Table S5. Genomic coordinates of the union of ATAC-seq peaks. Chromosome name, start position, end position, peak ID, genomic annotation (as in Fig. 4A).

Supplemental Table S6. List of differential high-confidence ATAC-seq peaks. Peak ID (as in Supplemental Table S4), genomic coordinates (chromosome name, start position, end position), genomic annotation (as in Fig. 4A), maximum reported $\log_2$ fold-change, summary of 28 pair-wise comparisons between individual domains and whole-embryo controls. 1: significant signal difference. 0: non-significant signal difference.

Supplemental Table S7. List of differential intervals from the union of ATAC-seq peaks. Peak ID (as in Supplemental Table S5), genomic coordinates (chromosome name, start position, end position), genomic annotation (as in Fig. 4A), summary of 28 pair-wise comparisons between individual domains and whole-embryo controls. 1: significant signal difference. 0: non-significant signal difference.

Supplemental Table S8. List of 88 AP enhancers. Name of the enhancer used in this study, genomic coordinates (chromosome name, start position, end position), REDfly identifier of the enhancer (REDfly v.5.4.3, Gallo et al. 2011), reference to the original study that identified and characterized the enhancer (source of the RNA *in situ* hybridization image used for estimation of the activity pattern), name of the enhancer's target gene, classification of the target gene according to the position in the AP gene regulatory network (gap, pair-rule, homeotic,

secondary downstream gene), Pearson's correlation coefficient of the linear regression, slope of the linear regression, corrected Pearson's correlation coefficient (for enhancers overlapping drivers of the nuclear tag), corrected slope of the linear regression (for enhancers overlapping drivers of the nuclear tag).

Supplemental Table S9. Positions of activity patterns of 88 AP enhancers along 1-100% of the AP axis. 1%: anterior tip. 100%: posterior tip. 1: TRUE. 0: FALSE.

References:

Berman BP, Pfeiffer BD, Laverty TR, Salzberg SL, Rubin GM, Eisen MB, Celniker SE. 2004. Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in Drosophila melanogaster and Drosophila pseudoobscura. *Genome Biol* **5**: R61.

Chen H, Xu Z, Mei C, Yu D, Small S. 2012. A system of repressor gradients spatially organizes the boundaries of Bicoid-dependent target genes. *Cell* **149**: 618–629.

Fisher WW, Li JJ, Hammonds AS, Brown JB, Pfeiffer BD, Weiszmann R, MacArthur S, Thomas S, Stamatoyannopoulos JA, Eisen MB, 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* **109**: 21330–21335.

Fujioka M, Emi-sarker Y, Yusibova GL, Goto T, Jaynes JB. 1999. Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers , and multi-stripe positioning by gap gene repressor gradients. *Development* **2538**: 2527–2538.

Gallo SM, Gerrard DT, Miner D, Simich M, Des Soye B, Bergman CM, Halfon MS. 2011. REDfly v3.0: Toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Res* **39**: 118–123.

Gao Q, Finkelstein R. 1998. Targeting gene expression to the head: the Drosophila orthodenticle gene is a direct target of the Bicoid morphogen. *Development* **125**: 4185–93.

Häder T, Wainwright D, Shandala T, Saint R, Taubert H, Brönner G, Jäckle H. 2000. Receptor tyrosine kinase signaling regulates different modes of Groucho-dependent control of Dorsal. *Curr Biol* **10**: 51–54.

Hoch M, Schröder C, Seifert E, Jäckle H. 1990. cis-acting control elements for Krüppel expression in the Drosophila embryo. *EMBO J* **9**: 2587–95.

Howard KR, Struhl G. 1990. Decoding positional information: regulation of the pair-rule gene hairy. *Development* **110**: 1223–31.

Jacob Y, Sather S, Martin JR, Ollo R. 1991. Analysis of Krüppel control elements reveals that localized expression results from the interaction of multiple subelements. *Proc Natl Acad Sci* **88**: 5912–5916.

Kazemian M, Blatti C, Richards A, McCutchan M, Wakabayashi-Ito N, Hammonds AS, Celniker SE, Kumar S, Wolfe SA, Brodsky MH, et al. 2010. Quantitative analysis of the Drosophila segmentation regulatory network using pattern generating potentials. *PLoS Biol* **8**: 51–52.

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

Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, Dickson BJ, Stark A. 2014. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature* **512**: 91–95.

La Rosée A, Häder T, Taubert H, Rivera-Pomar R, Jäckle H. 1997. Mechanism and Bicoid-dependent

control of hairy stripe 7 expression in the posterior region of the Drosophila embryo. *EMBO J*.

Margolis JS, Borowsky ML, Steingrímsson E, Shim CW, Lengyel JA, Posakony JW. 1995. Posterior stripe expression of hunchback is driven from two promoters by a common enhancer element. *Development* **121**: 3067–3077.

Ochoa-Espinosa A, Yucel G, Kaplan L, Pare A, Pura N, Oberstein A, Papatsenko D, Small S. 2005. The role of binding site cluster strength in Bicoid-dependent patterning in Drosophila. *Proc Natl Acad Sci U S A* **102**: 4960–4965.

Pankratz MJ, Busch M, Hoch M, Seifert E, Jäckle H. 1992. Spatial control of the gap gene knirps in the Drosophila embryo by posterior morphogen system. *Science (80-)* **255**: 986–989.

Perry MW, Boettiger AN, Levine M. 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. *Proc Natl Acad Sci* **108**: 13570–13575.

Riddihough G, Ish-Horowicz D. 1991. Individual stripe regulatory elements in the Drosophila hairy promoter respond to maternal, gap, and pair-rule genes. *Genes Dev* **5**: 840–854.

Rudolph KM, Liaw G-J, Daniel A, Green P, Courey AJ, Hartenstein V, Lengyel JA. 1997. Complex regulatory region mediating tailless expression in early embryonic patterning and brain development. *Development* **124**: 4297–4308.

Schröder C, Tautz D, Seifert E, Jäckle H. 1988. Differential regulation of the two transcripts from the Drosophila gap segmentation gene hunchback. *EMBO J* **7**: T2881-2887.

Schroeder MD, Greer C, Gaul U. 2011. How to make stripes: deciphering the transition from non-periodic to periodic patterns in Drosophila segmentation. *Development* **138**: 3067–3078.

Schroeder MD, Pearce M, Fak J, Fan H, Unnerstall U, Emberly E, Rajewsky N, Siggia ED, Gaul U. 2004. Transcriptional control in the segmentation gene network of Drosophila. *PLoS Biol* **2**: e271.

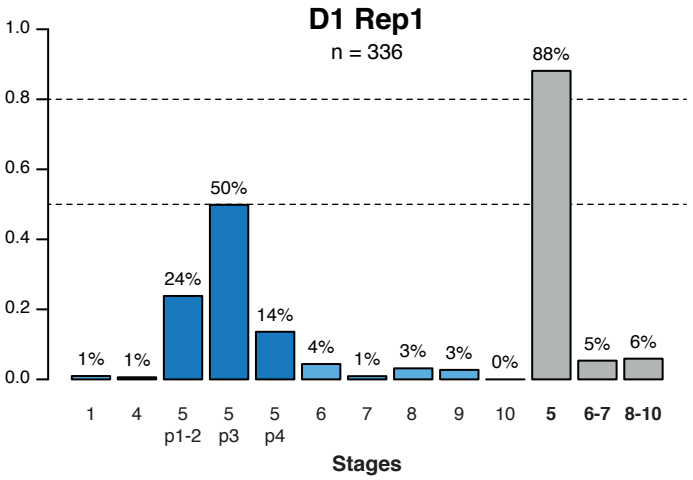
Segal E, Raveh-Sadka T, Schroeder M, Unnerstall U, Gaul U. 2008. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* **451**: 535–540.

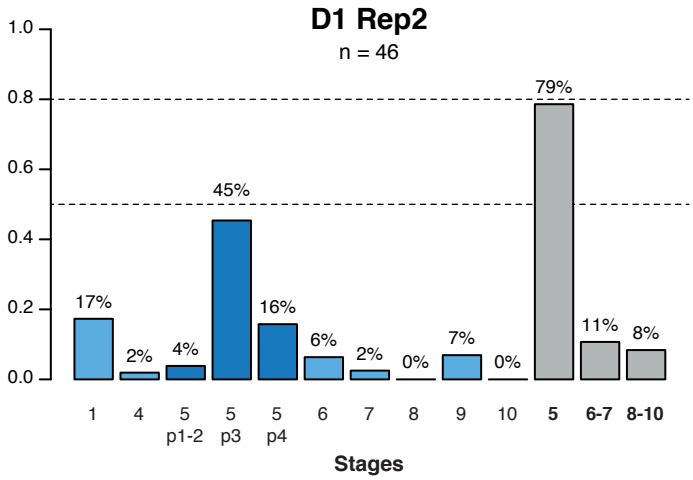
Small S, Blair A, Levine M. 1992. Regulation of even-skipped stripe 2 in the Drosophila embryo. *EMBO J* **11**: 4047–57.

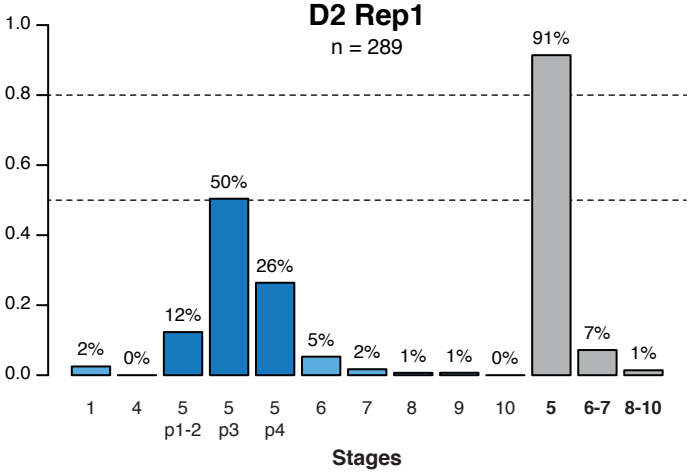
Small S, Blair A, Levine M. 1996. Regulation of two pair-rule stripes by a single enhancer in the Drosophila embryo. *Dev Biol* **175**: 314–324.

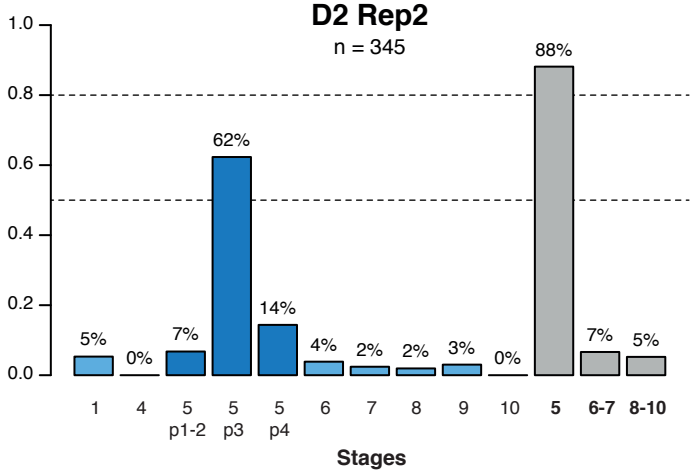
Wimmer EA, Simpson-Brose M, Cohen SM, Desplan C, Jäckle H. 1995. Trans- and cis-acting requirements for blastodermal expression of the head gap gene buttonhead. *Mech Dev* **53**: 235–245.

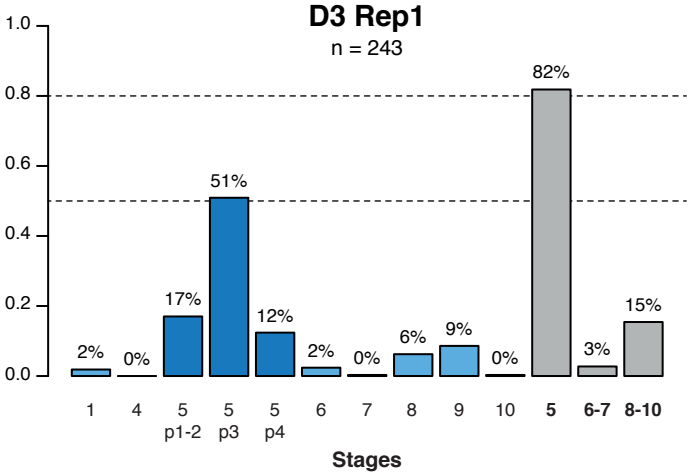
.

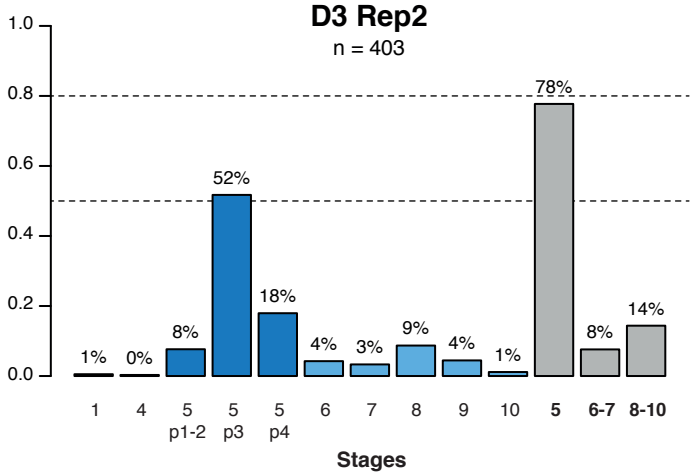

D1 Rep1
n = 336
1.0
0.8
0.6
0.4
0.2
0.0
1% 1% 24% 50% 14% 4% 1% 3% 3% 0% 88% 5% 6%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages

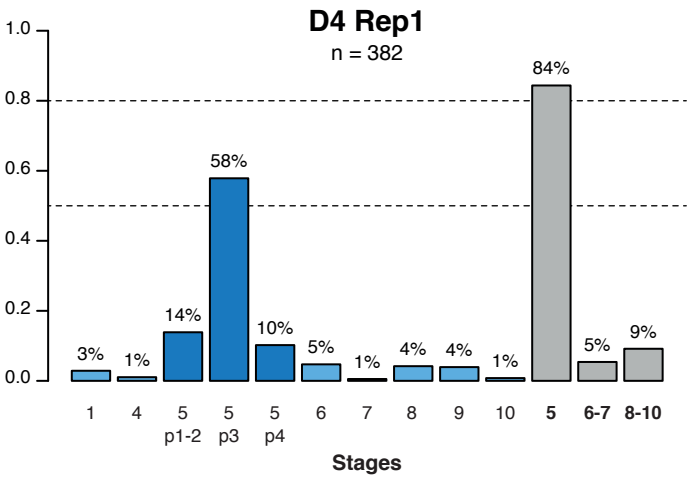

D1 Rep2
n = 46
1.0
0.8
0.6
0.4
0.2
0.0
17% 2% 4% 45% 16% 6% 2% 0% 7% 0% 79% 11% 8%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages

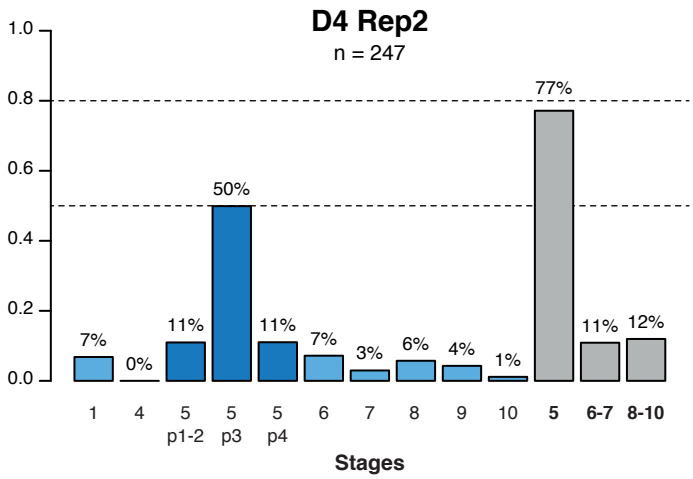

D2 Rep1
n = 289
1.0
0.8
0.6
0.4
0.2
0.0
2% 0% 12% 50% 26% 5% 2% 1% 1% 0% 91% 7% 1%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages

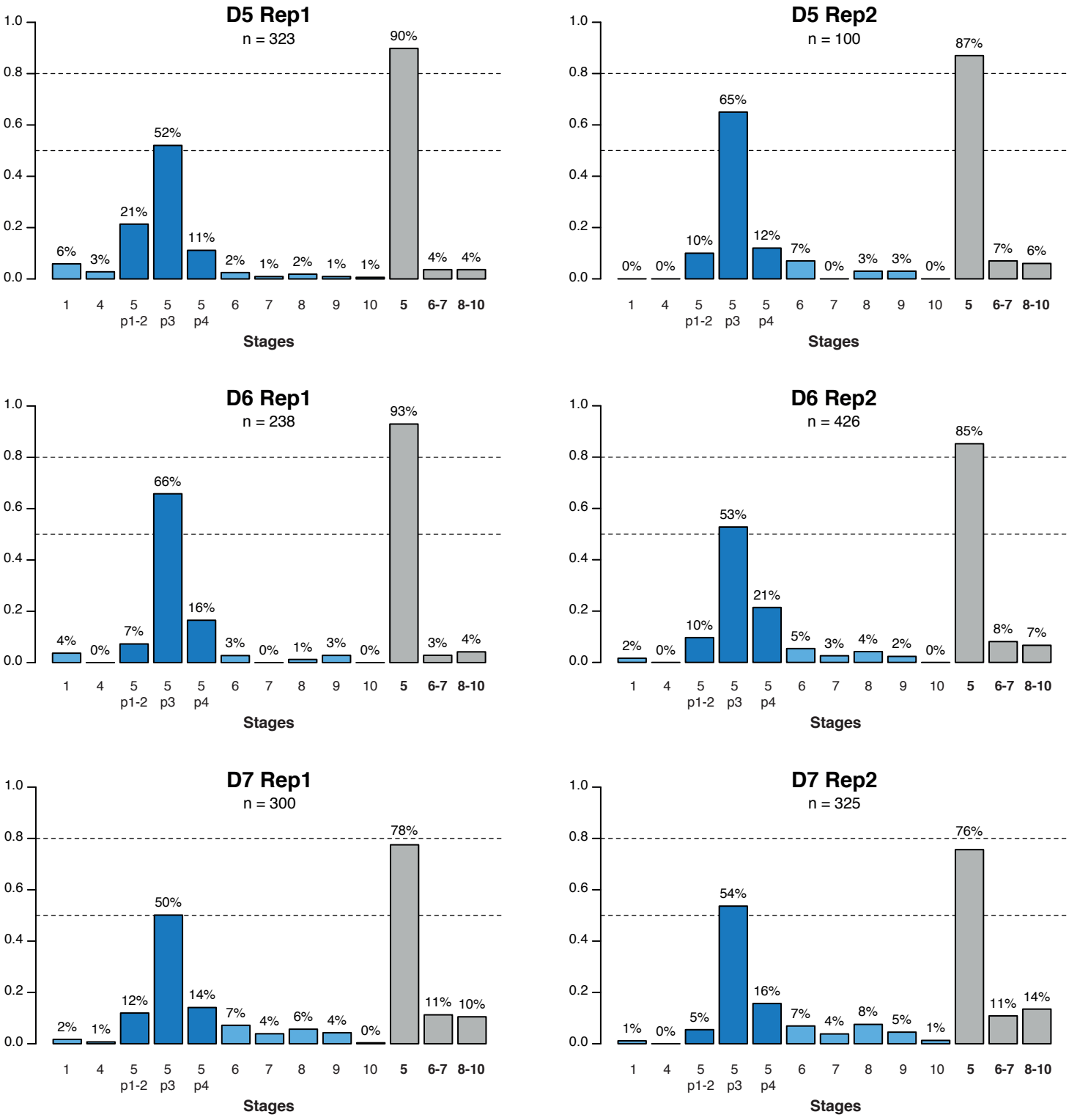

D2 Rep2
n = 345
1.0
0.8
0.6
0.4
0.2
0.0
5% 0% 7% 62% 14% 4% 2% 2% 3% 0% 88% 7% 5%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages


D3 Rep1
n = 243
1.0
0.8
0.6
0.4
0.2
0.0
2% 0% 17% 51% 12% 2% 0% 6% 9% 0% 82% 3% 15%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages


D3 Rep2
n = 403
1.0
0.8
0.6
0.4
0.2
0.0
1% 0% 8% 52% 18% 4% 3% 9% 4% 1% 78% 8% 14%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages


D4 Rep1
n = 382
1.0
0.8
0.6
0.4
0.2
0.0
3% 1% 14% 58% 10% 5% 1% 4% 4% 1% 84% 5% 9%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages


D4 Rep2
n = 247
1.0
0.8
0.6
0.4
0.2
0.0
7% 0% 11% 50% 11% 7% 3% 6% 4% 1% 77% 11% 12%
1 4 5 p1-2 5 p3 5 p4 6 7 8 9 10 5 6-7 8-10
Stages



Supplemental Figure S1. Temporal profiles of embryonic collections. A batch of embryos from each collection was fixed with formaldehyde and examined under the differential interference contrast microscope (DIC) to characterize their distribution of developmental stages (Supplemental Methods). Bar plots show proportional representation of the stages in individual collections of embryos (domains: D1-D7, replicates: Rep1-Rep2). Number of scored embryos is indicated. In blue: proportion of individual embryonic stages. Stage 5 is additionally subdivided into phase 1-2 (p1-2), phase 3 (p3) and phase 4 (p4), reflecting different degrees of blastoderm cellularization (after Schroeder et al. 2011). In grey: total proportion of embryos at stage 5 (cellularizing blastoderm), stage 6-7 (gastrulation) and stage 8-10 (germ band elongation). For higher consistency, stage 1 was omitted from these summary calculations as it corresponds to unfertilized eggs that minimally contribute to ATAC-seq libraries.

In each collection, more than 80% of embryos represent stage 5 of cellularizing blastoderm, with more than 50% of embryos corresponding to the mid-point of cellularization (phase 3). A negligible number of embryos represent developmental stages prior to the global zygotic genome activation (stage 4 or earlier). While contamination with older embryos extends up to stage 9, a considerable proportion of those corresponds to stages 6-7, each lasting 10 min only (Weigmann et al. 2003).

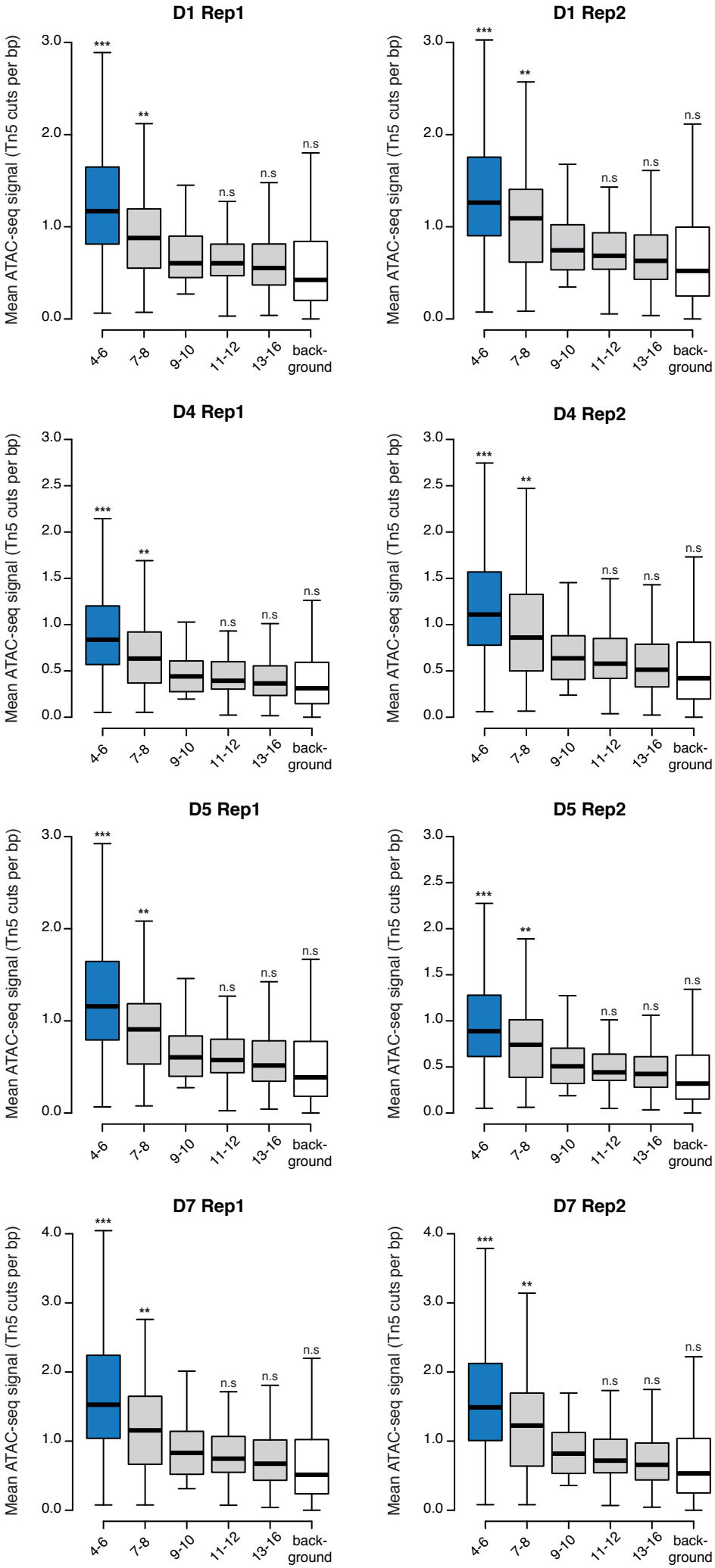
Exclusion of stages prior to the global zygotic genome activation ensures that we examine accessibility modulations during full operation of the patterning gene regulatory networks. High similarity of temporal profiles between independent collections (including comparable contribution of older stages) allows for a substantial reduction of the temporal dimension of our assay. As a result, we are confident that the relative signal differences between the tagged domains primarily reflect regional accessibility variation along the AP axis, rather than temporal discrepancies between the samples.

Contamination with older embryonic stages could have a potential confounding effect on the absolute measure of the ATAC-seq signal, possibly resulting in the observed residual accessibility of inactive elements. However, Haines and Eisen 2018, who specifically select stage-5 embryos by hand sorting, independently report the same property of axis patterning enhancers.

References:

Schroeder MD, Greer C, Gaul U. 2011. How to make stripes: deciphering the transition from non-periodic to periodic patterns in Drosophila segmentation. Development 138: 3067–3078.

Weigmann K, Klapper R, Strasser T, Rickert C, Technau G, Jäckle H, Janning W, Klämbt C. 2003. FlyMove - A new way to look at development of Drosophila. Trends Genet 19: 310–311.


D1 Rep1
D1 Rep2
D4 Rep1
D4 Rep2
D5 Rep1
D5 Rep2
D7 Rep1
D7 Rep2
Mean ATAC-seq signal (Tn5 cuts per bp)
***
**
n.s
4-6
7-8
9-10
11-12
13-16
back-
ground

Supplemental Figure S2. Accessibility of genes transcribed at different stages of embryogenesis. Mean ATAC-seq signal (transposase cuts per bp) was calculated over coding sequences of genes whose transcription is initiated at different stages of embryogenesis (classification according to RNA *in situ* hybridization patterns from Berkeley Drosophila Genome Project, Tomancak et al. 2002) and inaccessible background regions (no significant signal elevation in any tagged domain or a whole-embryo control). See Supplemental Methods for details. Boxplots show distribution of the mean ATAC-seq signal for individual replicates of whole-embryo controls (D1, D4, D5 and D7; Rep1 and Rep2). Blue: genes transcribed at stage 4-6. Grey: genes activated at later stages 7-8, 9-10, 11-12 and 13-16 (inactive in cellularizing blastoderm). White: inaccessible background regions. Asterisks indicate significant differences in reference to genes from stage 9-10 (Student's t-test): p-value < 0.001 (**), p-value < 0.0001 (***); n.s: non-significant.

Transcribed gene bodies have been demonstrated to display elevated signal in chromatin accessibility assays (Thomas et al. 2011; Bell et al. 2011; Weintraub and Groudine 1976; Bulger et al. 2003). We took advantage of this property to assess contribution of the ATAC-seq signal from embryonic stages older than stage 5 of cellularizing blastoderm (Supplemental Fig. S1). We compared accessibility profiles over different temporal classes of embryonic genes. Only coding sequences were considered in order to avoid confounding effects of intronic *cis*-regulatory elements. We only examined whole-embryo samples and did not consider spatial expression patterns of the analyzed genes.

The oldest embryos from the collections represent stage 9 (Supplemental Fig. S1), and contribution from this stage can be additionally enhanced by mitotic divisions that resume at the end of stage 7 (Foe et al. 1993). However, genes specifically activated at stage 9-10 are characterized by very similar distribution of their ATAC-seq signal as those transcribed at later stages 11-16, which are absent from the collections. ATAC-seq signal of stage 9-10 genes is significantly lower in comparison to genes active in cellularizing blastoderm (stage 4-6), while it is highly similar to that of inaccessible background regions. Overall, we conclude that accessibility of genes transcribed at stage 9-10 is minimal, which demonstrates negligible contribution of the oldest stages in the obtained ATAC-seq profiles.

While stage 7-8 genes show reduced accessibility in comparison to stage 4-6 genes, their signal is elevated in respect to inaccessible background regions. A reason for this discrepancy is not clear, taking into consideration comparable representation of stages 7-8 and 9-10 in the collections.

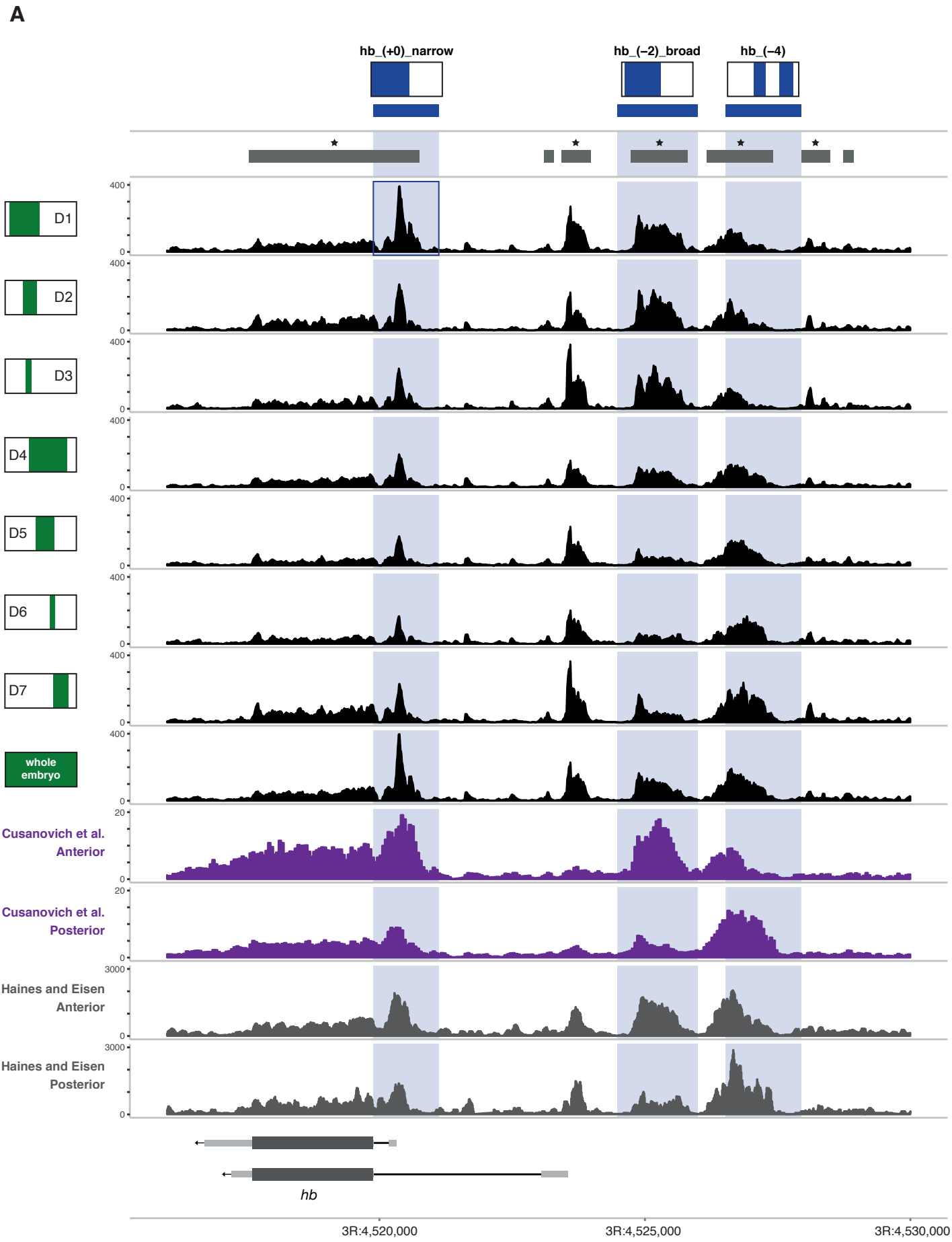
References:

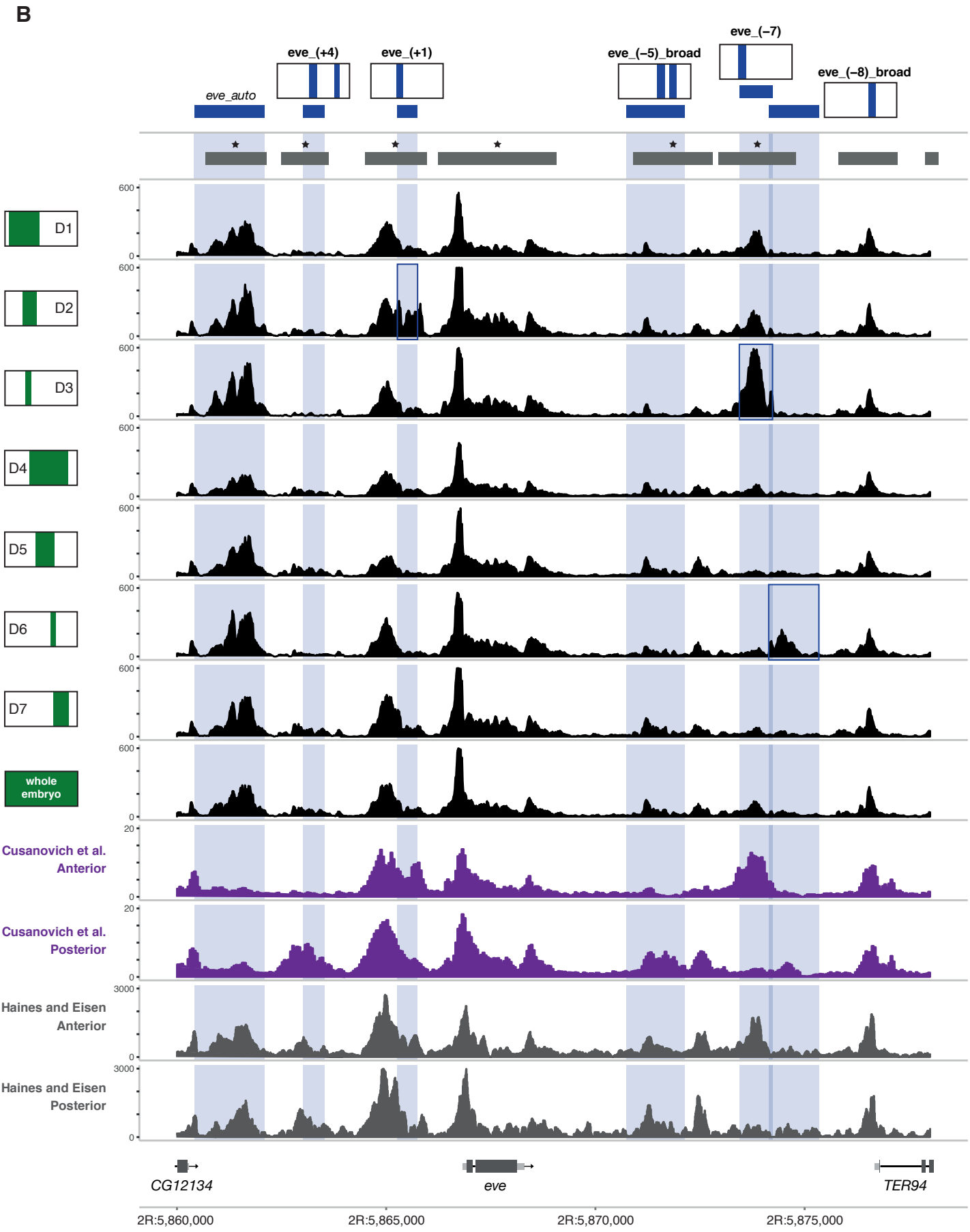
Bulger M, Schübeler D, Bender MA, Hamilton J, Farrell CM, Hardison RC, Groudine M. 2003. A Complex Chromatin Landscape Revealed by Patterns of Nuclease Sensitivity and Histone Modification within the Mouse β-Globin Locus. Mol Cell Biol 23: 5234–5244.

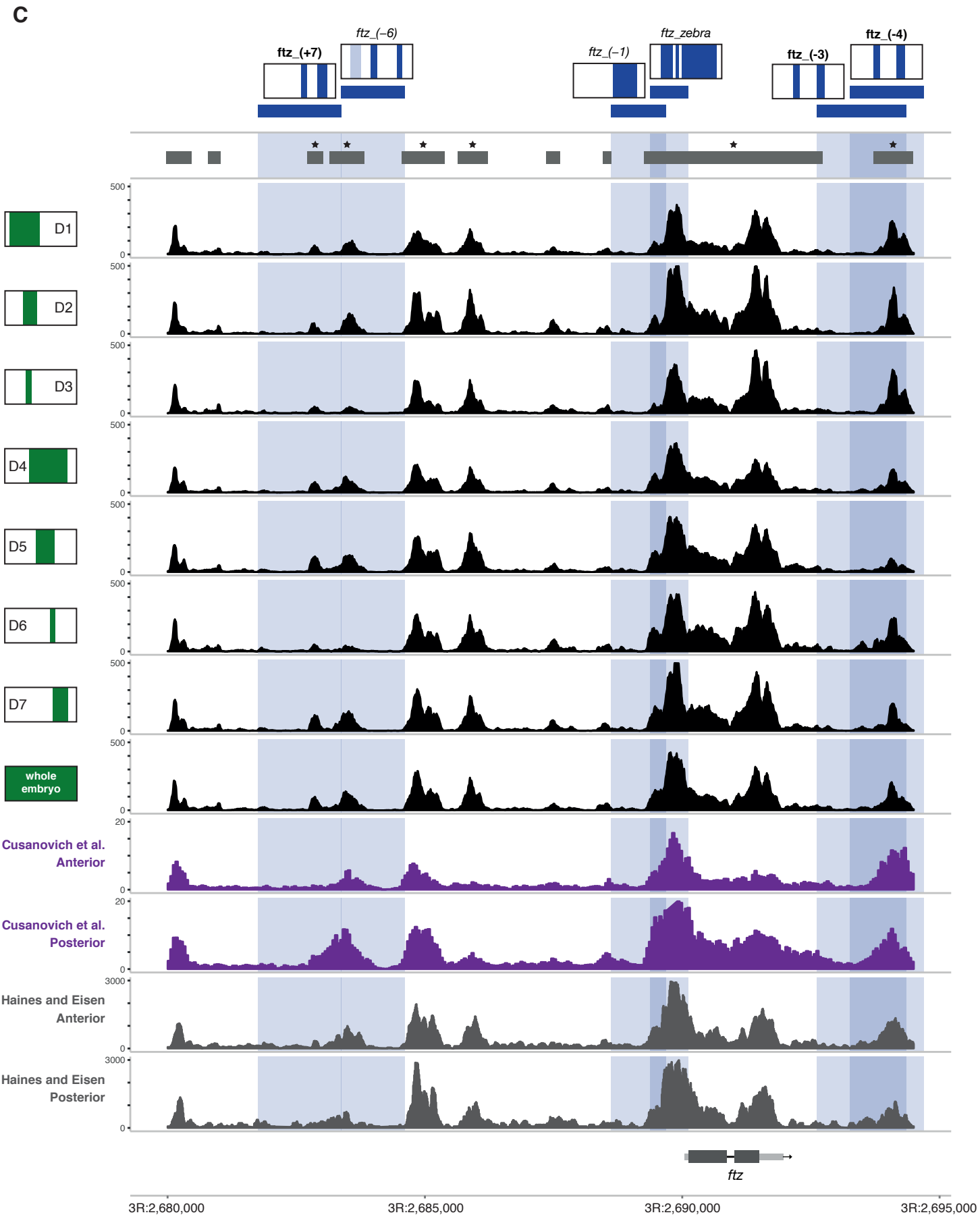
Foe VE, Odell GM, Edgar B a. 1993. Timetable of Drosophila Early Development. In *The Development of Drosophila melanogaster*, pp. 149–300.

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

Weintraub H, Groudine M. 1976. Chromosomal subunits in active genes have an altered conformation. Science (80-) 193: 848–856.


A
hb_(+0)_narrow
hb_(−2)_broad
hb_(−4)
D1
D2
D3
D4
D5
D6
D7
whole embryo
Cusanovich et al. Anterior
Cusanovich et al. Posterior
Haines and Eisen Anterior
Haines and Eisen Posterior
400
20
3000
0
hb
3R:4,520,000
3R:4,525,000
3R:4,530,000


B
eve_auto
eve_(+4)
eve_(+1)
eve_(−5)_broad
eve_(−7)
eve_(−8)_broad
D1
D2
D3
D4
D5
D6
D7
whole embryo
Cusanovich et al. Anterior
Cusanovich et al. Posterior
Haines and Eisen Anterior
Haines and Eisen Posterior
600
20
3000
0
CG12134
eve
TER94
2R:5,860,000
2R:5,865,000
2R:5,870,000
2R:5,875,000


C
ftz_(+7)
ftz_(−6)
ftz_(−1)
ftz_zebra
ftz_(-3)
ftz_(-4)
D1
D2
D3
D4
D5
D6
D7
whole embryo
Cusanovich et al. Anterior
Cusanovich et al. Posterior
Haines and Eisen Anterior
Haines and Eisen Posterior
ftz
3R:2,680,000
3R:2,685,000
3R:2,690,000
3R:2,695,000

Supplemental Figure S3. Regional differences in chromatin accessibility at the loci of selected AP regulators. (*A*) *hunchback* (*hb*), (*B*) *even skipped* (*eve*) and (*C*) *fushi tarazu* (*ftz*). Comparison of coverage tracks between tagged domains (black), anterior and posterior clusters from single-cell ATAC-seq by Cusanovich et al. 2018 (purple) and anterior and posterior halves of cryo-sliced embryos from Haines and Eisen 2018 (grey).

Blue horizontal bars and underlying shaded regions represent coordinates of known enhancers. Activity patterns of the elements in blastoderm embryos are indicated schematically above (blue shading; based on published RNA *in situ* hybridization images). In bold: elements included in the list of curated 88 AP enhancers (REDfly names and references provided in Supplemental Table S8). In italics: additional *cis*-regulatory elements, (*B*) *eve_auto* autoregulatory element from Jiang et al. 1991 (REDfly name: eve_HZE1600), (*C*) *ftz_(-6)* and *ftz_(-1)* from Schroeder et al. 2011; *ftz_zebra* from Hiromi et al. 1985 (REDfly name: ftz_zebra_element). Grey horizontal bars represent coordinates of high-confidence ATAC-seq peaks. Differential peaks are indicated with a black star. Genomic coordinates and gene models: Flybase Release 5.57.

Enhancers that coincide with drivers used for expressing UNC84-3xFLAG are marked with a dark blue frame in their respective tagged domain. Note that their elevated ATAC-seq signal likely results from the presence of two accessible copies of the element, one in its endogenous locus and one in the attP2 integration site (see Supplemental Methods).

Black: normalized coverage of 1-100 bp ATAC-seq fragments, smoothed over a sliding window of 15 bp. Accessibility profiles represent individual replicates of tagged domains (D4: replicate 1, other domains: replicate 2) and a single whole-embryo control (D1 replicate 2). AP positions of the profiled domains are indicated schematically on the left (green shading).

Purple: normalized coverage tracks represent a mean over two anterior blastoderm clusters (cluster 6 and 15) and a mean over three blastoderm posterior clusters (cluster 4, 7 and 16) identified *in silico* among single nuclei from the 2-4 h embryonic collection. Note that due to unsurpervised clustering the exact position and size of the clusters are not known. Data downloaded from: http://shiny.furlonglab.embl.de/scATACseqBrowser/.

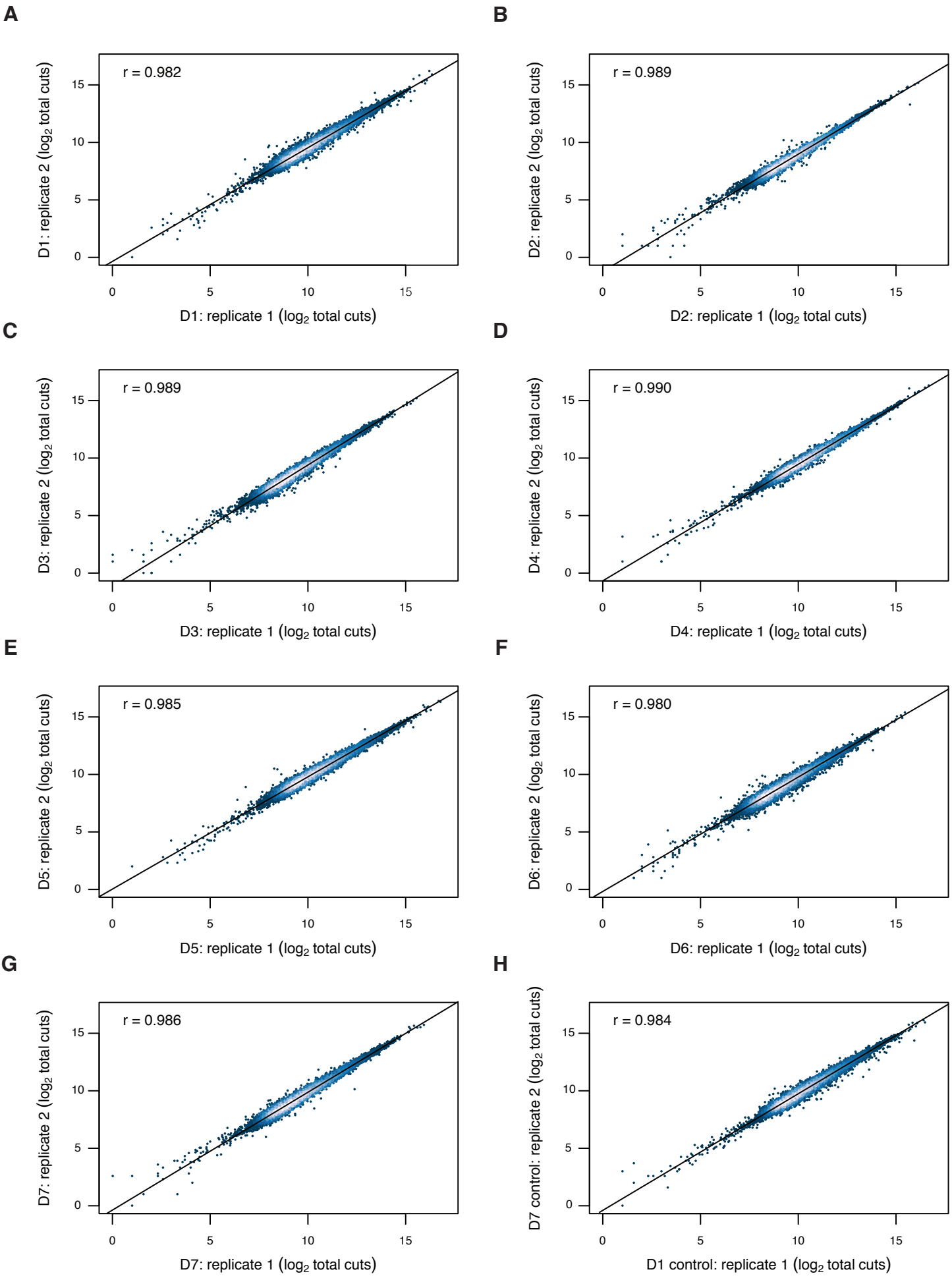
Grey: normalized coverage tracks represent anterior and posterior halves of cryo-sliced embryos in nuclear cycle 14. Data downloaded from Gene Expression Omnibus, accession number: GSE104957.

References:

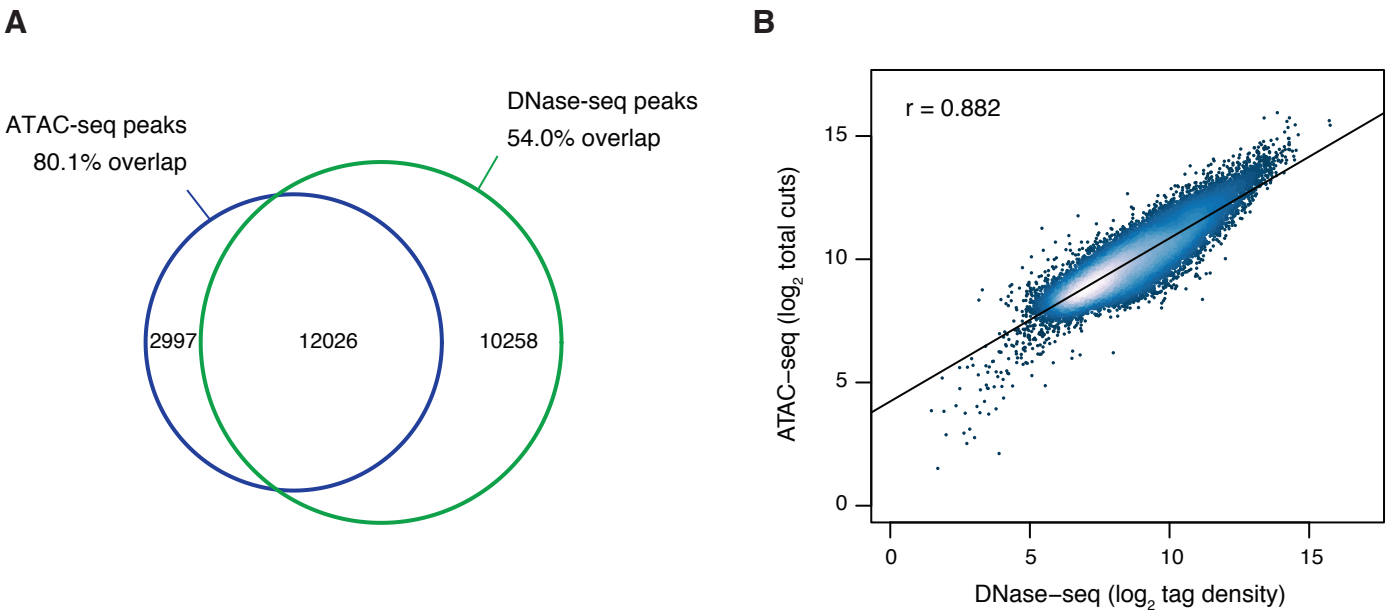
Hiromi Y, Kuroiwa A, Gehring WJ. 1985. Control elements of the Drosophila segmentation gene fushi tarazu. Cell 43: 603–13.

Jiang J, Hoey T, Levine M. 1991. Autoregulation of a segmentation gene in Drosophila: Combinatorial interaction of the even-skipped homeo box protein with a distal enhancer element. Genes Dev 5: 265–277.

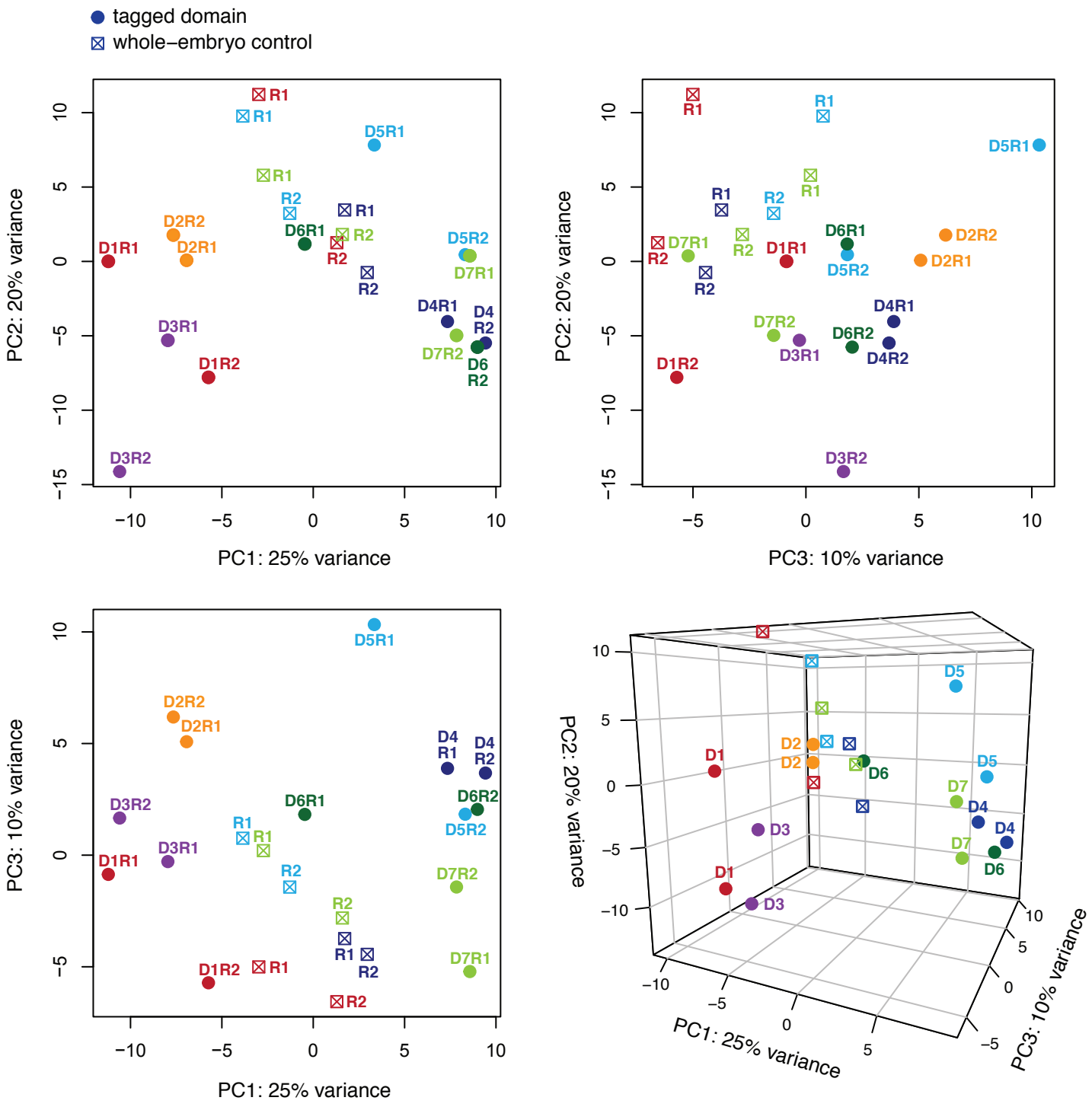
Schroeder MD, Greer C, Gaul U. 2011. How to make stripes: deciphering the transition from non-periodic to periodic patterns in Drosophila segmentation. Development 138: 3067–3078.


A
r = 0.982
D1: replicate 2 (log2 total cuts)
D1: replicate 1 (log2 total cuts)
B
r = 0.989
D2: replicate 2 (log2 total cuts)
D2: replicate 1 (log2 total cuts)
C
r = 0.989
D3: replicate 2 (log2 total cuts)
D3: replicate 1 (log2 total cuts)
D
r = 0.990
D4: replicate 2 (log2 total cuts)
D4: replicate 1 (log2 total cuts)
E
r = 0.985
D5: replicate 2 (log2 total cuts)
D5: replicate 1 (log2 total cuts)
F
r = 0.980
D6: replicate 2 (log2 total cuts)
D6: replicate 1 (log2 total cuts)
G
r = 0.986
D7: replicate 2 (log2 total cuts)
D7: replicate 1 (log2 total cuts)
H
r = 0.984
D7 control: replicate 2 (log2 total cuts)
D1 control: replicate 1 (log2 total cuts)
0
5
10
15

Supplemental Figure S4. High reproducibility of replicate ATAC-seq experiments. Pair-wise comparison of ATAC-seq signal between: (*A*-*G*) duplicates of individual tagged domains (D1-D7), and (*H*) representative whole-embryo controls (D1 control replicate 1 vs. D7 control replicate 2). Each panel shows a scatter plot of $\log_2$ transformed ATAC-seq signal (total count of Tn5 transposase cuts) over 17 345 high-confidence ATAC-seq peaks. r = Pearson's correlation coefficient.

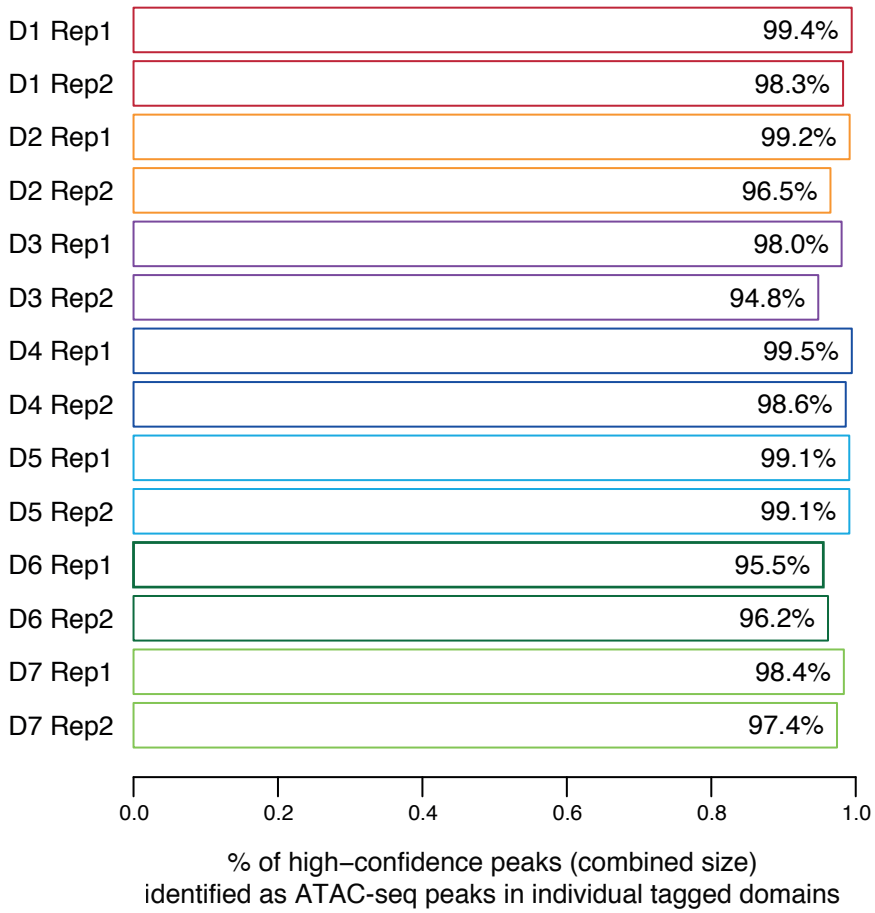


Supplemental Figure S5. High similarity between whole-embryo controls and published chromatin accessibility profiles from stage-5 embryos. Comparison of ATAC-seq data from whole-embryo controls with DNase-seq accessibility profiles of whole stage-5 embryos from Thomas et al. 2011. (*A*) Venn diagram represents the overlap between high-confidence ATAC-seq peaks (blue) and DNase-seq peaks (green). DNase-seq peaks were defined as an intersection of peaks called in two replicates (FDR = 5%). Note that different algorithms and significance thresholds were used for calling ATAC-seq and DNase-seq peaks (see Supplemental Methods). (*B*) Scatter plot represents a correlation between $\log_2$ transformed ATAC-seq and DNase-seq signal intensities, calculated for the set of 17 345 high-confidence ATAC-seq peaks. ATAC-seq signal was defined as the normalized count of Tn5 transposase cuts (average over eight whole-embryo controls). DNase-seq signal was defined as the normalized DNaseI tag density (average over two replicates). r = Pearson's correlation coefficient.

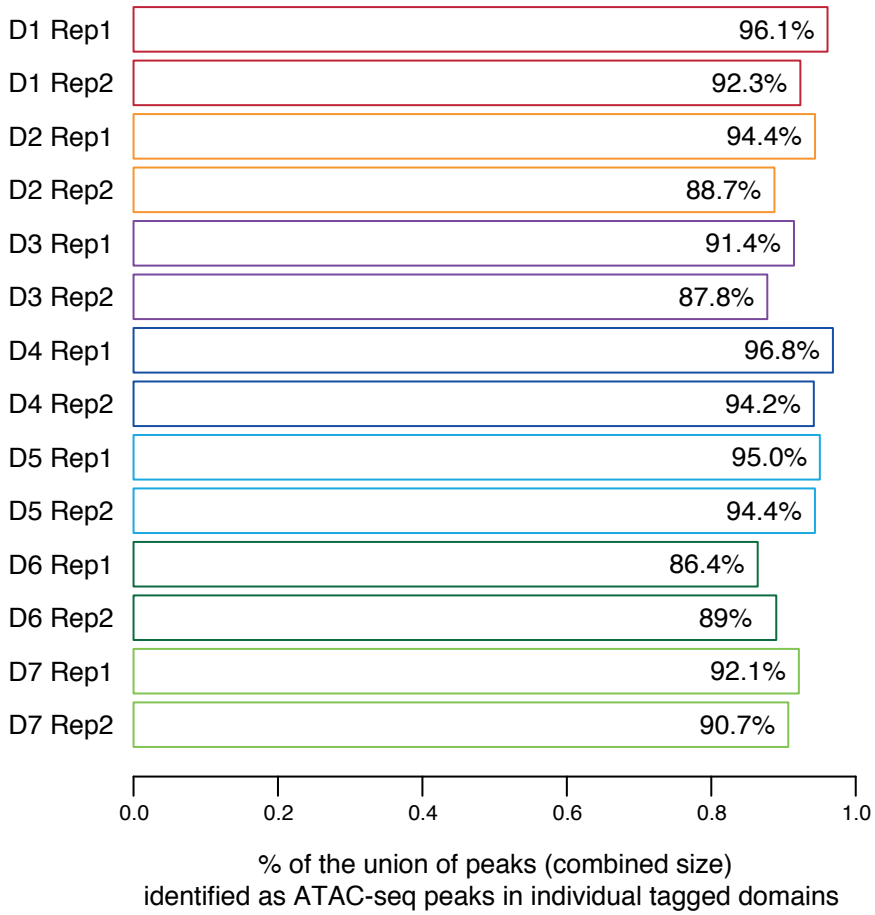


Supplemental Figure S6. Principal component analysis of genome-wide accessibility variation. Each panel shows distribution of individual tagged domains (solid circle) and whole-embryo controls (crossed square) against two principal components (individual planes of the 3D PCA plot from Fig. 3A). Replicates are represented as separate data points (R1: replicate 1, R2: replicate 2) and color-coded by genotype (D1: red; D2: orange; D3: purple; D4: dark blue; D5: light blue; D6: dark green; D7: light green). PCA is based on accessibility signal (total count of Tn5 transposase cuts) over 17 345 high-confidence ATAC-seq peaks.

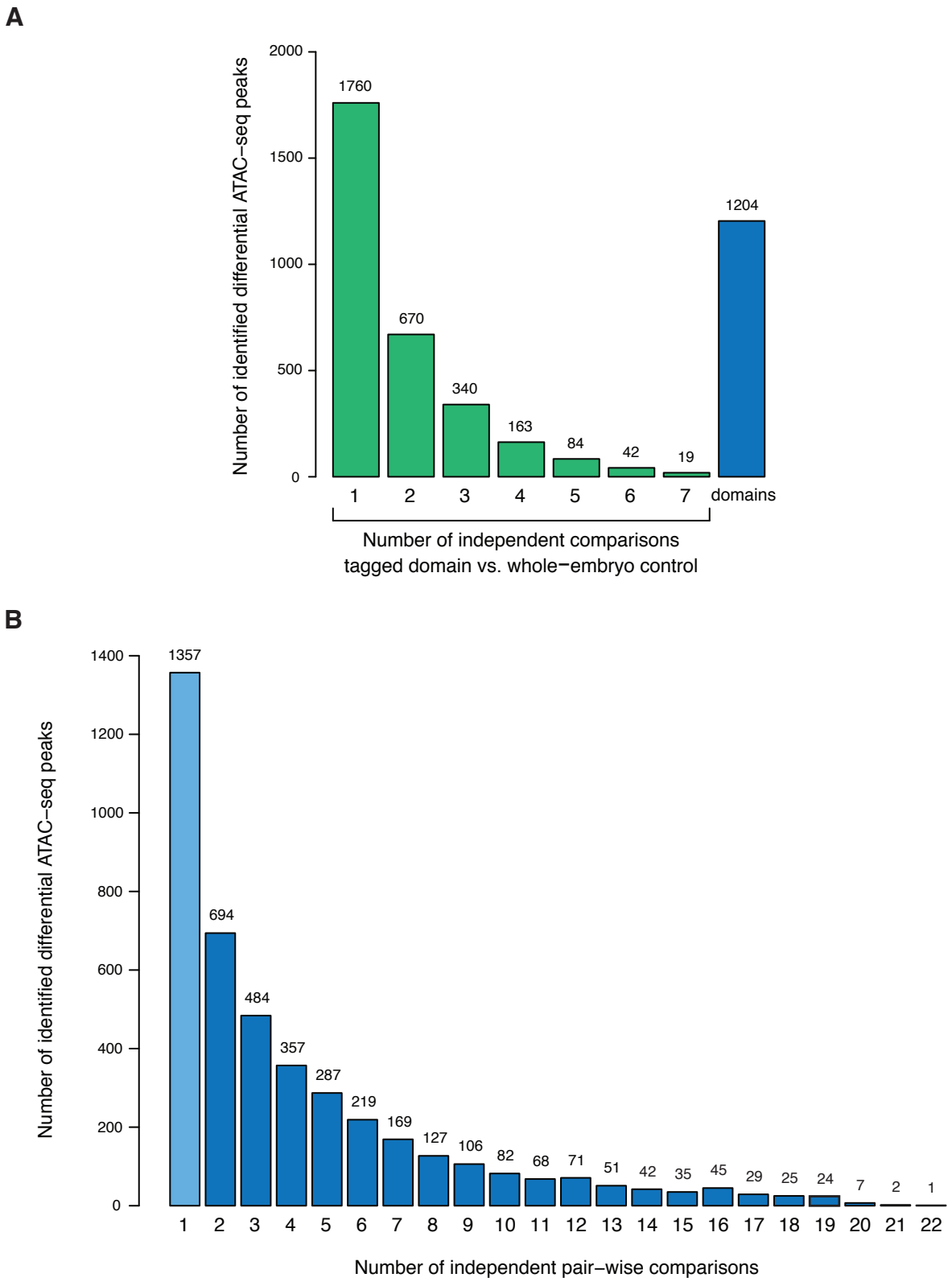
A


D1 Rep1
D1 Rep2
D2 Rep1
D2 Rep2
D3 Rep1
D3 Rep2
D4 Rep1
D4 Rep2
D5 Rep1
D5 Rep2
D6 Rep1
D6 Rep2
D7 Rep1
D7 Rep2
99.4%
98.3%
99.2%
96.5%
98.0%
94.8%
99.5%
98.6%
99.1%
99.1%
95.5%
96.2%
98.4%
97.4%
0.0
0.2
0.4
0.6
0.8
1.0
% of high−confidence peaks (combined size)
identified as ATAC-seq peaks in individual tagged domains

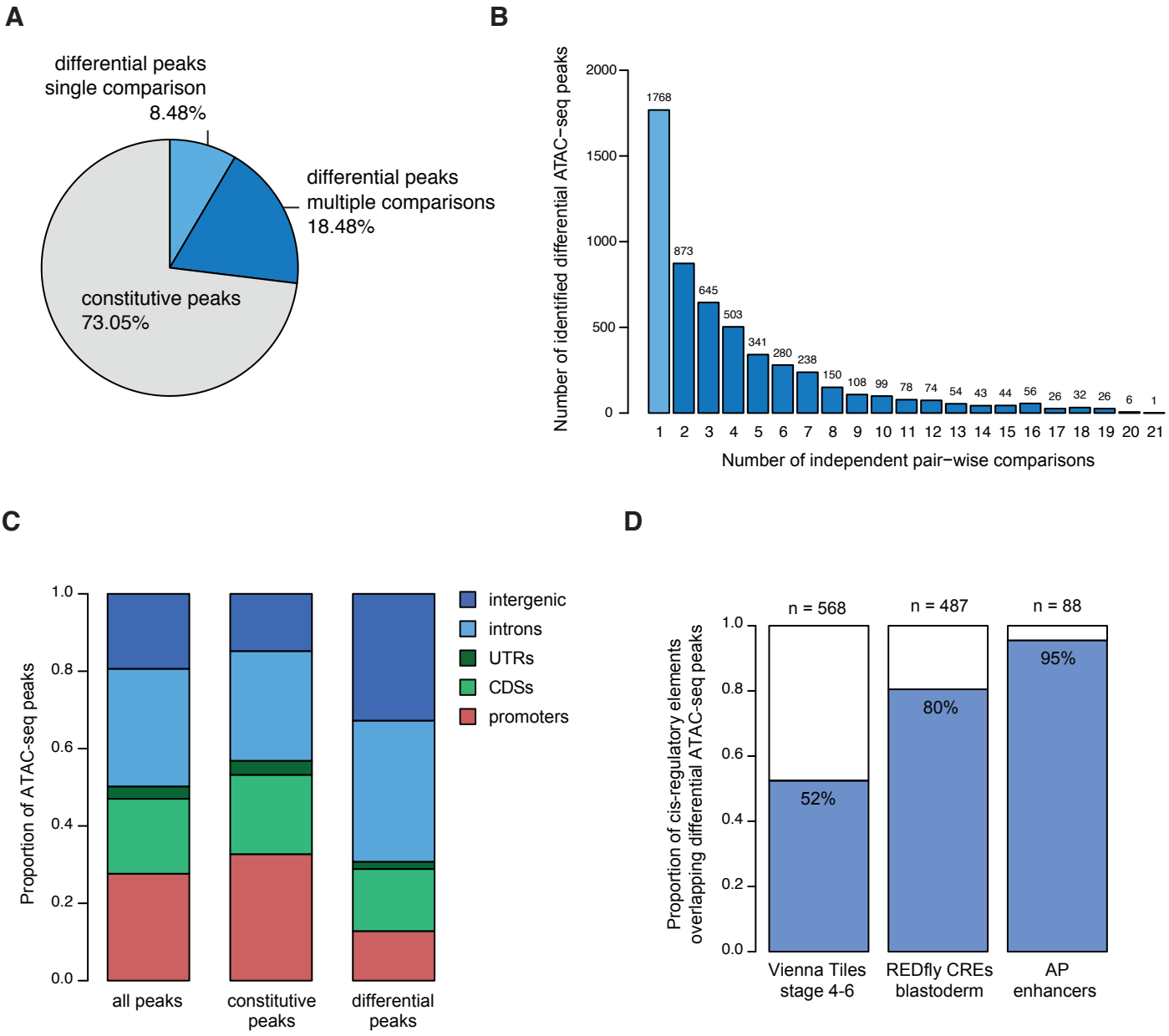
B


D1 Rep1
D1 Rep2
D2 Rep1
D2 Rep2
D3 Rep1
D3 Rep2
D4 Rep1
D4 Rep2
D5 Rep1
D5 Rep2
D6 Rep1
D6 Rep2
D7 Rep1
D7 Rep2
96.1%
92.3%
94.4%
88.7%
91.4%
87.8%
96.8%
94.2%
95.0%
94.4%
86.4%
89%
92.1%
90.7%
0.0
0.2
0.4
0.6
0.8
1.0
% of the union of peaks (combined size)
identified as ATAC-seq peaks in individual tagged domains

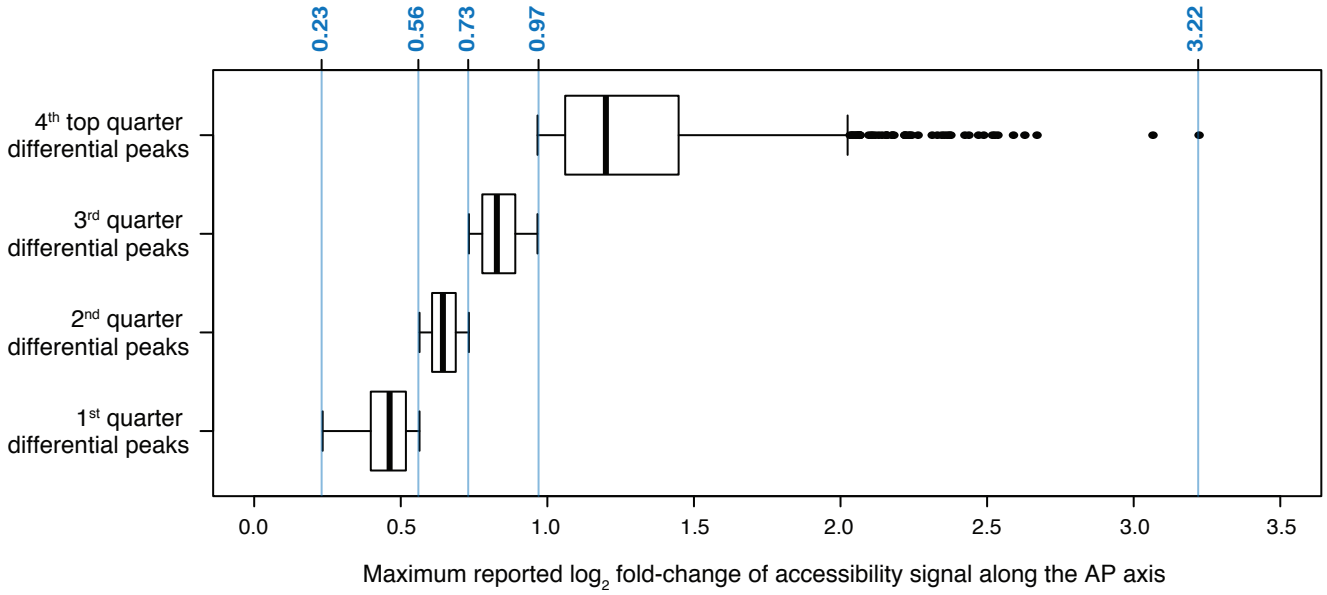
Supplemental Figure S7. Conserved genome-wide distribution of accessible regions. A vast majority of ATAC-seq peaks is characterized by significant signal enrichment in individual tagged domains. Bar plots show the proportion of (*A*) high-confidence peaks and (*B*) union of peaks that overlap ATAC-seq peaks identified in individual tagged domains (minimum overlap = 50 bp). High-confidence peaks represent an intersection of ATAC-seq peaks called in all whole-embryo controls, while the union of peaks represents a comprehensive set of ATAC-seq peaks called in all tagged domains and whole-embryo controls (Supplemental Methods). Proportion of overlap is defined in terms of the combined size of the overlapped peaks. Duplicates are plotted separately.



Supplemental Figure S8. Summary of differential high-confidence ATAC-seq peaks. (*A*) Bar plot shows the number of differential peaks supported by multiple pair-wise comparisons between tagged domains and whole-embryo controls (green) and identified uniquely in pair-wise comparisons between tagged domains (blue). (*B*) Bar plot shows the number of ATAC-seq peaks identified independently in all 28 pair-wise comparisons. Light blue: peaks identified in a single comparison, dark blue: peaks identified in multiple comparisons.

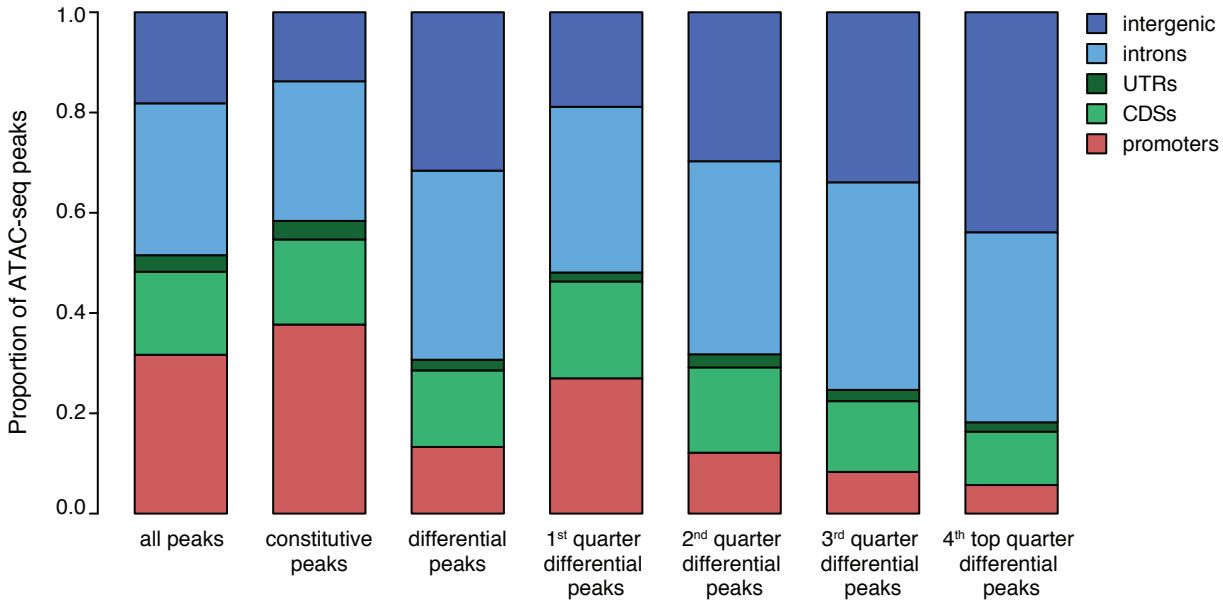


Supplemental Figure S9. Differential analysis of the union of ATAC-seq peaks. DESeq2 analysis was performed for a broader definition of accessible regions, representing the union of ATAC-seq peaks from a whole-embryo control and individual tagged domains (Supplemental Methods). (*A*) Pie chart shows proportion of the accessible genome (combined size of the union of ATAC-seq peaks; 21 414 intervals) represented by constitutive regions that show no significant variation in their accessibility signal along the AP axis (grey; 15 969 intervals), and differential peaks that are supported by a single pair-wise comparison (light blue; 1 768 intervals) and multiple pair-wise comparisons (dark blue; 3 677 intervals). (*B*) Bar plot shows the number of differential ATAC-seq peaks identified independently in 28 pair-wise comparisons. Light blue: peaks identified in a single comparison, dark blue: peaks identified in multiple comparisons. (*C*) Proportional distribution of genomic annotations among different classes of accessible regions: all intervals from the union of ATAC-seq peaks (all peaks), constitutive peaks and differential peaks. UTR: 5' and 3' untranslated regions. CDS: coding sequence. (*D*) Bar plot indicates the proportion of annotated *cis*-regulatory elements (CREs) that overlap differential ATAC-seq peaks. Total number of CREs from each category is indicated above the bars: Vienna Tiles (Kvon et al. 2014) active at stage 4-6, REDfly CREs (Gallo et al. 2011) active in blastoderm embryos and AP enhancers driving patterned expression specifically along the AP axis (Supplemental Table S8).

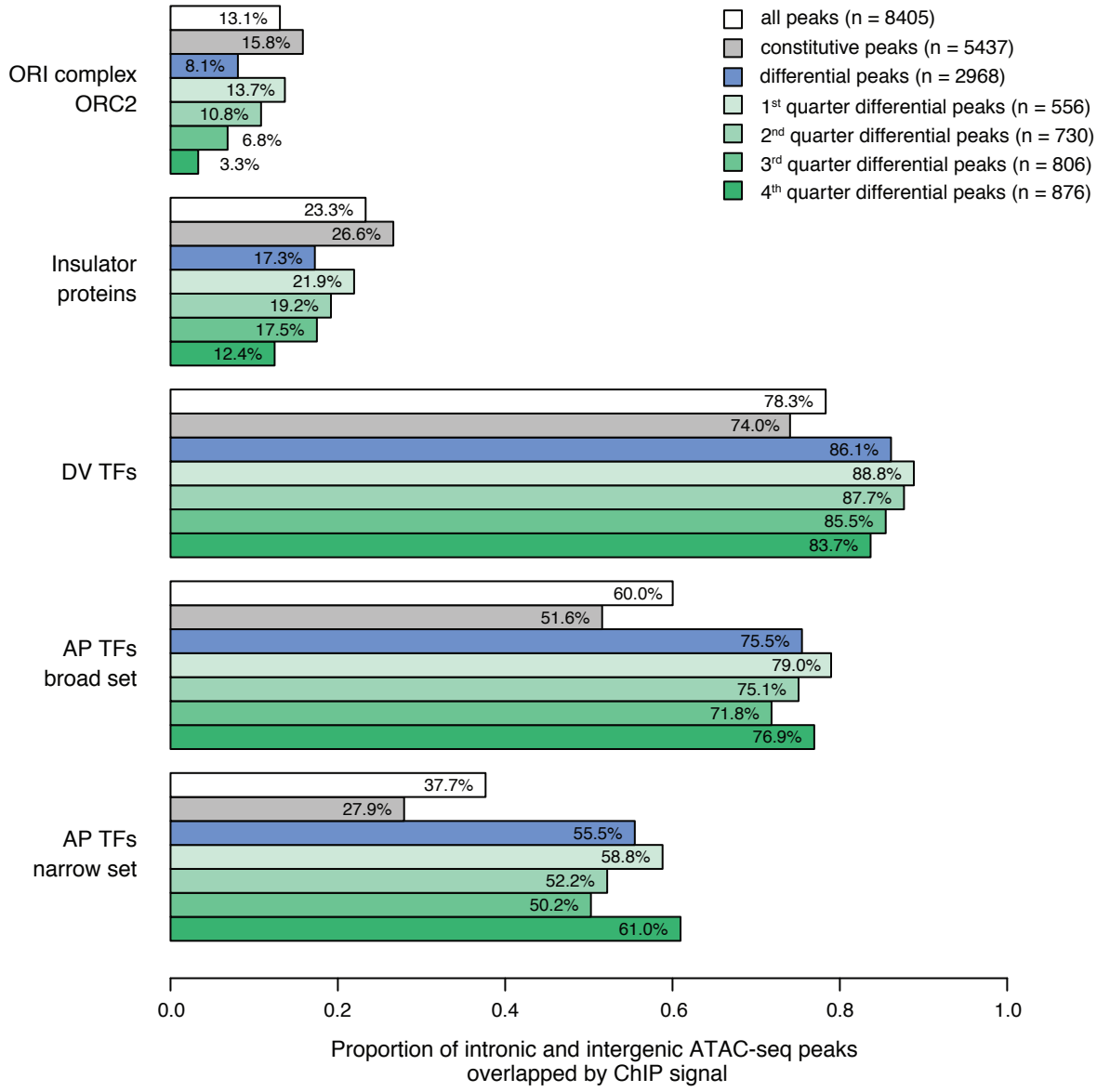


Supplemental Figure S10. Distribution of the magnitude of accessibility changes. Differential high-confidence peaks were binned into quarters based on the maximum $\log_2$ fold-change of their accessibility signal. Box plots represent distribution of $\log_2$ fold-change within each quarter. Blue lines represent quantiles.

A

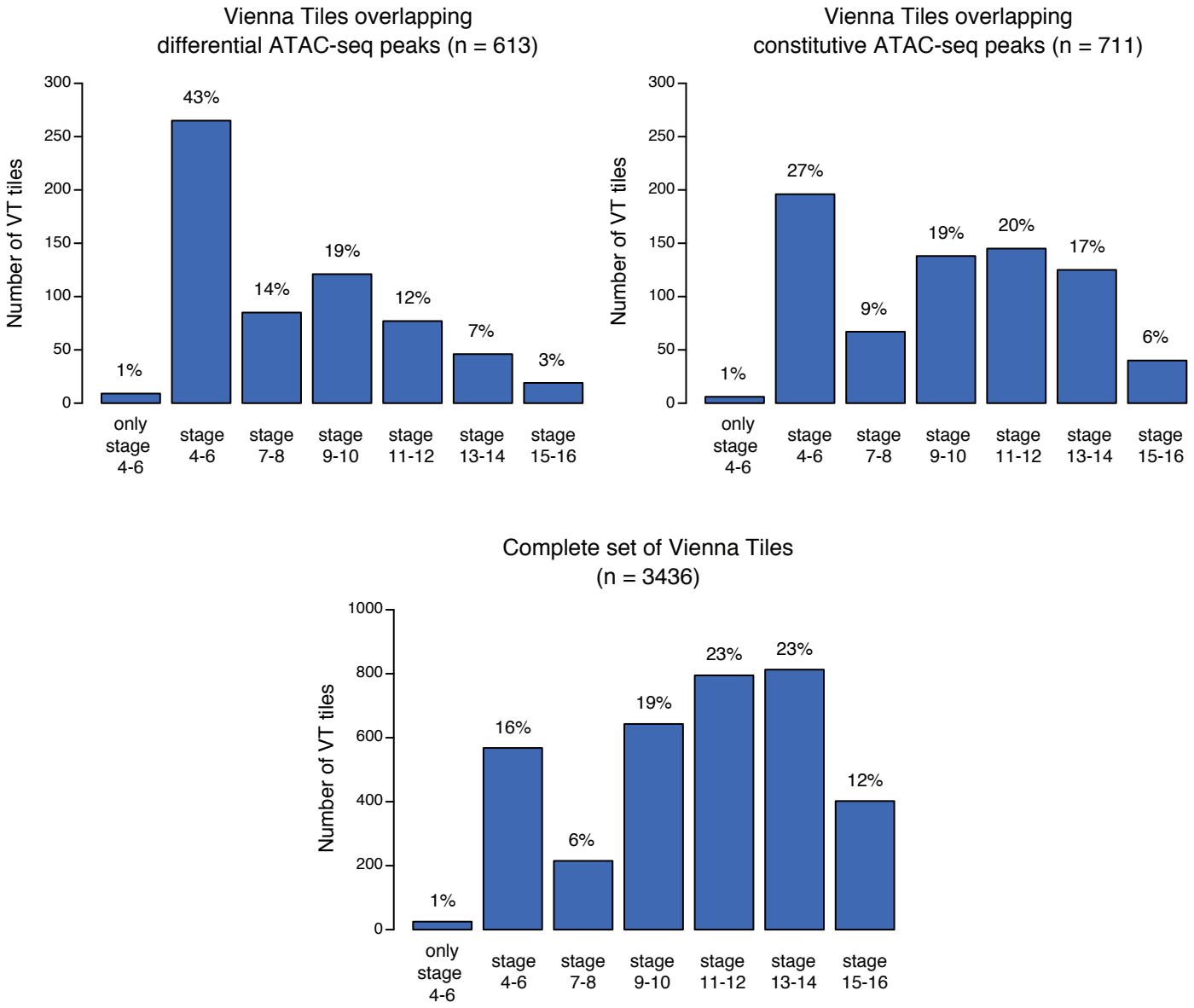

Proportion of ATAC-seq peaks
1.0
0.8
0.6
0.4
0.2
0.0
intergenic
introns
UTRs
CDSs
promoters
all peaks
constitutive peaks
differential peaks
1st quarter differential peaks
2nd quarter differential peaks
3rd quarter differential peaks
4th top quarter differential peaks

B

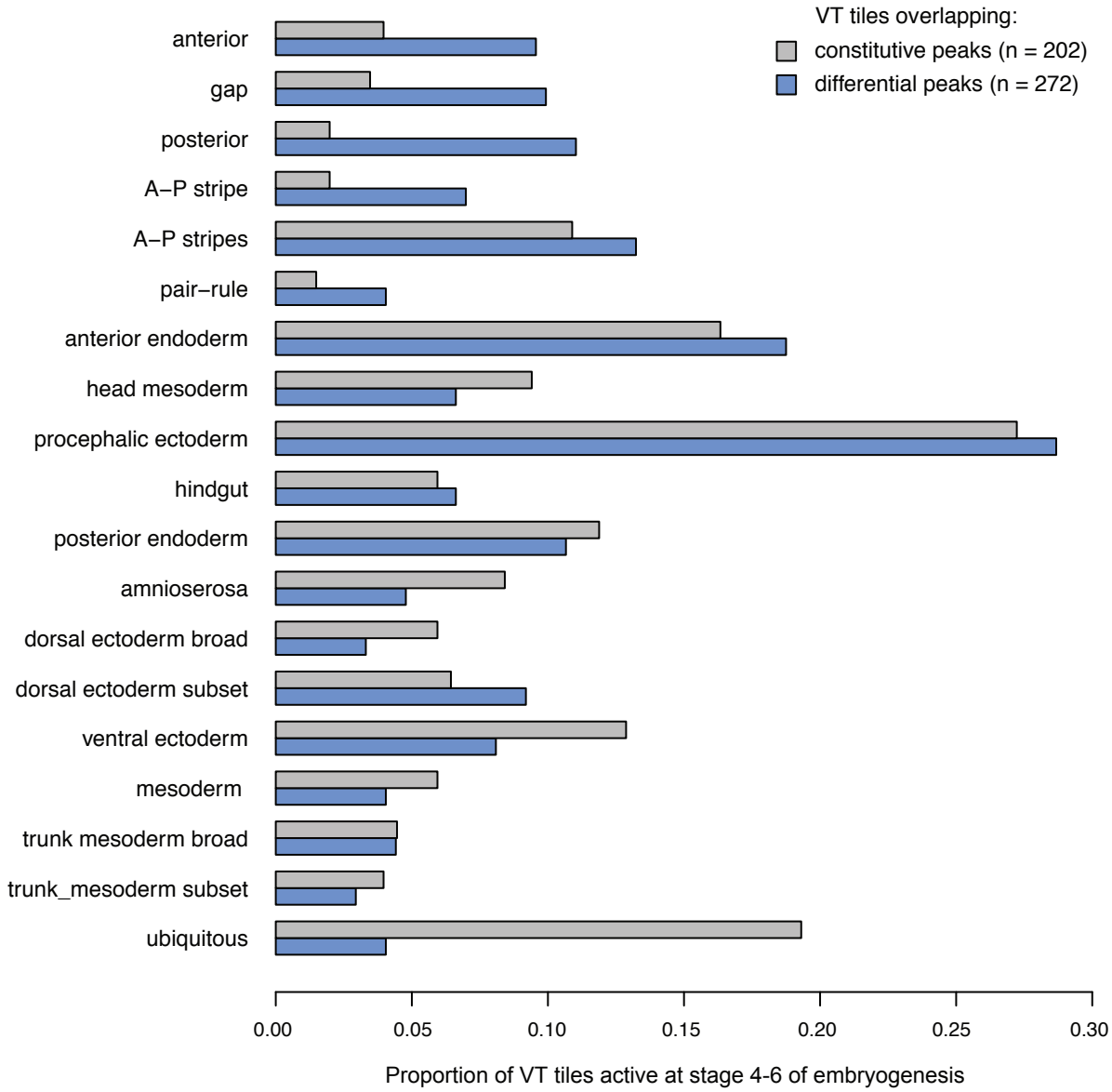

all peaks (n = 8405)
constitutive peaks (n = 5437)
differential peaks (n = 2968)
1st quarter differential peaks (n = 556)
2nd quarter differential peaks (n = 730)
3rd quarter differential peaks (n = 806)
4th quarter differential peaks (n = 876)
ORI complex ORC2
13.1%
15.8%
8.1%
13.7%
10.8%
6.8%
3.3%
Insulator proteins
23.3%
26.6%
17.3%
21.9%
19.2%
17.5%
12.4%
DV TFs
78.3%
74.0%
86.1%
88.8%
87.7%
85.5%
83.7%
AP TFs broad set
60.0%
51.6%
75.5%
79.0%
75.1%
71.8%
76.9%
AP TFs narrow set
37.7%
27.9%
55.5%
58.8%
52.2%
50.2%
61.0%
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of intronic and intergenic ATAC-seq peaks overlapped by ChIP signal

Supplemental Figure S11. Differential high-confidence ATAC-seq peaks display strong features of axis patterning enhancers. Extension of Fig. 4AB with all classes of accessible regions: all high-confidence ATAC-seq peaks (all peaks), constitutive peaks, differential peaks and four quarters of differential peaks (based on maximum $\log_2$ fold-change reported in DESeq2, Supplemental Fig. S10). (*A*) Proportional distribution of genomic annotations. UTR: 5' and 3' untranslated regions. CDS: coding sequence. (*B*) Bar plot represents the proportion of different classes of ATAC-seq peaks that map to intronic and intergenic regions (numbers of intervals in the legend) and co-localize with ChIP signal of different classes of proteins. ORI complex: ChIP-seq peaks of ORC2 (origin recognition complex subunit 2). Insulator proteins: ChIP-chip peaks of BEAF-32, CP190, CTCF and Su(Hw). DV TFs: ChIP-chip peaks of 4 maternal and zygotic DV TFs: Dorsal, Mothers against dpp, Snail and Twist. AP TFs (broad set): overlap with ChIP-chip peaks of 14 maternal, gap and pair-rule AP TFs: Bicoid, Caudal, Giant, Hunchback, Knirps, Kruppel, Huckebein, Tailless, Dichaete, Fushi-tarazu, Hairy, Paired, Runt and Sloppy-paired 1. AP TFs (narrow set): overlap with ChIP-seq peaks of 6 maternal and gap AP TFs: Bicoid, Caudal, Giant, Hunchback, Knirps, Kruppel.

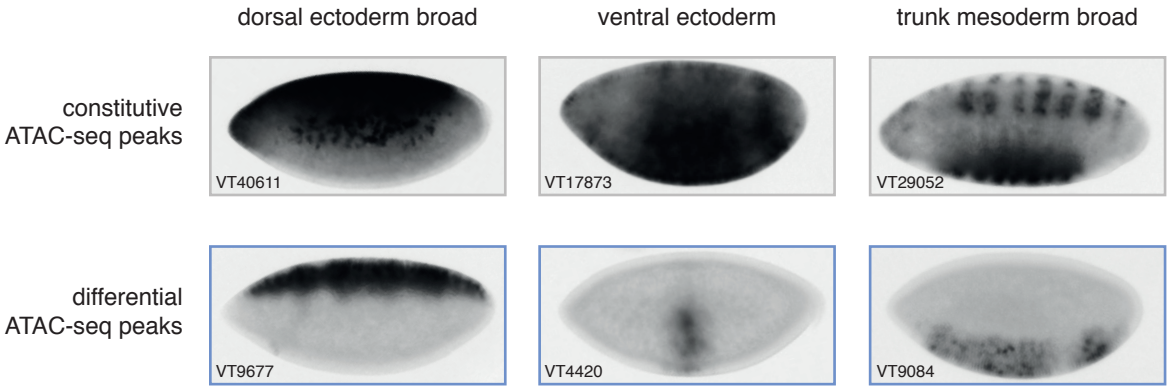
A


Vienna Tiles overlapping
differential ATAC-seq peaks (n = 613)
Number of VT tiles
300
250
200
150
100
50
0
1%
43%
14%
19%
12%
7%
3%
only stage 4-6
stage 4-6
stage 7-8
stage 9-10
stage 11-12
stage 13-14
stage 15-16
Vienna Tiles overlapping
constitutive ATAC-seq peaks (n = 711)
Number of VT tiles
300
250
200
150
100
50
0
1%
27%
9%
19%
20%
17%
6%
only stage 4-6
stage 4-6
stage 7-8
stage 9-10
stage 11-12
stage 13-14
stage 15-16
Complete set of Vienna Tiles
(n = 3436)
Number of VT tiles
1000
800
600
400
200
0
1%
16%
6%
19%
23%
23%
12%
only stage 4-6
stage 4-6
stage 7-8
stage 9-10
stage 11-12
stage 13-14
stage 15-16

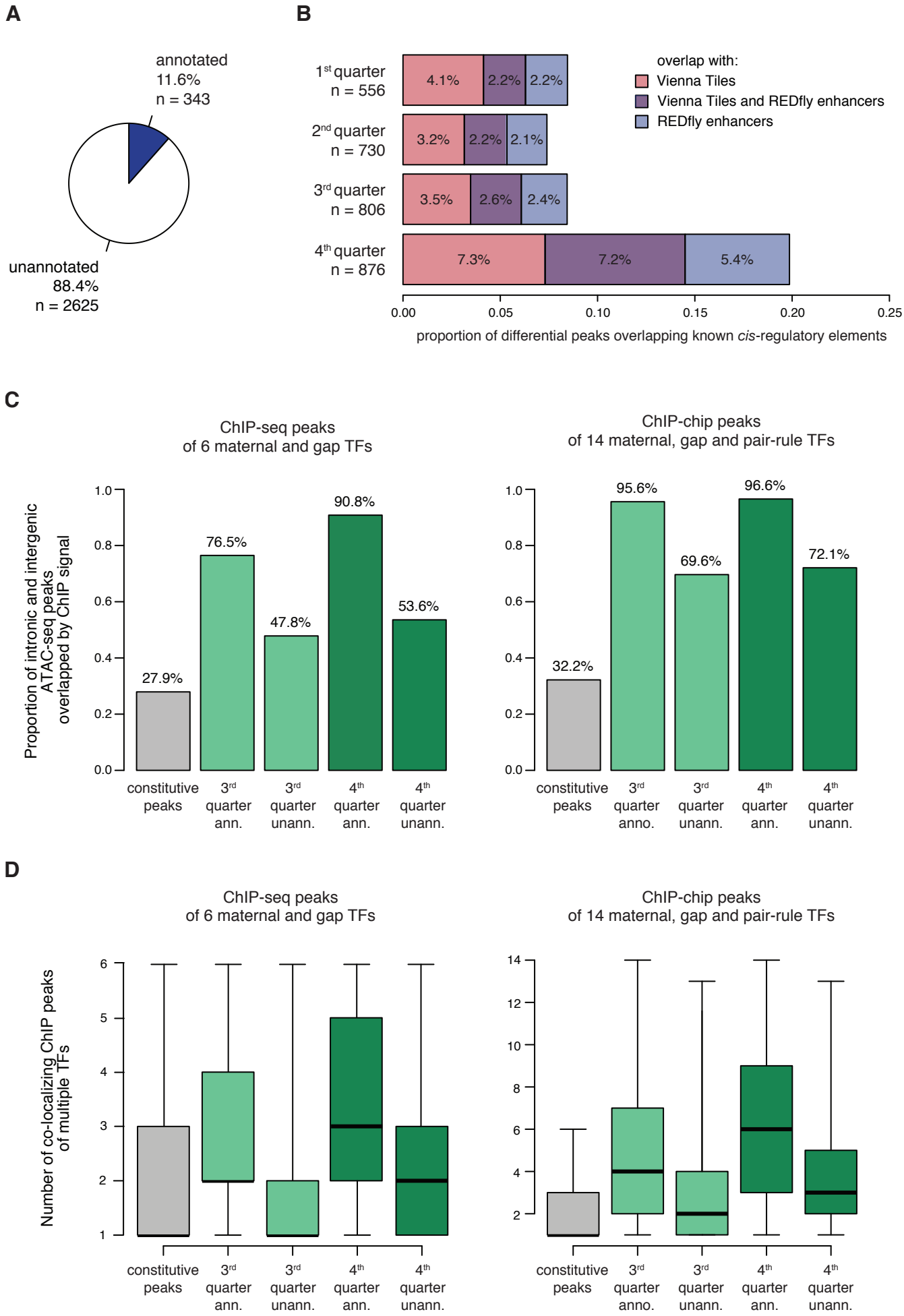
B


VT tiles overlapping:
constitutive peaks (n = 202)
differential peaks (n = 272)
anterior
gap
posterior
A–P stripe
A–P stripes
pair–rule
anterior endoderm
head mesoderm
procephalic ectoderm
hindgut
posterior endoderm
amnioserosa
dorsal ectoderm broad
dorsal ectoderm subset
ventral ectoderm
mesoderm
trunk mesoderm broad
trunk_mesoderm subset
ubiquitous
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Proportion of VT tiles active at stage 4-6 of embryogenesis

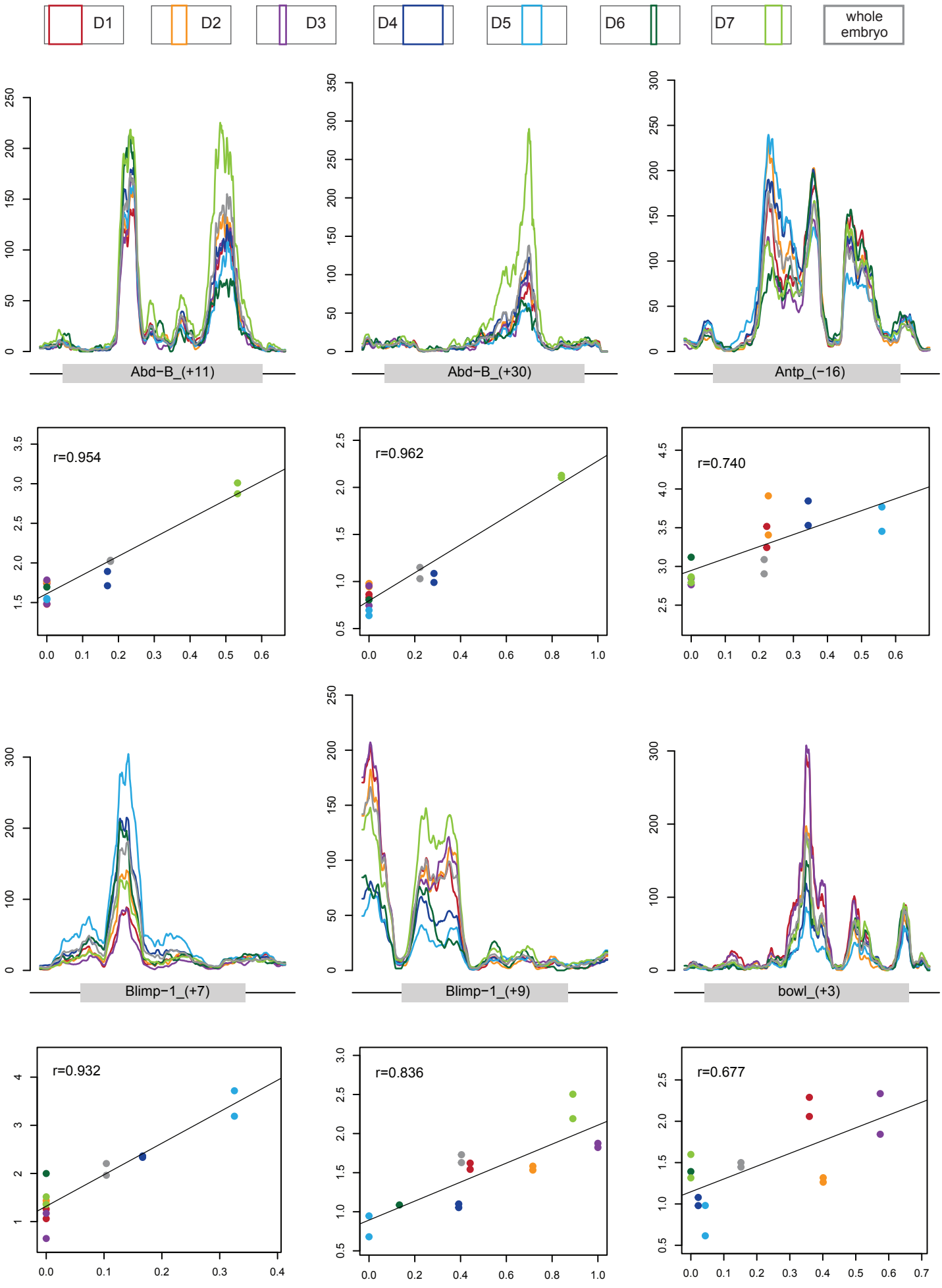
C

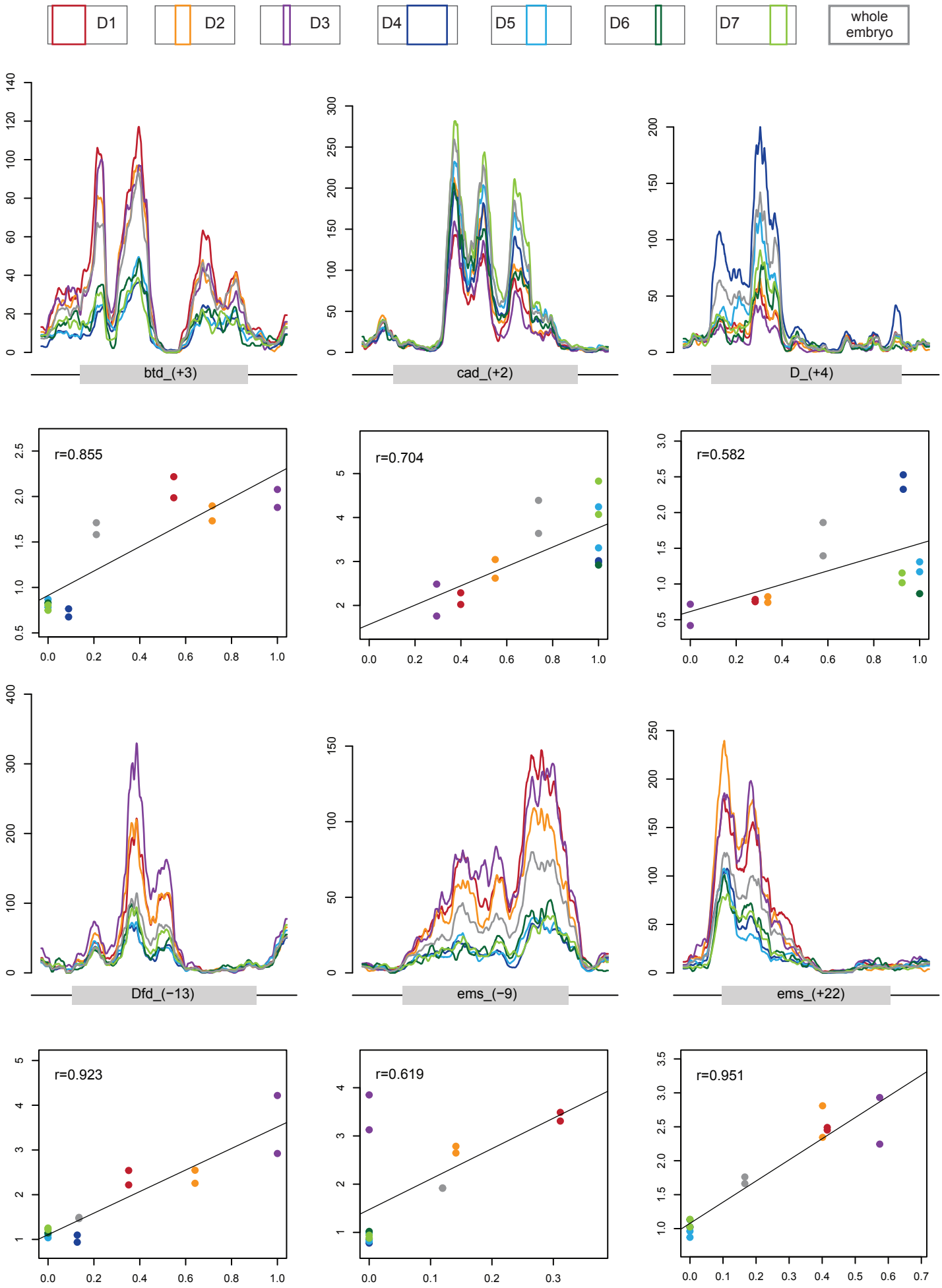

dorsal ectoderm broad
ventral ectoderm
trunk mesoderm broad
constitutive ATAC-seq peaks
differential ATAC-seq peaks
VT40611
VT17873
VT29052
VT9677
VT4420
VT9084

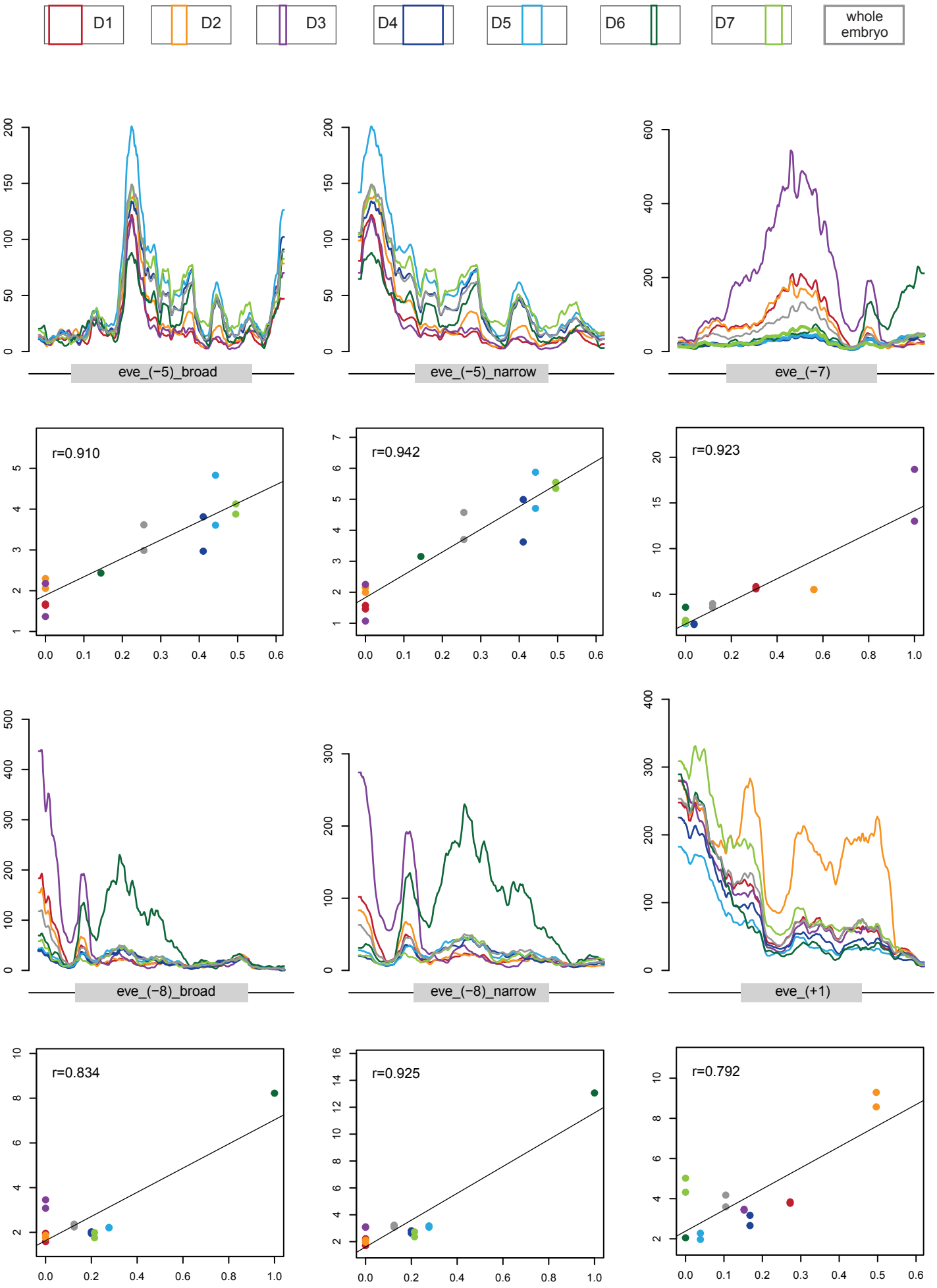
Supplemental Figure S12. Temporal and spatial activity of Vienna Tiles overlapping differential and constitutive ATAC-seq peaks. (*A*) Bar plots show the number of Vienna Tiles (VTs) active at individual stages of *Drosophila* embryogenesis, with indicated proportional representation among all elements from each considered category: VTs overlapping differential peaks, VTs overlapping constitutive peaks and a complete set of VTs. Since the reported enhancer activity of VTs usually spans several consecutive stages of embryogenesis, the elements were classified based on the earliest stage at which they drive expression of the reporter construct. (*B*) Bar plot shows the proportional representation of annotation terms among VTs that are specifically active in stage 4-6 embryos and overlap either differential peaks (blue) or constitutive peaks (grey). Note that individual Vienna Tiles are often annotated with multiple terms. (*C*) Examples of Vienna Tiles that drive patterned expression along the DV axis (annotation terms: "dorsal ectoderm broad", "ventral ectoderm", "trunk mesoderm broad"). Comparison between activity patterns of Vienna Tiles that overlap constitutive (grey frame) and differential (blue frame) ATAC-seq peaks. Note that the differential DV elements display more pronounced modulation of their activity along the dissected AP axis. Images downloaded from http://enhancers.starklab.org/.

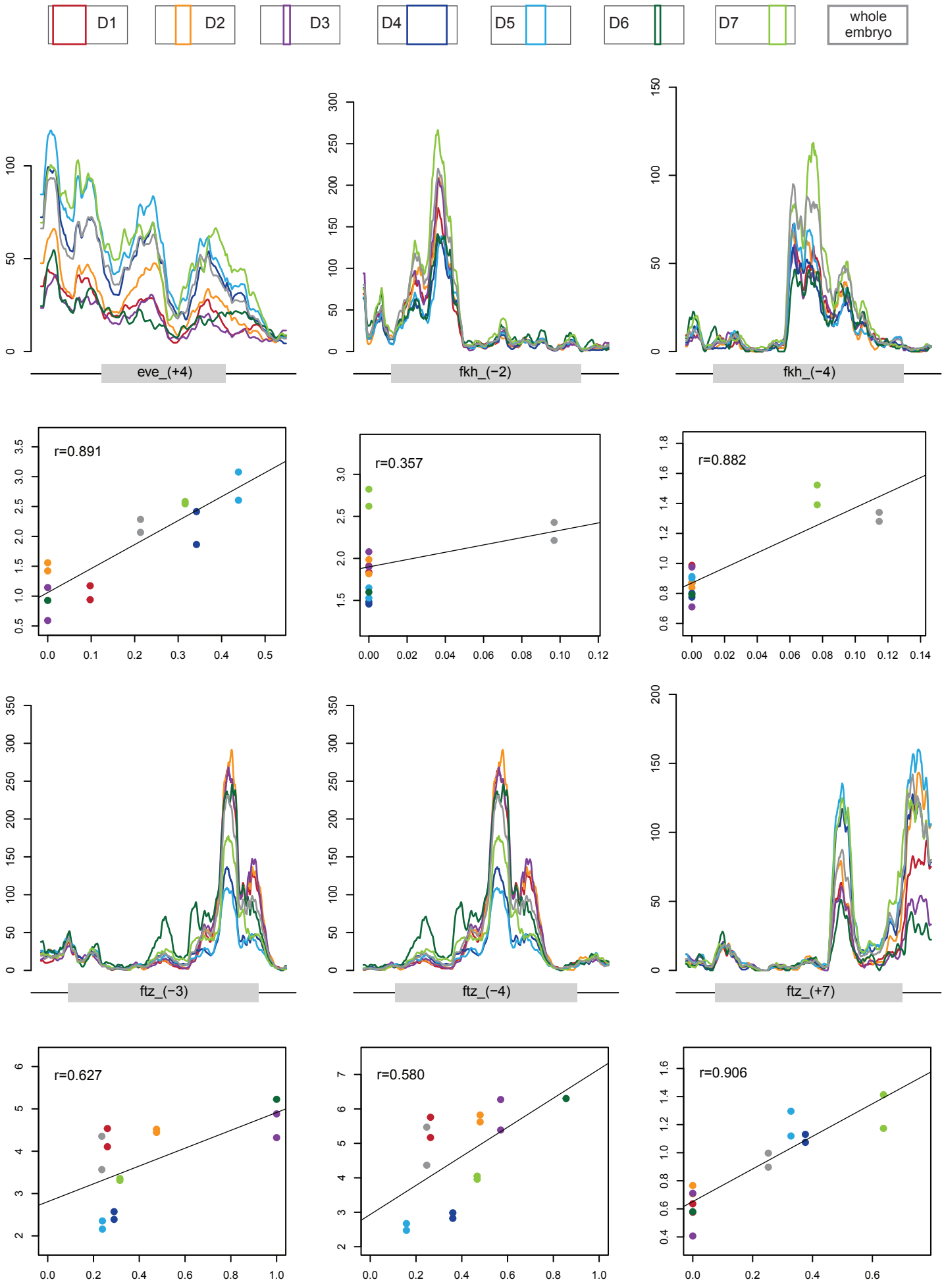

A
annotated
11.6%
n = 343
unannotated
88.4%
n = 2625
B
1st quarter
n = 556
2nd quarter
n = 730
3rd quarter
n = 806
4th quarter
n = 876
4.1%
2.2%
2.2%
3.2%
2.2%
2.1%
3.5%
2.6%
2.4%
7.3%
7.2%
5.4%
overlap with:
Vienna Tiles
Vienna Tiles and REDfly enhancers
REDfly enhancers
0.00
0.05
0.10
0.15
0.20
0.25
proportion of differential peaks overlapping known cis-regulatory elements
C
ChIP-seq peaks
of 6 maternal and gap TFs
Proportion of intronic and intergenic
ATAC-seq peaks
overlapped by ChIP signal
27.9%
76.5%
47.8%
90.8%
53.6%
constitutive peaks
3rd quarter ann.
3rd quarter unann.
4th quarter ann.
4th quarter unann.
ChIP-chip peaks
of 14 maternal, gap and pair-rule TFs
32.2%
95.6%
69.6%
96.6%
72.1%
constitutive peaks
3rd quarter anno.
3rd quarter unann.
4th quarter ann.
4th quarter unann.
D
ChIP-seq peaks
of 6 maternal and gap TFs
Number of co-localizing ChIP peaks
of multiple TFs
ChIP-chip peaks
of 14 maternal, gap and pair-rule TFs

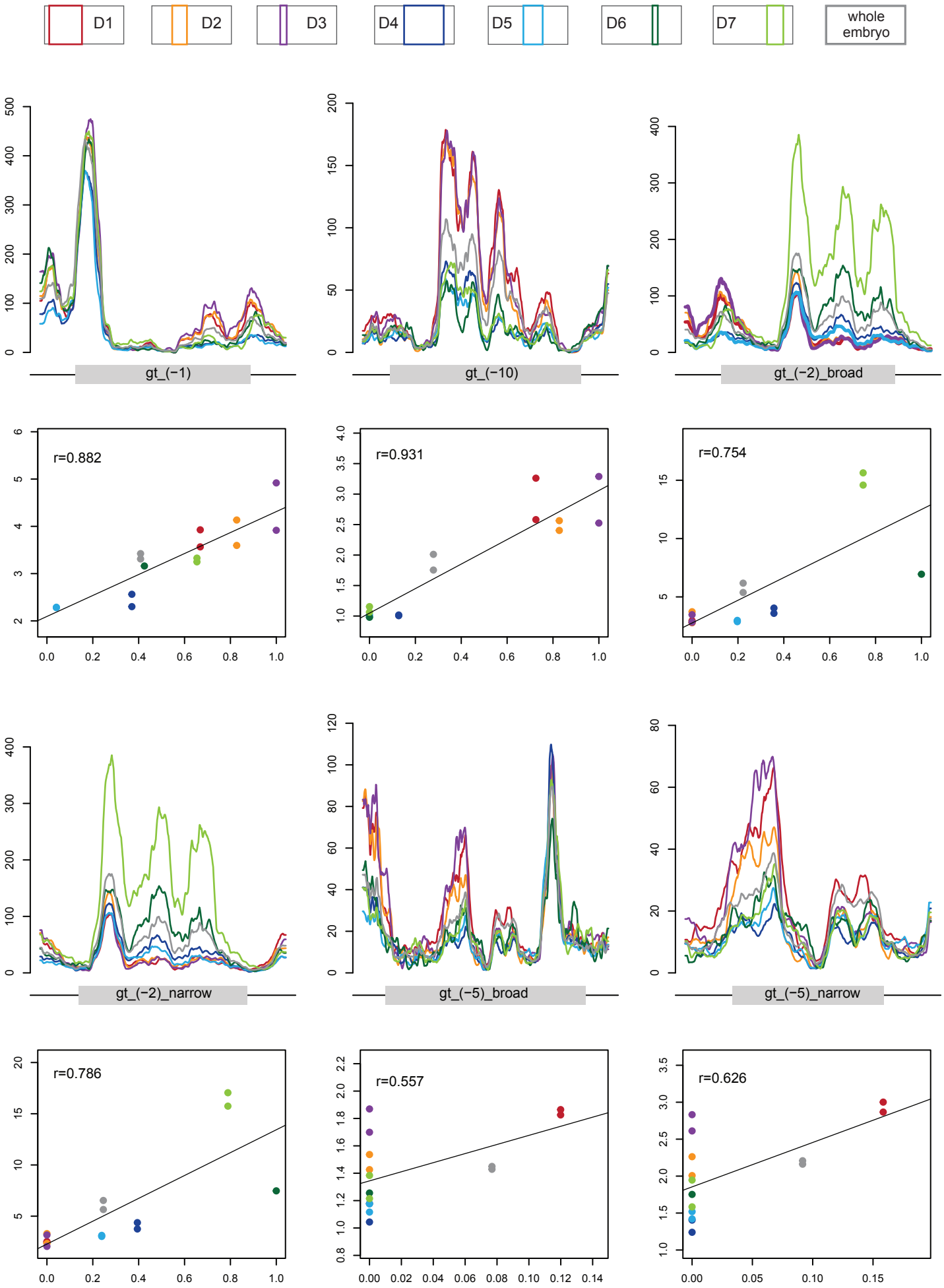
Supplemental Figure S13. Comparison of annotated and unannotated differential peaks. Differential peaks were divided into those that overlapped (ann.: annotated) and did not overlap (unann.: unannotated) known *cis*-regulatory elements: Vienna Tiles active at stage 4-6 and REDfly elements active in blastoderm embryos (including 88 AP enhancers). Only peaks mapping to intronic and intergenic regions were considered. (*A*) Pie chart represents the proportion and number of annotated (blue) and unannotated (white) intervals among all intronic and intergenic differential peaks. (*B*) Bar plot shows the proportion of differential peaks from each quarter that are represented by Vienna Tiles (red), REDfly enhancers (blue), both Vienna Tiles and REDfly enhancers (purple). The total number of intronic and intergenic intervals in each quarter is indicated. (*C,D*) Comparison of constitutive peaks with annotated and unannotated differential peaks from the 3rd and 4th quarters. The peaks were overlapped with ChIP-seq peaks of 6 maternal and gap AP TFs (left; narrow set from Fig. 4B) and ChIP-chip peaks of 14 maternal, gap and pair-rule AP TFs (right; broad set from Fig. 4B). 3rd quarter: 68 annotated and 738 unannotated peaks. 4th quarter: 174 annotated and 702 unannotated peaks. (C) Bar plots represent the proportion of differential and constitutive ATAC-seq peaks that are overlapped by ChIP signal of at least one TF. Note that the unannotated differential peaks are more frequently targeted by patterning TFs than the constitutive ATAC-seq peaks. (D) Boxplots represent the number of different TFs that co-bind within the ATAC-seq peaks. Peaks with no overlap were excluded. Note that combinatorial regulation by multiple TFs is a distinguishing feature of axis patterning enhancers.

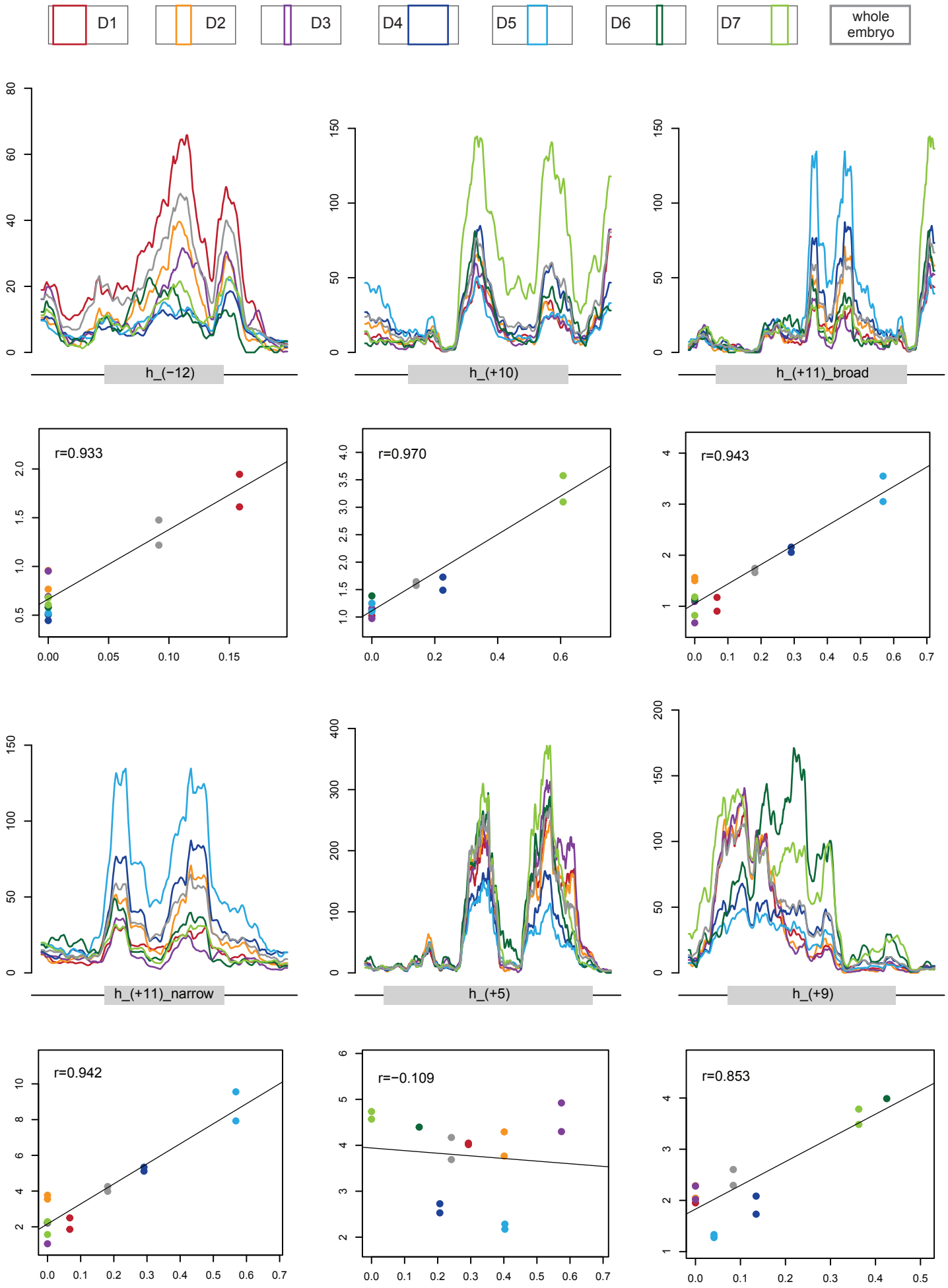

D1
D2
D3
D4
D5
D6
D7
whole embryo
Abd−B_(+11)
Abd−B_(+30)
Antp_(−16)
r=0.954
r=0.962
r=0.740
Blimp−1_(+7)
Blimp−1_(+9)
bowl_(+3)
r=0.932
r=0.836
r=0.677

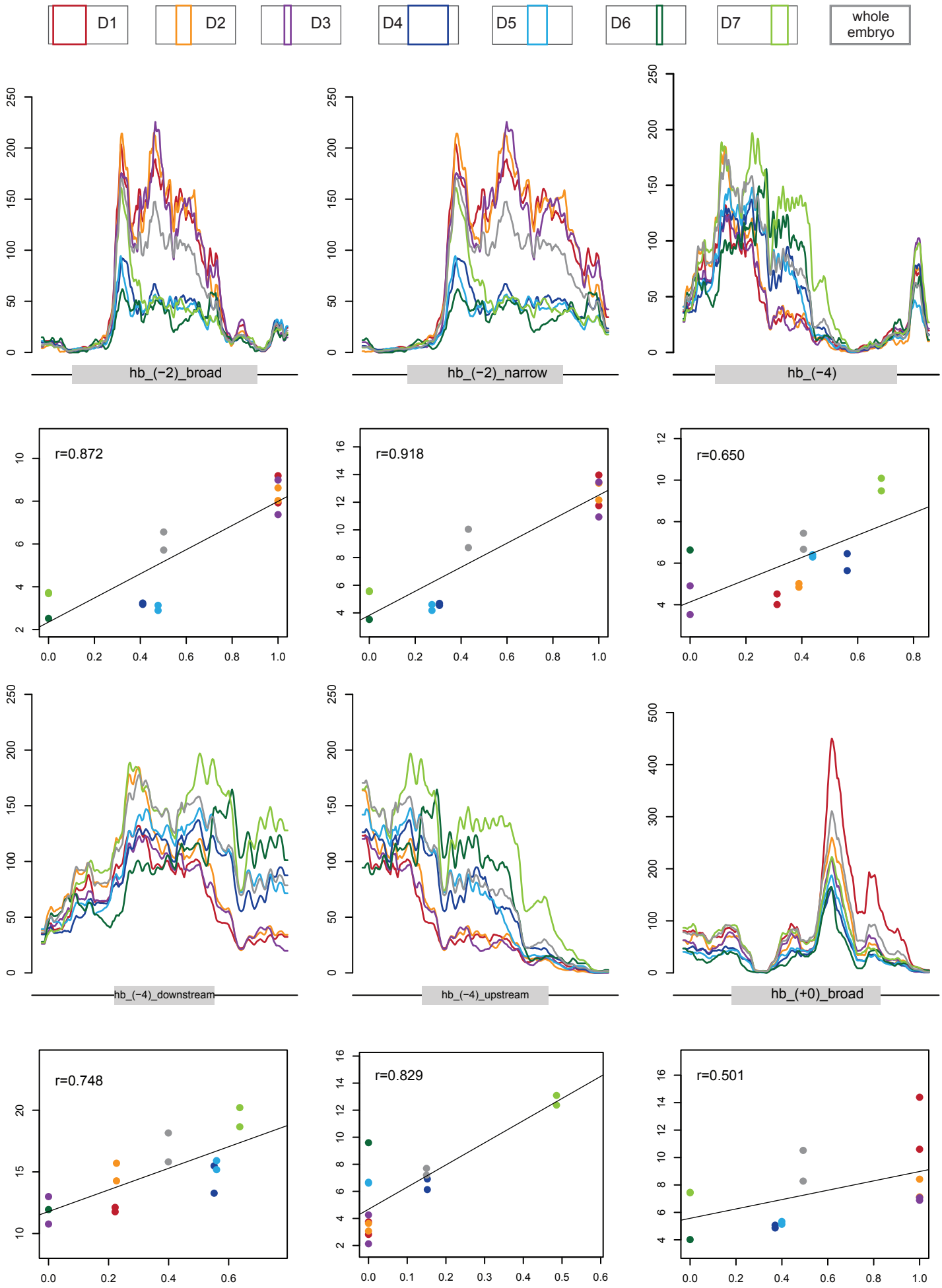

D1
D2
D3
D4
D5
D6
D7
whole embryo
btd_(+3)
cad_(+2)
D_(+4)
r=0.855
r=0.704
r=0.582
Dfd_(−13)
ems_(−9)
ems_(+22)
r=0.923
r=0.619
r=0.951

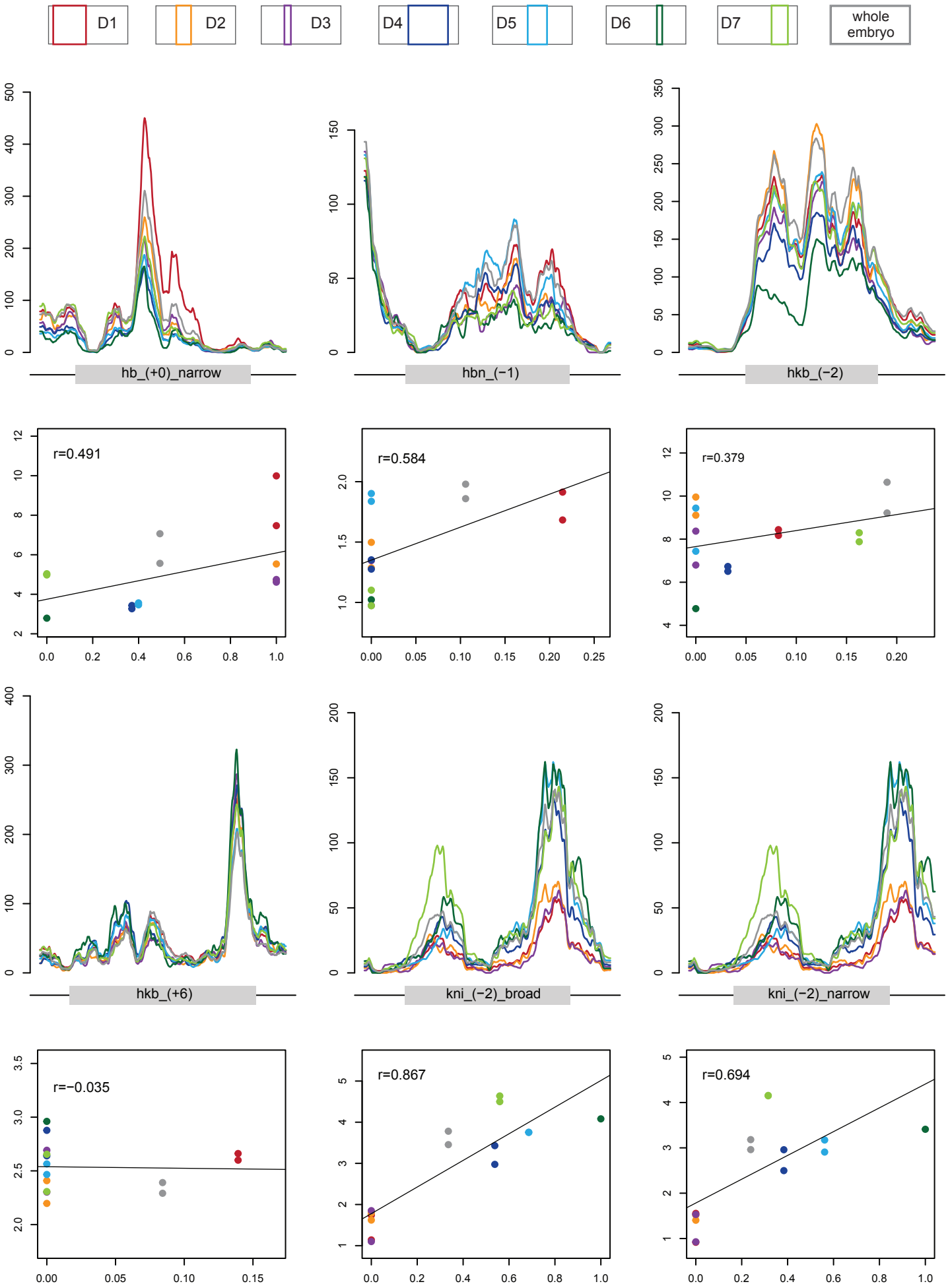

D1
D2
D3
D4
D5
D6
D7
whole embryo
eve_(−5)_broad
eve_(−5)_narrow
eve_(−7)
r=0.910
r=0.942
r=0.923
eve_(−8)_broad
eve_(−8)_narrow
eve_(+1)
r=0.834
r=0.925
r=0.792

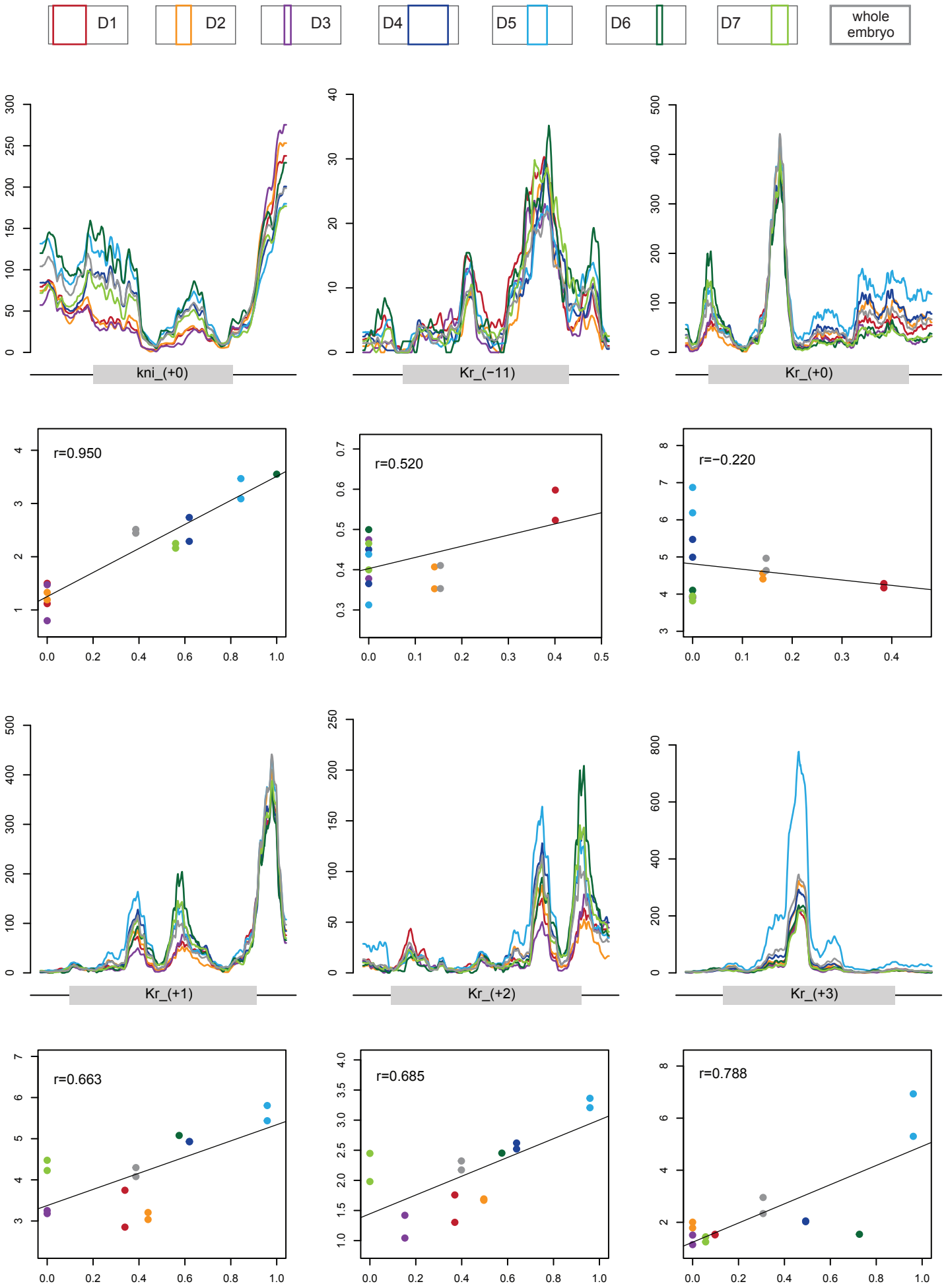

D1
D2
D3
D4
D5
D6
D7
whole embryo
eve_(+4)
fkh_(−2)
fkh_(−4)
r=0.891
r=0.357
r=0.882
ftz_(−3)
ftz_(−4)
ftz_(+7)
r=0.627
r=0.580
r=0.906

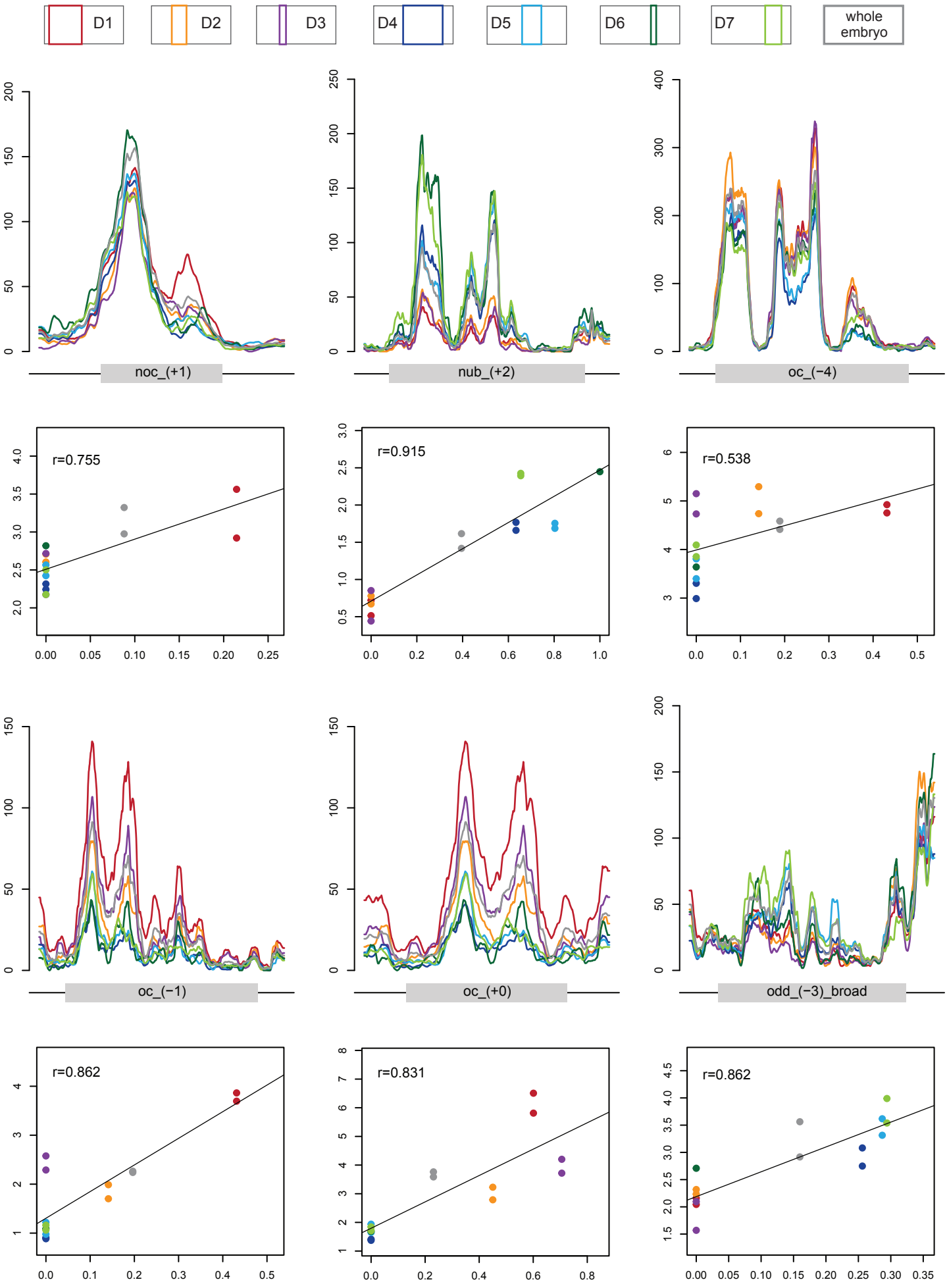

D1
D2
D3
D4
D5
D6
D7
whole embryo
gt_(−1)
gt_(−10)
gt_(−2)_broad
r=0.882
r=0.931
r=0.754
gt_(−2)_narrow
gt_(−5)_broad
gt_(−5)_narrow
r=0.786
r=0.557
r=0.626

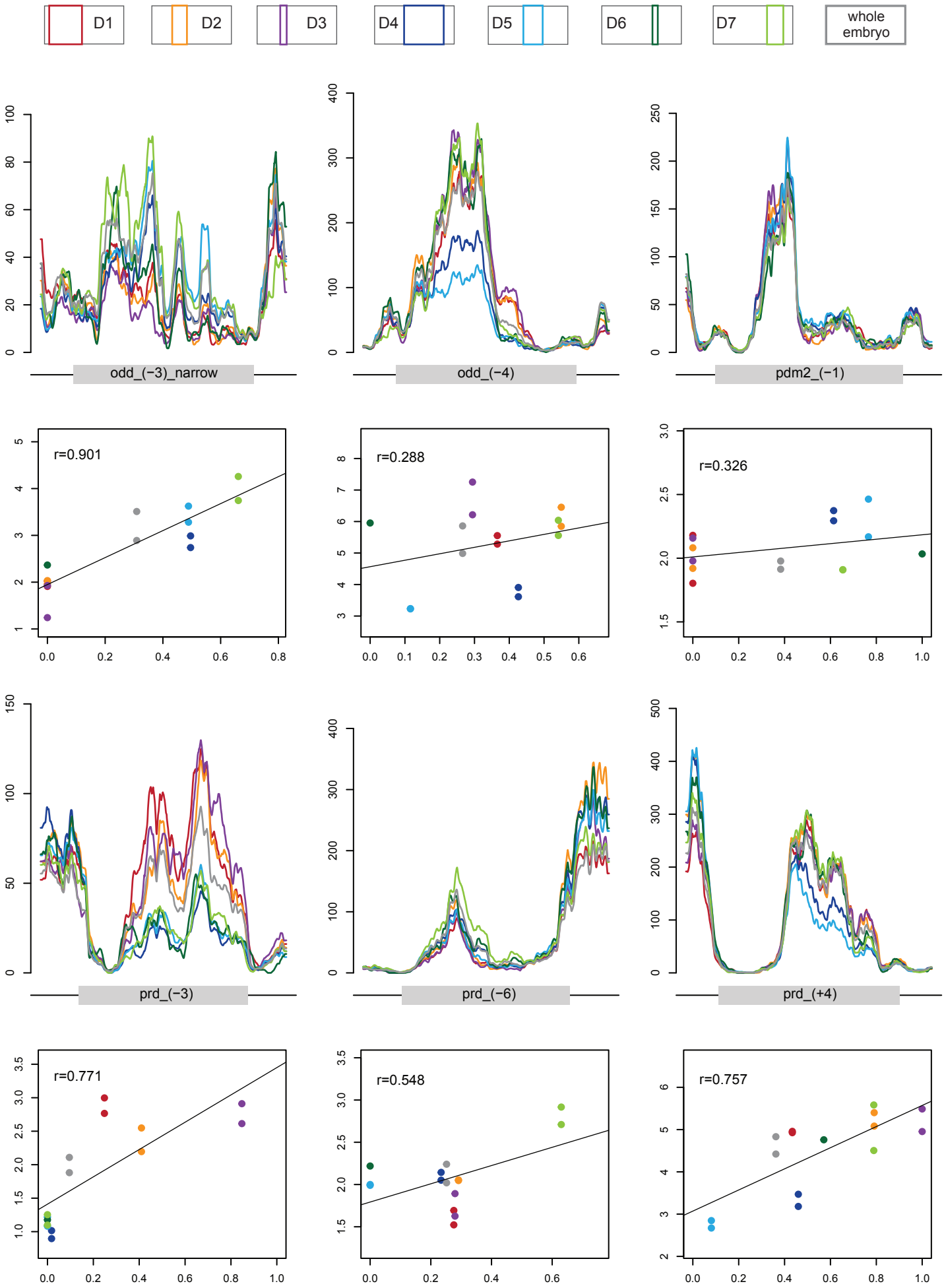

D1
D2
D3
D4
D5
D6
D7
whole embryo
h_(−12)
h_(+10)
h_(+11)_broad
r=0.933
r=0.970
r=0.943
h_(+11)_narrow
h_(+5)
h_(+9)
r=0.942
r=−0.109
r=0.853

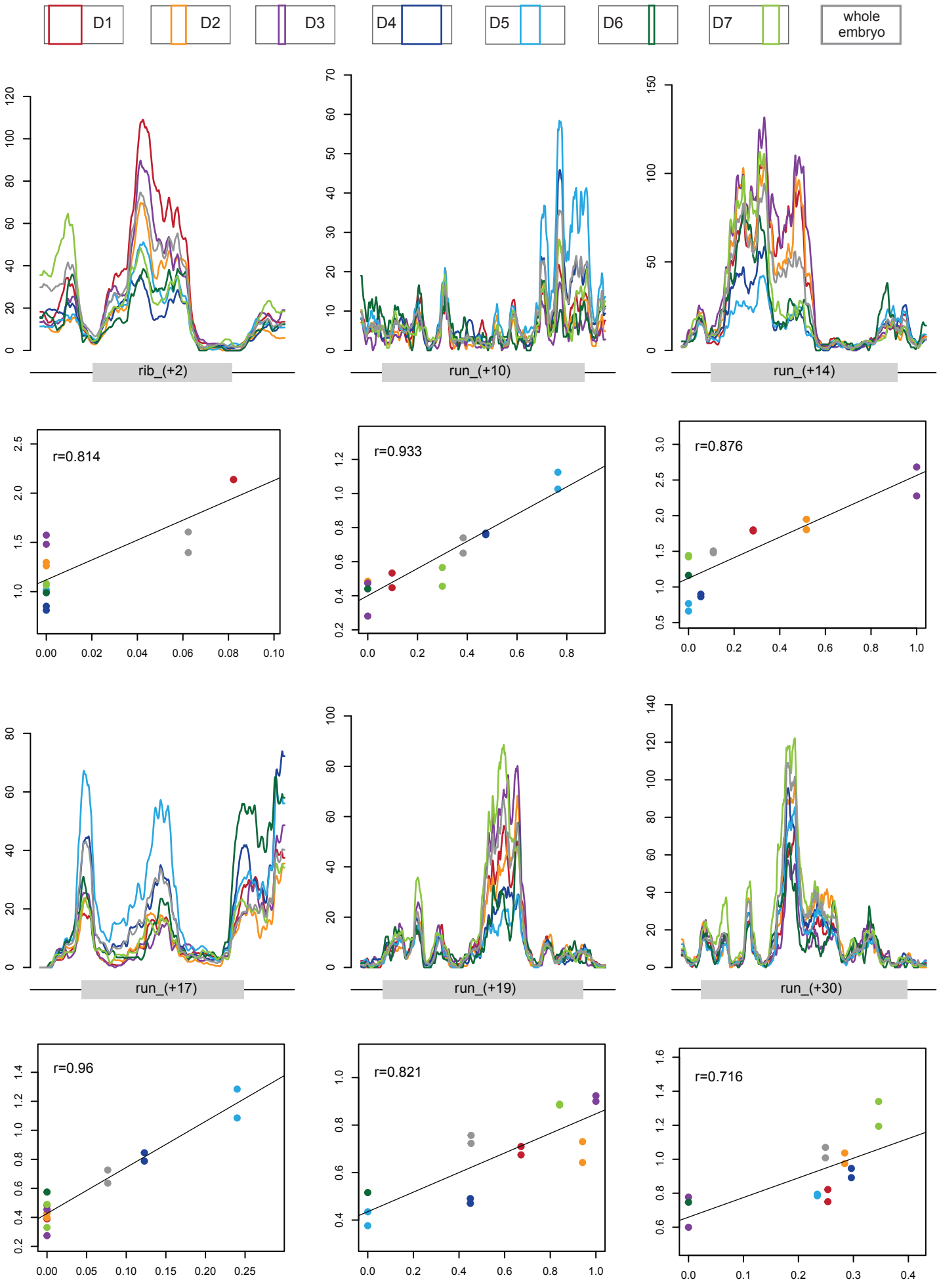

D1
D2
D3
D4
D5
D6
D7
whole embryo
hb_(−2)_broad
hb_(−2)_narrow
hb_(−4)
r=0.872
r=0.918
r=0.650
hb_(−4)_downstream
hb_(−4)_upstream
hb_(+0)_broad
r=0.748
r=0.829
r=0.501

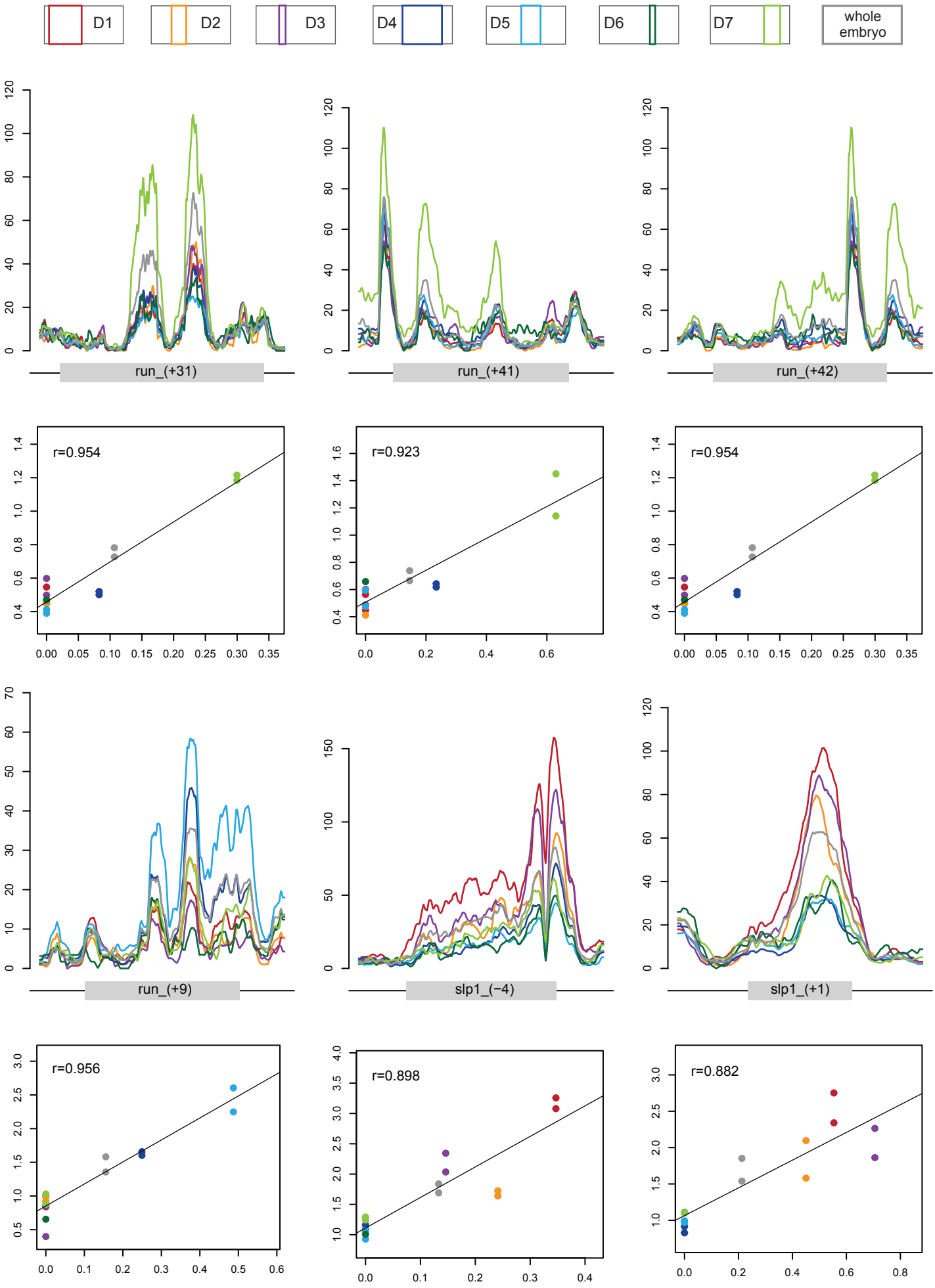

D1
D2
D3
D4
D5
D6
D7
whole embryo
hb_(+0)_narrow
hbn_(−1)
hkb_(−2)
r=0.491
r=0.584
r=0.379
hkb_(+6)
kni_(−2)_broad
kni_(−2)_narrow
r=−0.035
r=0.867
r=0.694

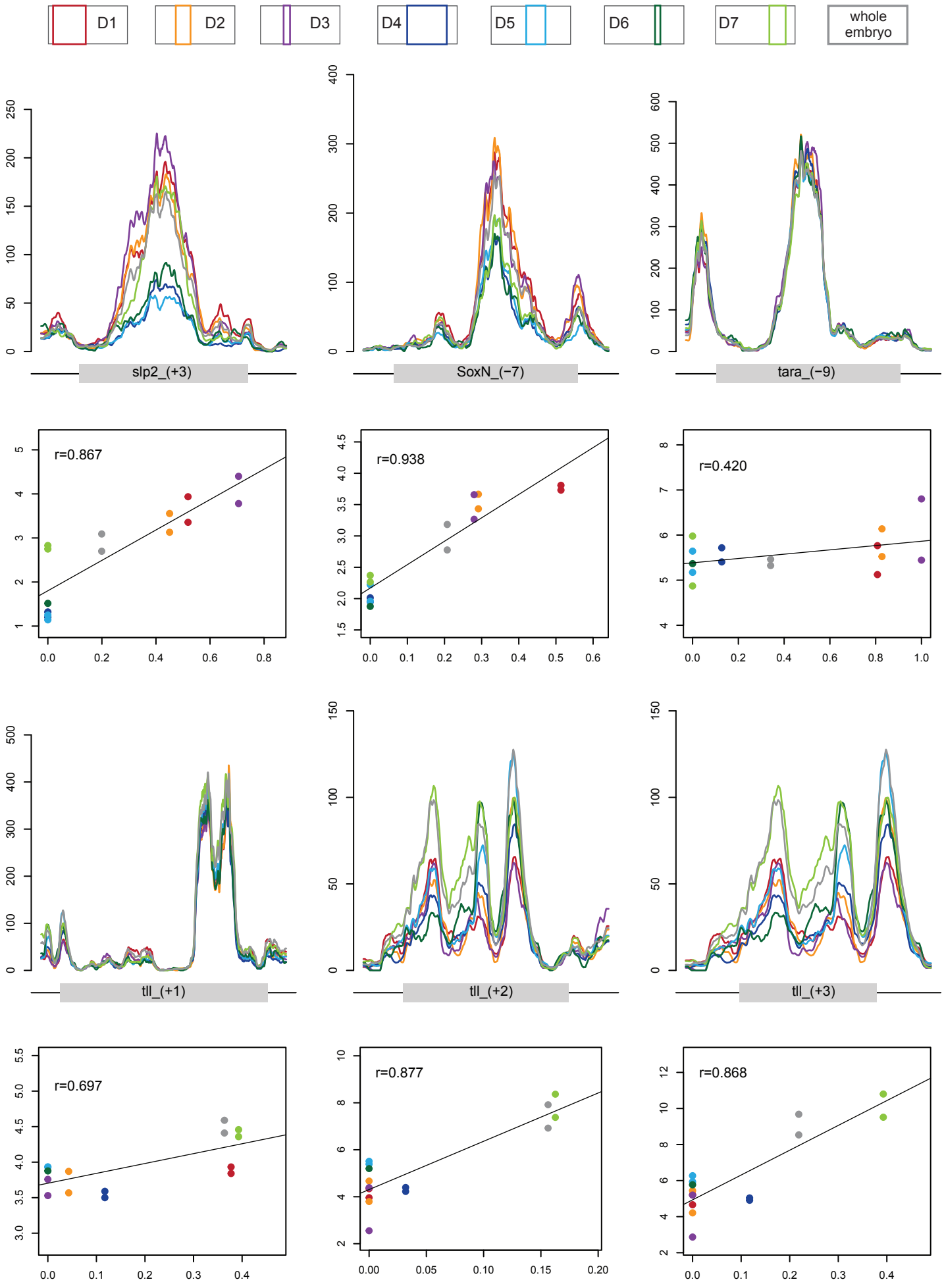

D1
D2
D3
D4
D5
D6
D7
whole embryo
kni_(+0)
Kr_(−11)
Kr_(+0)
r=0.950
r=0.520
r=−0.220
Kr_(+1)
Kr_(+2)
Kr_(+3)
r=0.663
r=0.685
r=0.788

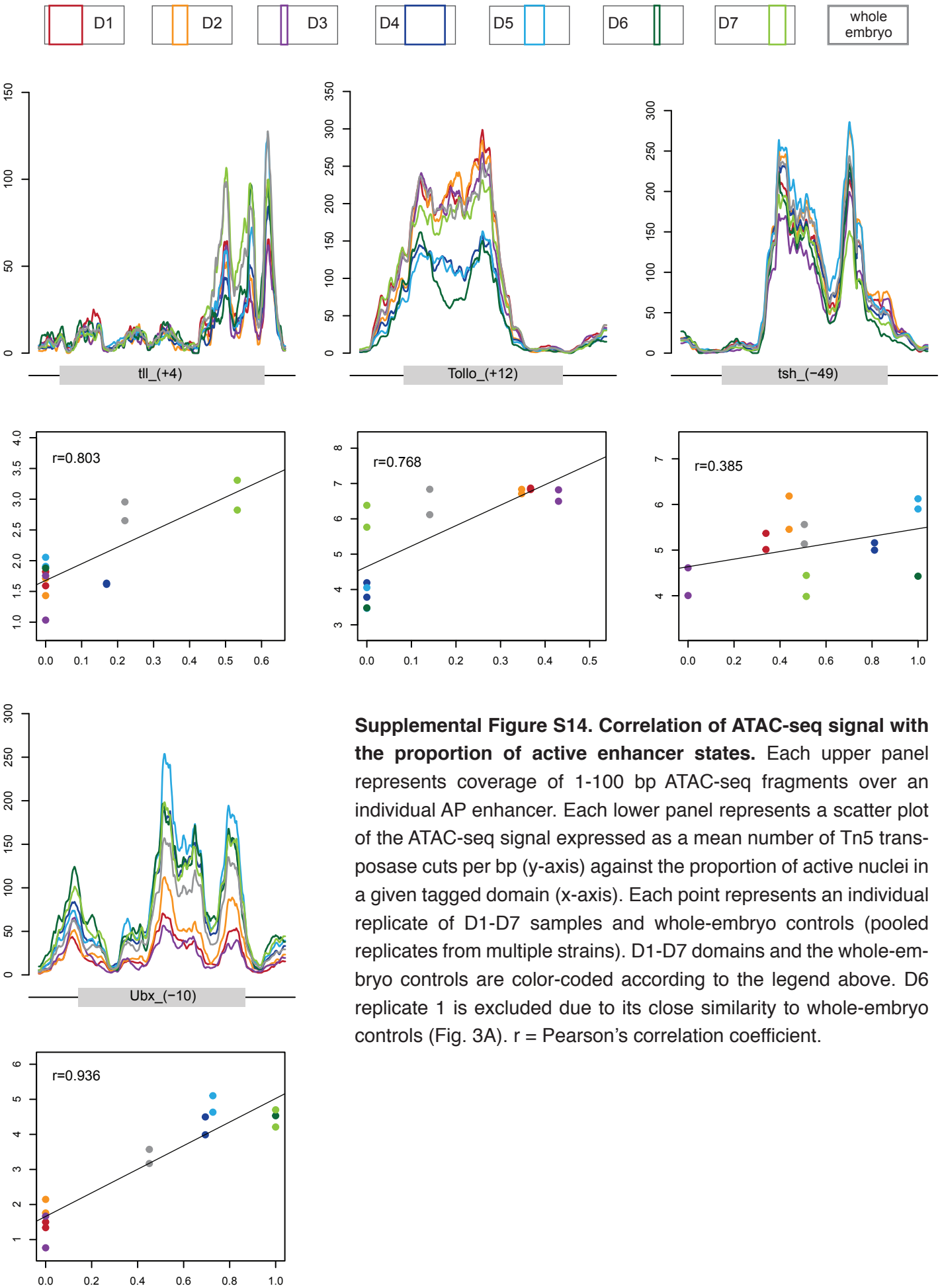

D1
D2
D3
D4
D5
D6
D7
whole embryo
noc_(+1)
nub_(+2)
oc_(−4)
r=0.755
r=0.915
r=0.538
oc_(−1)
oc_(+0)
odd_(−3)_broad
r=0.862
r=0.831
r=0.862


D1
D2
D3
D4
D5
D6
D7
whole embryo
odd_(−3)_narrow
odd_(−4)
pdm2_(−1)
r=0.901
r=0.288
r=0.326
prd_(−3)
prd_(−6)
prd_(+4)
r=0.771
r=0.548
r=0.757

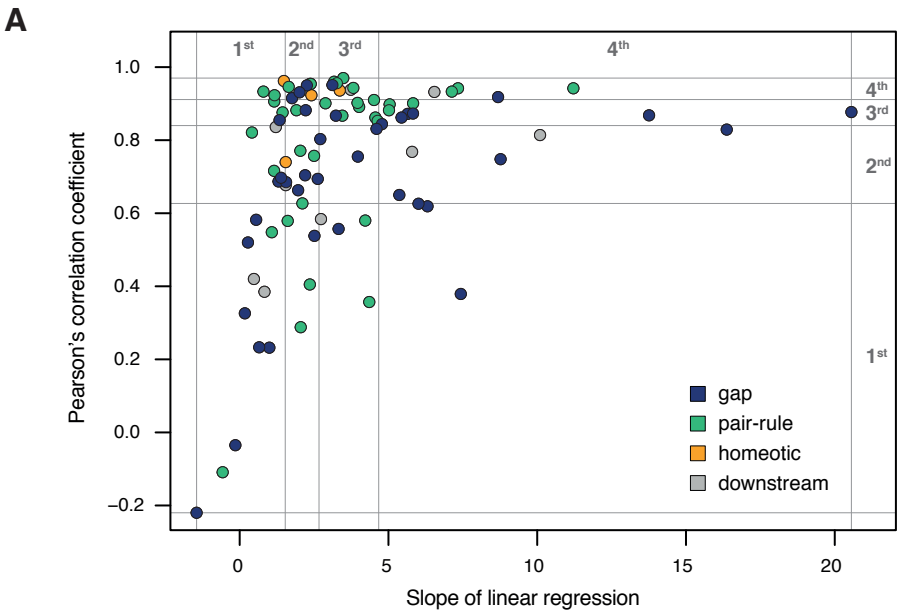

D1
D2
D3
D4
D5
D6
D7
whole embryo
rib_(+2)
run_(+10)
run_(+14)
r=0.814
r=0.933
r=0.876
run_(+17)
run_(+19)
run_(+30)
r=0.96
r=0.821
r=0.716

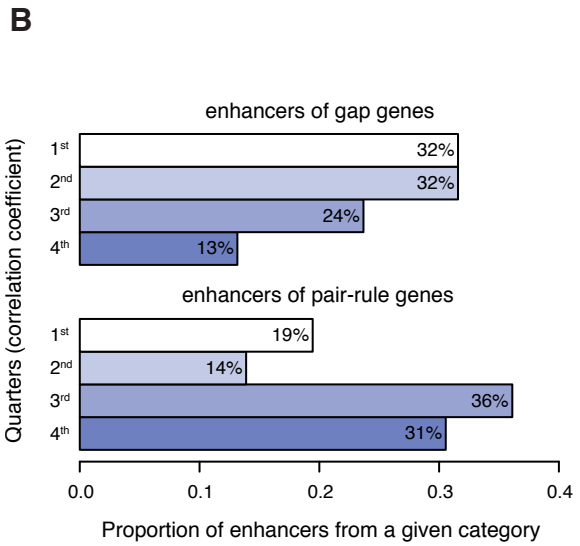

D1
D2
D3
D4
D5
D6
D7
whole embryo
run_(+31)
run_(+41)
run_(+42)
r=0.954
r=0.923
r=0.954
run_(+9)
slp1_(−4)
slp1_(+1)
r=0.956
r=0.898
r=0.882

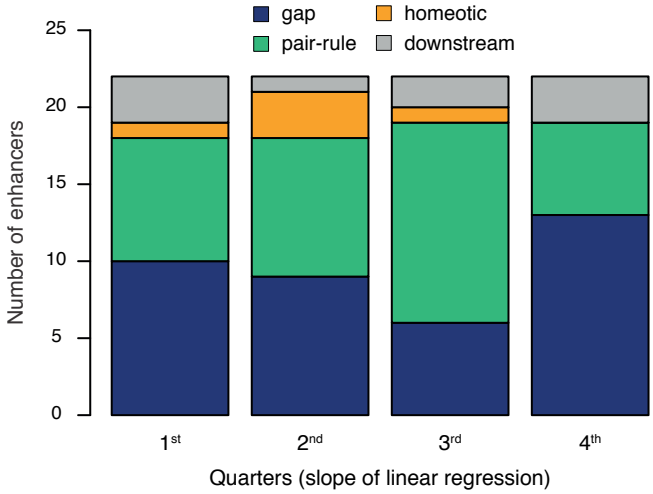

D1
D2
D3
D4
D5
D6
D7
whole embryo
slp2_(+3)
SoxN_(−7)
tara_(−9)
r=0.867
r=0.938
r=0.420
tll_(+1)
tll_(+2)
tll_(+3)
r=0.697
r=0.877
r=0.868

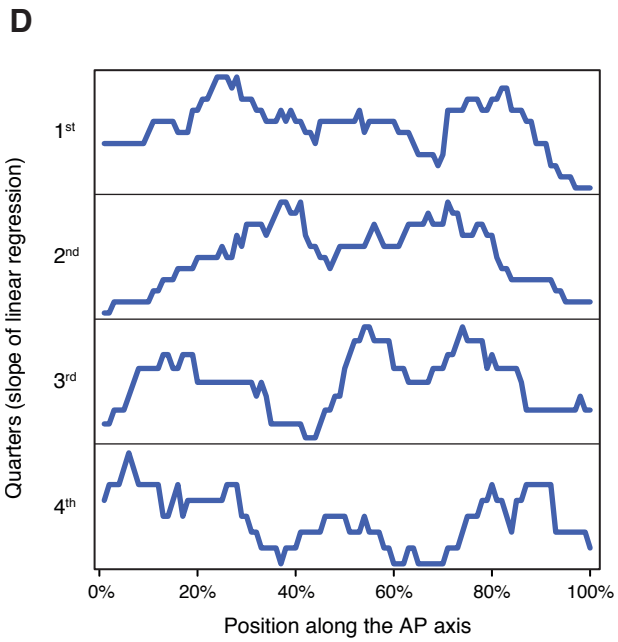


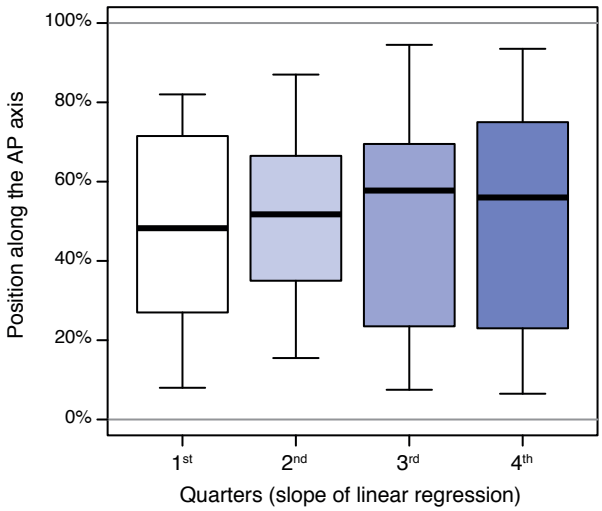
Supplemental Figure S14. Correlation of ATAC-seq signal with the proportion of active enhancer states. Each upper panel represents coverage of 1-100 bp ATAC-seq fragments over an individual AP enhancer. Each lower panel represents a scatter plot of the ATAC-seq signal expressed as a mean number of Tn5 transposase cuts per bp (y-axis) against the proportion of active nuclei in a given tagged domain (x-axis). Each point represents an individual replicate of D1-D7 samples and whole-embryo controls (pooled replicates from multiple strains). D1-D7 domains and the whole-embryo controls are color-coded according to the legend above. D6 replicate 1 is excluded due to its close similarity to whole-embryo controls (Fig. 3A). r = Pearson's correlation coefficient.


A
1st
2nd
3rd
4th
Pearson's correlation coefficient
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0
5
10
15
20
Slope of linear regression
gap
pair-rule
homeotic
downstream


B
enhancers of gap genes
1st
2nd
3rd
4th
32%
32%
24%
13%
enhancers of pair-rule genes
19%
14%
36%
31%
Quarters (correlation coefficient)
0.0
0.1
0.2
0.3
0.4
Proportion of enhancers from a given category


C
gap
pair-rule
homeotic
downstream
Number of enhancers
25
20
15
10
5
0
1st
2nd
3rd
4th
Quarters (slope of linear regression)


D
Quarters (slope of linear regression)
1st
2nd
3rd
4th
0%
20%
40%
60%
80%
100%
Position along the AP axis


E
Position along the AP axis
100%
80%
60%
40%
20%
0%
1st
2nd
3rd
4th
Quarters (slope of linear regression)

Supplemental Figure S15. Linear correlation between accessibility and transcriptional activity: properties of different classes of AP enhancers. Note that the analysis considered in this figure includes corrections of ATAC-seq signal over enhancers that were used as drivers of the nuclear marker in their respective tagged domains (Supplemental Methods). (*A*) Distribution of the slope of linear regression (x-axis) and Pearson's correlation coefficient (y-axis) across 88 AP enhancers (Supplemental Table S8). Each enhancer is represented as an individual data point and color-coded according to the position of its target gene in the AP regulatory network (Supplemental Methods). Four classes of target genes are considered: gap (dark blue), pair-rule (green), homeotic (orange) and secondary downstream genes (grey). Quantiles of distributions of the regression slope and correlation coefficient are marked with grey lines; corresponding quarters are marked in grey above and on the right. (*B*) Comparison of the strength of linear correlation between enhancers of gap and pair-rule genes. Bar plots show the proportion of enhancers from each class that fall into different quarters of the correlation coefficient distribution. (*C*) Comparison of the magnitude of accessibility changes (i.e. linear regression slope) between enhancers of gap (dark blue), pair-rule (green), homeotic (orange) and secondary downstream (grey) genes. Bar plot summarizes representation of the four classes of enhancers in individual quarters of the regression slope distribution. (*D*) Comparison of the AP position of activity domains between enhancers that display different magnitudes of accessibility changes. Profiles show the relative frequency with which each % of the AP axis is represented by activity patterns of 22 enhancers from individual quarters of the distribution of the linear regression slope. (*E*) In order to correct for different sizes of activity domains, positions of their midpoints along the AP axis are compared. Boxplots represent distribution of the midpoints within individual quarters (defined for the linear regression slope). 0%: anterior tip, 100%: posterior tip of the embryo.

Enhancers from different tiers of the AP gene regulatory network do not display any significant bias regarding the strength of linear correlation (Pearson's correlation coefficient) or the magnitude of accessibility changes between active and inactive enhancer states (slope of linear regression). Additionally, anterior and posterior elements are not characterized by any distinguishing properties. In fact, there is no correlation between the magnitude of accessibility changes displayed by the enhancers and the position of their activity patterns along the AP axis.

Supplemental Methods

Immunostaining of transgenic embryos

Specific localization of the nuclear tag in transgenic embryos was validated with anti-FLAG antibody staining. In short, embryos were collected in population cages (broad window: 0-4 h AEL), dechorionated in 50% bleach (2 min), fixed for 20 min in 4% formaldehyde–heptane and devitellinised with methanol. Tagged domains were visualized by anti-FLAG immunostaining according to standard procedures (Müller 2008). Primary antibody: monoclonal anti-FLAG M2 (Sigma-Aldrich: F1804) in 1:500 dilution. Secondary antibody: Alexa Fluor 647-conjugated goat anti-mouse antibody (Jackson ImmunoResearch: 115-605-166) in 1:500 dilution. The embryos were imaged under LSM 710 Carl Zeiss laser-scanning confocal microscope.

Positions of tagged domains (Supplemental Table S2) were measured by projection of the anti-FLAG signal onto the axis connecting the anteriormost and posteriormost tips of the embryo, and were expressed as percent of the axis length (1-100%).

Definitions of the domains show high agreement with previously reported expression patterns of the enhancers (Segal et al. 2008), except for D1, D2 and D7 domains. Compared to the activity pattern of *hb_anterior_actv*, D1 is characterized by retraction of the tag from the anterior tip; however, this pattern is in agreement with reduced levels of native *hunchback* mRNA in the anteriormost region of the blastoderm (Tomancak et al. 2002). UNC84-3xFLAG is expressed in a wider D2 domain than the endogenous stripe 2 of *even skipped*, which is consistent with the *minimal stripe 2 element* initially driving expression in a broad anterior region at the beginning of stage 5 (Small et al. 1992). D7 shows strong posterior expansion in comparison to the reported activity domain of *gt_(-3)* enhancer. This is consistent with posterior extension of the endogenous expression domain of *giant* gene at the beginning of stage 5, which gradually gets sharpened and anteriorly shifted (Eldon and Pirrotta 1991).

Replicates

Homozygous lines representing independent insertions of the same construct in the attP2 site were used as biological duplicates, except for the eve_stripe1_DSCP_UNC84-3xFLAG construct (D3 domain) for which only one transgenic line was obtained (Supplemental Table S3).

Affinity-purification of tagged nuclei

Staged embryos were homogenized on ice by 10 strokes of Dounce homogenizer (tight pestle) in Buffer A (15 mM Tris-HCl pH = 7.4, 60 mM KCl, 15 mM NaCl, 0.1 mM EGTA, 5 mM MgCl2 and EDTA-free protease inhibitor (Roche: 04693132001). Homogenate was filtered through 10-µm nylon net filter (Merck Millipore) and centrifuged at 500 g for 7 min at 4°C. The nuclear pellet was resuspended in Buffer A + 0.5% NP-40 (Tergitol, Sigma), incubated on ice for 3 min and centrifuged at 500 g for 7 min at 4°C. Purified nuclei were resuspended in 500 µl of Buffer A. 50 µl of Dynabeads protein G (ThermoFisher: 10009D) were adsorbed to 1 µl of

monoclonal anti-FLAG M2 antibody (Sigma-Aldrich: F1804) in 200 µl PBS + 0.02% Tween-20 for 1 h at 4°C. Magnetic beads were washed once with PBS + 0.02% Tween-20, resuspended in 100 µl of Buffer A and combined with the nuclei. After 30-min incubation with rotation at 4°C, the beads with tagged nuclei were bound to a magnetic stand, washed three times with Buffer A and resuspended in 700 µl of Buffer A.

Whole-embryo controls were generated from the same collections of staged embryos as isolations of tagged domains D1, D4, D5 and D7 (including duplicates). After homogenization and nuclei purification, the nuclear pellet was resuspended directly in 700 µl of Buffer A, excluding incubation with magnetic beads.

ATAC-seq on purified nuclei

In order to estimate the final yield of nuclear isolations, 25 µl of nuclear suspension (+beads) was combined with 10 µl of 10 mg/ml proteinase K and 1.25 µl of 10% SDS. Following incubation at 55°C for 20 min and vortexing at maximum speed for 5 min, concentration of genomic DNA was measured with Qubit 2.0 fluorometer (Invitrogen). Nuclear suspension corresponding to 340 ng of genomic DNA was used for each ATAC-seq reaction.

Fragmentation and amplification of ATAC-seq libraries were performed according to the standard protocol (Buenrostro et al. 2015). An appropriate volume of nuclear suspension (+beads) was pelleted at 500 g for 7 min at 4°C. The pellet was resuspended in Nextera Tagment DNA Buffer with 6 µl of Nextera Tn5 transposase (Illumina) in the final volume of 25 µl. After incubation at 37°C for 30 min, the reaction was terminated by addition of Buffer PB (Qiagen). Magnetic beads were bound to the magnetic stand and the supernatant was purified with Qiagen MinElute PCR Purification Kit, followed by elution in 10 µl Elution Buffer (10 mM Tris buffer, pH = 8). 5 µl of eluted DNA was combined with custom Nextera PCR primers (from Buenrostro et al. 2013) and NEBNext Q5 Hot Start HiFi PCR Master Mix (New England Biolabs) in the total volume of 50 µl, followed by amplification for 12 cycles (PCR program according to Buenrostro et al. 2015). Purification was performed with Agencourt AMPure XP beads (Beckman Coulter: A63881), using double size selection (left ratio: 2.0x, right ratio: 0.5x) according to the manufacturer's protocol. Libraries were eluted in 30 µl of 10 mM Tris-HCl, pH = 8.0. Paired-end sequencing was performed on Illumina HiSeq 1500. A summary of all ATAC-seq experiments and generated samples is provided in Supplemental Table S3.

ATAC-seq on genomic DNA

In order to control for potential sequence bias of Tn5 transposase, a reference ATAC-seq library was prepared from purified genomic DNA (gDNA). gDNA was isolated from larval tissues of a wild-type *Drosophila* strain, using a high-salt extraction method (Aljanabi 1997) with minor modifications: a step of vortexing was omitted to reduce shearing of gDNA and a final step of RNase digestion was included (with 50 ng/µl of Ambion RNaseA, 30-min incubation at 37°C). 12 ng of gDNA was combined with 2.5 µl of Tn5 transposase in Nextera Tagment DNA Buffer (final volume: 50 µl), followed by incubation at 37°C for 30 min. The reaction was directly terminated by purification with Qiagen MinElute PCR Purification Kit and elution in 11 µl of nuclease-free water. 10 µl of eluted DNA was combined with custom Nextera

PCR primers (from Buenrostro et al. 2013) and NEBNext Q5 Hot Start HiFi PCR Master Mix (New England Biolabs) in the total volume of 50 µl, followed by amplification for 12 cycles (according to Buenrostro et al. 2015). The library was subsequently purified and processed as other ATAC-seq libraries.

ATAC-seq data processing

After demultiplexing and removal of adaptor sequences, ATAC-seq reads were aligned to a reference genome (UCSC: dm3) using Bowtie2 v.2.2.9 (Langmead and Salzberg 2012), with the following settings: bowtie2 --local --very-sensitive-local. Mapped reads were filtered for mapping quality (q >10) and proper pairing (read paired & read mapped in proper pair) using SAMtools v.1.3.1 (Li et al. 2009), with the following settings: samtools view –f 0x3 – q 10.

As a general measure of sensitivity to Tn5 transposase fragmentation, ATAC-seq signal was defined as a number of transposase cuts mapping to each bp (or a total count of cuts mapping to a selected genomic interval). The cuts were defined as 5' ends of ATAC-seq reads, with additional shifting by +4 bp and -5 bp for reads mapping to the plus and minus strands, respectively (Buenrostro et al. 2013). Total pool of ATAC-seq fragments was considered, with no prior size selection.

For visualization purposes, ATAC-seq signal was alternatively defined as coverage of short digestion products (1-100 bp, with *in silico* size selection). ATAC-seq fragments shorter than 100 bp have been previously demonstrated to correspond to nucleosome-free regions, accessible to binding by TFs and other small proteins (Buenrostro et al. 2013; Schep et al. 2015).

Peak calling

ATAC-seq peaks were called with MACS2 v.2.1.1.2 (Zhang et al. 2008), based on the enrichment of Tn5 transposase cuts (smoothed by +/- 100 bp extension). ATAC-seq library from genomic DNA was used as a common control sample. Peak calling was performed separately for each replicate, with the q-value cut-off = 0.01 and the following non-default parameters: macs2 callpeak --keep-dup all -q 0.01 --nomodel --shift -100 --extsize 200 -f BAM -g dm -B.

A set of 17 345 high-confidence accessible regions (Supplemental Table S4) was identified by intersection of ATAC-seq peaks from all eight whole-embryo controls. A broader set of 21 414 accessible regions (Supplemental Table S5) was defined as the union of ATAC-seq peaks from all samples: tagged domains and whole-embryo controls (after intersection of peaks from replicates). In both cases, only peaks mapping to chromosomes X, 2L, 2R, 3L, 3R and 4 were considered.

Normalization of ATAC-seq signal

To normalize for differences in sequencing depth between the samples, their ATAC-seq signal was divided by scaling factors representing a median of signal ratios over a complete set of high-confidence ATAC-seq peaks (calculated with the function estimateSizeFactorsForMatrix,

from DESeq2 v.1.20.0, Love et al. 2014). Size factors were calculated separately for ATAC-seq signal expressed as the total number of cuts and coverage of 1-100 bp fragments, in order to control for differences in size distribution of the sequencing libraries (provided in Supplemental Table S3).

Comparison of ATAC-seq signal between temporal classes of embryonic genes

Different temporal classes of embryonic genes were defined based on annotations of RNA *in situ* hybridization images from Berkeley Drosophila Genome Project (Tomancak et al. 2002), and represented genes whose transcription was initiated specifically at stage 4-6 (exclusion of maternal genes), 7-8, 9-10, 11-12 and 13-16 of embryogenesis. ATAC-seq signal was calculated over coding sequences of the genes (gene models: Flybase Release 5.57, Gramates et al. 2017), and expressed as the mean number of Tn5 transposase cuts per base-pair, after normalization to the sequencing depth with the corresponding scaling factors. Introns were omitted from the analysis in order to avoid potential confounding effects of active and accessible intronic *cis*-regulatory elements. Spatial expression patterns of the analysed genes were not considered, and thus only whole-embryo controls were examined.

Correlation of ATAC-seq profiles with published DNase-seq data

Published DNase-seq accessibility data from whole stage-5 embryos (Thomas et al. 2011) was downloaded from the UCSC Genome Browser, dm3 release (Kent et al. 2002): coordinates of DNase-seq peaks (intersection of peaks from two replicates, FDR = 5%) and wig tracks with normalized DNaseI tag density (both replicates). In order to control for the different size and genomic distribution of DNase-seq and high-confidence ATAC-seq peaks, first a union of these accessible regions was defined, followed by identification of peaks from each dataset that overlapped the union (minimum overlap = 1 bp). Correlation of accessibility signal was calculated over a complete set of high-confidence ATAC-seq peaks. DNaseI tag density and the total count of Tn5 transposase cuts was calculated separately for each replicate, followed by estimation of a Pearson's correlation coefficient between mean ATAC-seq signal from eight whole-embryo controls and mean DNase-seq signal from both duplicates.

Principal component analysis

Principal component analysis (PCA) was performed with DESeq2 package v.1.20.0 (Love et al. 2014) on the normalized and rlog-transformed total count of Tn5 transposase cuts, over a complete set of high-confidence ATAC-seq peaks. Function plotPCA was customized in order to include all input genomic intervals.

Identification of differential ATAC-seq peaks

High-confidence ATAC-seq peaks that showed significant differences in the total count of Tn5 transposase cuts between different tagged domains were identified with DESeq2 package v.1.20.0 (Love et al. 2014), using default parameters and an adjusted p-value below 0.01. A total of 28 unique pair-wise comparisons was performed: a) between domains D1, D4, D5,

D7 and their corresponding whole-embryo controls with the same genotype (multiple-factor design: domain/control + transgenic line A/line B), b) between domains D2, D3, D6 and all eight whole-embryo controls (single-factor design: domain/control) and c) between individual domains (single-factor design: domain A/ domain B). As D6 replicate 1 showed strong clustering with whole-embryo controls (Fig. 3A), it was excluded from the analysis. A complete set of differential ATAC-seq peaks is provided in Supplemental Table S6. A corresponding differential analysis was performed for the union of ATAC-seq peaks (differential intervals are listed in Supplemental Table S7).

Constitutive ATAC-seq peaks were defined as a complementary subset of high-confidence ATAC-seq peaks that showed no significant difference in their accessibility signal in any of the pair-wise comparisons (adjusted p-value ≥ 0.01). Differential peaks were divided into quarters based on the maximum absolute value of $\log_2$ fold-change reported among all 28 pair-wise comparisons.

Genomic annotations of ATAC-seq peaks

Drosophila gene models were downloaded from FlyBase release 5.57 (Gramates et al. 2017). Promoters were defined as intervals 100 bp upstream and downstream of transcription start sites (TSS). Peaks were tested consecutively for their overlap with promoters, CDSs and introns, with a minimum overlap of 100 bp. Remaining peaks were tested for their overlap with UTR regions (5' and 3' UTRs), with a minimum overlap of 1 bp. The remaining peaks, without any gene annotations, were assigned as "intergenic" peaks.

Overlap of ATAC-seq peak with published ChIP data

ATAC-seq peaks were filtered to include only the "intronic" and "intergenic" annotations and tested for their overlap with ChIP peaks from published datasets (minimum overlap = 50 bp). Coordinates of ORIs (origins of replication) were downloaded from FlyBase release 5.57 (Gramates et al. 2017) and represented a union of ChIP-seq peaks of ORC2 (origin recognition complex subunit 2) from three *Drosophila* cell lines (Eaton et al. 2011). Coordinates of ChIP-chip peaks of insulator proteins from 2-4 h embryos were downloaded from modENCODE (Celniker et al. 2009), with the following IDs: BEAF-32 (5130), CP190 (5131), CTCF (5057) and Su(Hw) (5066).

Definitions of ChIP-chip peaks of TFs from the AP and DV patterning networks in stage 4-5 embryos (Li et al. 2008; MacArthur et al. 2009) were downloaded from the UCSC Genome Browser (dm3) as wig files representing signal intensity above FDR = 1% (single replicate, indicated as "best antibody"): Bicoid AB2, Caudal AB1, Giant AB2, Hunchback AB1, Knirps AB2, Kruppel AB2, Huckebein AB1, Tailless AB1, Dichaete AB1, Fushi-tarazu AB3, Hairy AB2, Paired AB1, Runt AB1, Sloppy-paired1 AB1, Dorsal AB3, Mothers against dpp AB2, Snail AB2 and Twist AB2. In order to represent every bp of ChIP-chip peaks, intervals from wig files were extended by +/- 25 bp.

Definitions of ChIP-seq peaks of TFs from the AP patterning network in stage 4-5 embryos (Bradley et al. 2010) were downloaded from the NCBI Gene Expression Omnibus

repository with the following accession numbers (single replicate): Bicoid AB1 (GSM511083), Caudal AB1 (GSM511087), Giant AB2 (GSM511086), Hunchback AB1 (GSM511081), Knirps AB2 (GSM511088) and Kruppel AB2 (GSM511085). For greater specificity, ChIP-seq peaks were additionally filtered based on their overlap with peak summits from ChIP-chip experiments generated with the same antibody (Li et al. 2008). Definitions of ChIP-chip summits were downloaded from Berkeley *Drosophila* Transcription Network Project (primary peaks defined by the symmetric null test at FDR = 1%).

Overlap of ATAC-seq peaks with annotated *cis*-regulatory elements

Differential and constitutive peaks were tested for their overlap with three categories of annotated *cis*-regulatory elements (minimum overlap = 100 bp). Vienna Tiles (Kvon et al. 2014) were filtered with the following parameters: verification_status = correct, positive = 1 (reported as active at any stage of embryogenesis). Temporal activity was assigned based on the earliest stage of embryogenesis at which a Vienna Tile was reported as active. REDfly enhancers were downloaded from the REDfly database v.5.4.2 (Gallo et al. 2011) with the following criteria: 1) CRM (data type), 2) expression + only (restictions), and 3) blastoderm embryo (ontology/expression term).

AP enhancers represent a carefully selected set of blastoderm REDfly enhancers that drive patterned expression specifically along the AP axis (Supplemental Table S8). Only elements that met the following criteria were considered: 1) availability of high-quality RNA *in situ* hybridization images, 2) annotated and validated target gene, 3) agreement of RNA *in situ* hybridization images with an expression pattern of the target gene, and 4) non-background accessibility signal (indicating an active functional element). Additionally, the elements were excluded if they: 1) produced very weak or spurious expression, 2) their expression pattern was characterized by strong modulation along the DV axis, and 3) overlapped promoters, coding sequences and UTR regions. For greater clarity, enhancers that defined only one or two activity domains along the AP axis were selected, resulting in omission of enhancers of segment polarity genes. Overall, AP enhancers represent *cis*-regulatory elements of gap, pair-rule and homeotic genes of the AP gene regulatory network (Nasiadka et al. 2002). Those elements whose target genes could not be easily classified are referred to as enhancers of secondary downstream genes.

Names of AP enhancers were standardized to represent their distance in kb to the closest TSS of their target gene. Overlapping AP enhancers that represent the same regulatory element are indicated as “broad” and “narrow”.

Overlap of tagged domains with activity patterns of AP enhancers

Activity patterns of AP enhancers were measured using published RNA *in situ* hybridization images of reporter constructs (references listed in Supplemental Table S8) and defined as distribution of the signal along the axis connecting the anteriormost and posteriormost tips of the embryo. Activity patterns of the enhancers, expressed as percent of the axis length (1-100%), are provided in Supplemental Table S9.

Proportion of a tagged domain being overlapped by the activity pattern of an AP enhancer, i.e. the proportion of active nuclei, was additionally scaled based on the density of blastoderm nuclei along the AP axis. The number of nuclei within each percent of the axis was calculated using the embryo model from *Drosophila* Virtual Expression eXplorer (DVEX.org, Karaiskos et al. 2017).

Linear correlation between ATAC-seq signal and proportion of active nuclei

Linear regression was performed between ATAC-seq signal defined as the mean normalized number of Tn5 transposase cuts per base-pair and the proportion of nuclei in a given tagged domain that displayed transcriptional activity of the considered enhancers. Duplicates of tagged domains were considered as separate data points, apart from D6 Rep1 that was omitted from the analysis. For the whole-embryo controls, average values of four Rep1 and four Rep2 samples were considered.

Enhancers that overlapped drivers of the nuclear marker displayed unusually elevated ATAC-seq signal in their respective tagged domain. This likely resulted from the presence of two accessible copies of the element, one in its endogenous locus and one in the attP2 integration site. As this property biased the calculated correlation coefficient and slope of the linear regression, accessibility signal of these enhancers was additionally corrected by halving their ATAC-seq signal in their respective tagged domains (for the analysis summarized in Supplemental Fig. S15). Values of Pearson's correlation coefficients and slope of the linear regression, before and after correction, are provided in Supplemental Table S8.

Identification of background regions

Background regions were defined by subtraction of the union of ATAC-seq peaks that were called in all samples (tagged domains and whole-embryo controls) from the entire *Drosophila* genome. Only intervals 1 kb – 50 kb long were considered. After normalization, background ATAC-seq signal was defined as an average across all tagged domains. Intervals from the top and bottom deciles were removed before comparison with ATAC-seq signal of active and inactive enhancers.

Nucleosome occupancy predictions

Nucleosome occupancy over AP enhancers was predicted with the NucleoATAC tool v.0.3.4 (Schep et al. 2015) based on the local size distribution of ATAC-seq fragments. Input intervals represented extension of AP enhancers by 10 kb upstream and downstream.

Tracks with nucleosome occupancy score (…occ.bedgraph…) were used for comparison with coverage profiles of 1-100 bp ATAC-seq fragments (smoothed over a sliding window of 15 bp). Correlations were calculated over individual base-pairs, using mean signal from duplicates. Only domains representing inactive (0% active nuclei) and active (100% active nuclei) enhancer states were considered. When multiple tagged domains represented either state of an individual enhancer, mean correlation over the multiple comparisons was considered for the box plot in Fig. 7D.

References:

Aljanabi S. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR- based techniques. *Nucleic Acids Res* **25**: 4692–4693.

Bradley RK, Li XY, Trapnell C, Davidson S, Pachter L, Chu HC, Tonkin LA, Biggin MD, Eisen MB. 2010. Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related drosophila species. *PLoS Biol* **8**: e1000343.

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**: 1213–1218.

Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* **2015**: 21–29.

Celniker SE, Dillon LAL, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, et al. 2009. Unlocking the secrets of the genome. *Nature* **459**: 927–930.

Eaton ML, Prinz JA, MacAlpine HK, Tretyakov G, Kharchenko P V., MacAlpine DM. 2011. Chromatin signatures of the Drosophila replication program. *Genome Res* **21**: 164–174.

Eldon ED, Pirrotta V. 1991. Interactions of the Drosophila gap gene giant with maternal and zygotic pattern-forming genes. *Development* **111**: 367–378.

Gallo SM, Gerrard DT, Miner D, Simich M, Des Soye B, Bergman CM, Halfon MS. 2011. REDfly v3.0: Toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Res* **39**: 118–123.

Gramates LS, Marygold SJ, Dos Santos G, Urbano JM, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, et al. 2017. FlyBase at 25: Looking to the future. *Nucleic Acids Res* **45**: D663–D671.

Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, Kocks C, Rajewsky N, Zinzen RP. 2017. The Drosophila embryo at single-cell transcriptome resolution. *Science* **358**: 194–199.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The Human Genome Browser at UCSC. *Genome Res* **12**: 996–1006.

Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, Dickson BJ, Stark A. 2014. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature* **512**: 91–95.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li XY, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, Hechmer A, Simirenko L, Stapleton M, Luengo Hendriks CL, et al. 2008. Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. *PLoS Biol* **6**: 0365–0388.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 1–21.

MacArthur S, Li XY, Li J, Brown JB, Cheng HC, Zeng L, Grondona BP, Hechmer A, Simirenko L, Keränen SVE, 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**.

Müller HA. 2008. Immunolabeling of embryos. *Methods Mol Biol* **420**: 207–218.

Nasiadka A, Dietrich BH, Krause HM. 2002. Anterior – posterior patterning in the Drosophila embryo. *Adv Dev Biol Biochem* **12**: 155–204.

Schep AN, Buenrostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. 2015. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within

regulatory regions. *Genome Res* **25**: 1757–1770.

Segal E, Raveh-Sadka T, Schroeder M, Unnerstall U, Gaul U. 2008. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* **451**: 535–540.

Small S, Blair A, Levine M. 1992. Regulation of even-skipped stripe 2 in the Drosophila embryo. *EMBO J* **11**: 4047–57.

Thomas S, Li XY, Sabo PJ, Sandstrom R, Thurman RE, Canfield TK, Giste E, Fisher W, Hammonds A, Celniker SE, et al. 2011. Dynamic reprogramming of chromatin accessibility during Drosophila embryo development. *Genome Biol* **12**: R43.

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

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.

Sequences of enhancer drivers, DNA linker, basal promoters, UNC84-3xFLAG and PCR primers

enhancer: hb_anterior_actv

AACAATTGCAACAGGCATTAGTTTATATATCGCTCAGGTAGACGGATGCACGCGTCAAGGGATT
AGATGGGCAGAGGTGACGGGAAGTCAGGTACAGGTCGCGGATCGGTCGGGAATCGAGGATC
ACGGATCGCGGATTGAGGATTGCGCTCTTGATCCATTCTGGATTAGAGCAGAAACAAAAAATTA
TGCGCACTTGGATTTGGATGATCCGGGAGCTTAGCGGATGGCCAGCTTAGCAGCGAGCTGCG
AATTTTCCACCGGTTTTCTATGGGGATTACGTTGGTCAGGAGTCGACAGCAGGAGTAGGCAGC
TAGCGTGGGCAGTTTCGTAGTTAATAATAAAAAGTAAAAAGGATTGCGGGACTTAACTAAATTA
ACGGATCAGAACTGCTTACACCTGCGGGAAAACTCTAAGGACCAACTAAACTATATGCATAATA
TGTGCAGTATAATTATTACACACCCATTTGAAAAACATTTTCCTGACAACAATTTTCCGCCAGAC
ATTTCACTTTGATTTGCGTAGTTTTTCTAATAATTCTCGCATTAAAATTGCTTGTTGCCTATATTTT
TTCCATTTCCAATTTCACACTGAAAAATTGTGCAGTTGCTGCATTTTTGGCTAATTGTTTGTGCT
TTCAAGTAAATATTATTAAAAACGCAAAACGGGAAAAAGGGGCATTTACGGAATATTATTATGG
GAGGATGGTGCTGTGCTA

enhancer: eve_stripe2

AATATAACCCAATAATTTGAAGTAACTGGCAGGAGCGAGGTATCCTTCCTGGTTACCCGGTACT
GCATAACAATGGAACCCGAACCGTAACTGGGACAGATCGAAAAGCTGGCCTGGTTTCTCGCTG
TGTGTGCCGTGTTAATCCGTTTGCCATCAGCGAGATTATTAGTCAATTGCAGTTGCAGCGTTTC
GCTTTCGTCCTCGTTTCACTTTCGAGTTAGACTTTATTGCAGCATCTTGAACAATCGTCGCAGTT
TGGTAACACGCTGTGCCATACTTTCATTTAGACGGAATCGAGGGACCCTGGACTATAATCGCA
CAACGAGACCGGGTTGCGAAGTCAGGGCATTCCGCCGATCTAGCCATCGCCATCTTCTGCGG
GCGTTTGTTTGTTTGTTTGCTGGGATTAGCCAAGGGCTTGACTTGGAATCCAATCCCGATCCCT
AGCCCGATCCCAATCCCAATCCCAATCCCTTGTCCTTTTCATTAGAAAGTCATAAAAACACATAA
TAATGATGTCGAAGGGATTAGGGGCGCGCAGGTCCAGGCAACGCAATTAACGGACTAGCGAA
CTGGGTTATTTTTTTGCGCCGACTTAGCCCTGATCCGCGAGCTTAACCCGTTTTGAGCCGGGC
AGCAGGTAGTTGTGGGTGGACCCCACGA

enhancer: eve_1_ru

GGCCTAATCACTTCCCTGAAATGCATAATTGTGCCGCGGCTTTTGATACGCTCCTGGCGGAGA
GGGAGATGAGGAAAGGATGCACGGGAACCGCAGCCAAGTGGCAGTCGAGATTGGCAAATCC
GCCAGCGGACAATGCCCAGAGAATGGGCAACAAGTAGCGGCGAATTAGCAATCCTATCATGCT
TTTATGGCCGGCCAACTCTTGCCCGCGCATCTCAGTTCATCCGAAGCGGGACCAGGTCCAGG
TTCAAGTCGAGGTCCAGTACCCCTGCTATCCCGTCAACCCCTTTAGGGCGATAATCCTTCTAAA
TGTTTGCATTAATTTCGAGGCGTGGACGGATTAGGGCGTGCTGGCTGGGCGGAACCCGCAGC
AGAAACCGCCGAGGACACTGCACCGACTGACCTGCAGCCTACAGATCTCTGATCTTCGATCTC
TAATCCTTTCGCATTTGCAACTGACTTCTGCACTGGGTCCGCCCCTAATCCTTCCGCCGAGAAG
GCGGCAGAGTCGCGAGGTACTGGCCCGGGGTAATGGGATTATCTGCGATTACCCCAGATGAT
CCGCAGAAAGTCAATCTGGTTCAGGGGCTAATTGTCAGCGAAGTCAACTAAATCCAATCCTTTC
GCGCCCCCTTCTGTTTATTTGTTTGTTTTCGTTTGTTTTGAGAATTTCTGGCAATTAAGTTGCCC
GTTTTGATGCGCGGGGGCGGGTGCATCAAATCCTTTCGGCATACCTGTCCTGCACAAATGCTG
AATTCCGCATCCCATGGATACCCAGATATTCAGATATCCCAAGGCCGCAAAG

enhancer: D_(+4)

CGGGGAGATCCTTAAATACGCAGAAAACTACCAGAGTTACAAAACAGAGAACAGCTTGACCCC
TTTGCAAAACGATCTTGGAAGAGTCCCGCCAGCAAGTTAATGTTTATGGCAGCCACAACCCGA
ACGATAAACATAAACTGAACCTTCCGCGTAAGAAAGGCATTTGGAAATTGCACGGAAAATGTGG
AAAATGAGAACAGCCGACGAGGTCGTCACAATGATTTCACTAGCAAATTGTATCACCGTTATTA
GCTAATCCGTCTGGTCCTTTGTCGCTAATCCGGCGGATTAGCGCGTGCATCTTTCGCCTAAAA
AAGTGCACAAAAACAGACGGATTTTTGGGAAAACTTGAAGGAAATCAATAGTCTGGCTAATCCG
GGTCCAAAGTTACCACAGGTAGCCAAAACACTGCACGAAAAGAGCAACAACAACCCAAACAGA
CAGCGCTAGAAGCATAAAAATCTCGCGACATTTCGCACACTCCTTTCTGCGGACCGGACGCTA
ATCCCCGGTCATAAAGTCGTAAAAAAATCGAAGCAACCGGGAGGCTGCAAATTATATTTGTATA
CAGACGTGACCAAATGGGACGAAAGGCGTGGCTTTTGGGAGTAGGAGGCGTAGGCGAGATG
GGTCGTAAAACGTTCTCGACAATCCTTAAAAACCCAACGAGTTGAAATGGAATTTCAATTATTC
GTGCATCATTTTTTTTGGATGAAATGTTTAAACTATGGATCAGGAAGGCGTGAGCGGGCTTATG
TTTTAAGCTTGGCTTAGTCGGATGGATGGATTGGGCGTTTAGCGAAAGTTAATCACTTCCGTTT

GGGTTCGGCAAAGGCAACATCCGATATTTGATCATAGATTGTTTTAATGGCTGATTATGGTTGT
AACTAGTTTTATGATCAGAATGTTATGATGCGAAAAATGGCATAAAATTTAACTTGTAAGATTCA
AAGTAGTTACGAAATATAATGTAATAAAGTAGAATAATAAAATACAATATGTTACTTCAGCAGCG
CGAATTCTTGAGCCGCTATTATCTAAAGGATGTCAAATTAAAATGCAATTTTTAAAATGCGGATG
CCGTTCTTTTAATTGATATTTATTATTTACAAAATCTGTAATAGTTTCTAGATTTGAAAAATAGTTA
TCCAATCAAAATACTATAAAAACCTCTTTCCTTACAATACAAATGTATTGTTGAACCCTTTGGAA
GCCTTTACCGCTGTGCACACAAAGTAATGTTCCAAAGAGCTAAGCCAGAAGAATCCAATTTTTA
TATAATTATTCAATATCACGTCAATGCTGGGTAAATATTTGTTTGCAGAGCAAACGGAGAATGGA
GAAACAAGGCATGGAAAAAAGACCAATATCAACAATACGCATAATAACAACAATAATCAATATTT
GATATTTTAACACATTTGACATACAACGTGGAAAACGAAGGGGAGCGGGGCAGAGGGCATTGA
ATAGACTCCCAAAGACTTGAAAACGGCAATCGAAGAGTCAGAGTTCGGGGGAAGACAATGCAA
GTGGATTTATTAATGACCGCAATTATTCTTATTATTATTACGAAAAGAAGCTAAAAATCAAAACG
GAAAAACACGCGTTGATTTCTTTCGCCAGGGTATTCAGTTACACTCCGAATAGTGAGGTCTTAG
AGGGCCCAGGCCCCCCGTTGATCCCGAAAAAAAAACACAATGTGACCCCCATCCAAAGAGC

enhancer: Kr_CD1_ru

AAATCGGGATCCTAAGTTAACTATAATCCAGGCTTAATCACTGGATCAATAACTAAGTAGCATTT
TCCGGGATGGAAATATGAAGTTACCTGCATATGCCTACCGATCCTGAAAACTGCTTTAACTTAA
TCGACATGCATGATCATAAAAAGCAATTTGCTACAATTTATATTTTTTTGCTTTTCCTTCTTTTAA
GCATCTGGGATCTGGATCAGAAAAGAAAAAGTGTAACGCCTACCTTCAGAAACGGATTAAATTT
TTTCAGACAAATAATCCAGCCTTAAGCATGGTGATTAAGCTTGATCCCCTACCAAGGGGCGTAA
TATTGACGGATTTTCCTTAAATCCCTCTGTTAATCTCCGGCTTAGAGCGCGACGCGTTTTTTCG
CGACTCCGCCTGCATTGTTTTTTTTTTCAGTTTCTTCAATTCGCAAGAAGGCAGGCCTATGGAC
CGAATGAGGATCATAATTATGGAATTCCTAAATAAACTAAGAAGGGCAGTCGGCATAGTATTGA
TCTACCTGTAAGCGTGGGTTCTATCTTTGCCCCTCGCATTCGAGACTCTCTAGTCACAGGTAGA
CCAGCCTTGAGTTCGTCGGCAATTAAGAAGTCAAATTTCTCTTAAAAACAACAAAAAATGTCAAA
GTAAAAACAATGCAAAAAATATGTGTAACTGAACTAAATCCGGCTTAGGATTCTTGCGTCATAAA
CATGACTAGGGAGCCATTAAAATTTGATAATTTGCATGTCACTTGTGCACCGTAAACGAAGCCA
ACACATCGGCTGAAACCCAGCGTCATTTATGCTATGCTATCTCACTCTTTTGCATAATTTTTTTA
AATTATTGACTTGTCTTTTTTTACGAATGCAAAGGAAGTACTTGTGTACTTCCGTTTTTCAAAGTC
CGATTTTATTATCTACCTGCAATTTATAATGACAACCAAGTTTTTAACATTCCAAAACGTGGTTTT
TTGAGTTCAATAAGCTTCATTTTTTAAAGGCTCTTTTCGAACCACATTAACATTTGCGGTTACGT
TTTTGTAGAAATTTTCCATAGCTGCATAAAGGTACAAGTTTCATTTATTTGATGGGAAAATGAATT
AGAGACCAGCACGATCTTTAACAGAAACAAAAAGGCGTGGATATTGTTATATGTACAAAAACAA
CCATGGCATATACGTATGTATTGTATATCCTAATTCGGCCTTCCTCTTATTTAATCACTTGCAAT
AATTTTCAACTATACAGATCAGCAATCAAAGTTCTTGATCCAATCGCTATAGTAACCAGATGCGC
AAGCTGATAATTTATTGAAAACTTTTTTCTTAGGATTCTTCCATATTAAGCATTTCCCTTGCTGTG
ACCTCGATTCAAAAGAAAGCACTGATTTAATGCAGGTAATCGTCAAAAAAATAA

enhancer: eve_stripe5

GGGCGGGTGCATCAAATCCTTTCGGCATACCTGTCCTGCACAAATGCTGAATTCCGCATCCCA
TGGATACCCAGATATTCAGATATCCCAAGGCCGCAAAGTCAACAAGTCGGCAGCAAATTTCCC
TTTGTCCGGCGATGTGTTTTTTTTTTAGCCATAACTCGCTGCATTGTTTGGGCCAAGTTTTTCTT
CTGCCAAATTGCGGAGATGATGCGGGGATTATGCGCTGATTGCGTGCAATTATGGACATCCTG
CGAGGCCCCGAGGAACTTCCTGCTAAATCCTTTCATCCGCCTACAGAACCCCTTTGTGTCCCG
TTCGCCGGGAGTCCTTGACGGGTCCTTCGACTATTCGCTTACAGCAGCTTGCGTAAAATTTCAT
AACCCTACGAGCGGCTCTTCCGCGGAATCCCTGGCATTATCCTTTTTACCTCTTGCCAATCCGT
TGGCTAAAAAACGGCTTCGACTTCCGCGTAACTGCTGGACAACAAAGACAAAAAACGGCGAAA
GGACGGCGATTTCCAGGTAGCATTGCGAATTCCGTCAAACTAAAGGACCGGTTATATAACGGG
TTTATATGGCCAGAATCTCTGCATCTCCACGACCGCCAGAAGCTGCGTAAAACTGCAGGCTCT
GTTTTGATTTCTGCAACTTCAGTTAATTGCCCGGGATGGCCAGCAATTGCCGGCAATTATAAAA
CAGCGCAGATGTGACTCAGCTTCCATATCTAACTCTATATCTCATGCCGAAAATCTAGGGTGGG
GAGCGGAGGGGCGGGGTGCGTGGGTGACTTGCCTGCCA

enhancer: gt_(-3)

CTGCTCGTGTTTGCCCTCCTCCTTAAACATCCTTGATTCTACGTCCTTTATCCCTTTGGCTTCCG
CTGCGGCGTGGGATGTGGGATCCTCGACAAGTTGTTTCACTTTGTTAACTGTTTGTCCTAACG
GTCGCCTCGACAGCCAACGGAAGTGGGTAGCGGTACAAAGTTGGCTATCATCCATGGGAAAT
GTTATGCTAGATACTTACTGCAGCTAATTCGAGGACATAACGTAGTTTAGATTTGTTACGCTGAA
ATGTTATCAAATCCTATCAACAATACCTACATTATATACTATCTTTCATTCATTCTCTTTTTACAAC
TGCCCATTCAGGGGGATTGGGTGAATTGCAAATAAAAAGAGGGATCGATCGATGGTAATCAGT

AAAAAACCGCGAACTTATCCGCTAATCCGTGCGCTAATCCAACCGTAAAAATGGTGGATTAAAC
TAAAACTCGGACCGTCACGTCGCCGTCATTGAGTCGATCCCTTACCTGCGGTTACCGGGCCCG
AACCCGAGCCCTTCCAGATCCTGGACATACGTATAGCCCAGCCCAATCGATCCCGAGCAGATG
CCATAAAAAGTCGCGGCCGAAGTCGAAGCCAAAGAGTTGCCAAGTTCGGTGTAAAAACCAGAT
TCAGCGAAAAAAAAACGTAAAAAACAGCAGGCCATAAAACCATATATTCATATACTCGGACTGC
TGTTCCTTCGCCATAATGAGCTCCGAACCAAAGATGCAATATTTGTGCAATCCTTTGAGTTGTG
ACCTTGGGCGCGTTTTTTACGGAATCATCTACCTGATGGCCGAACGGGTTGCCCACCCACGGT
ATGCGCTGTATTAGGCTGTTATGCTGTTGCCTAAAAAACATCCAAAGTTTTATTGCATAATGGAC
GCTTCGTGCGAAAAGGCCTCATTCCGAGTCCGCGGGCCCTAAAACCGCTGCAGTTTTTTAAGG
AGCCATAAACTGATATTTCGGACACCGGCGGTCCTTAAAAAAAAAAGGGTTCTCTCCTCCGACC
ATTGTGGCTGCGATTTTTATATTGGTCCGCTGCTAGGTGACAAAGAAAATCAGCGACTAATTGA
CTGAAAGGGGTTTTGCCTTATATGACCCGTCCCACCCGTCCAAAAAAGGGTGAGGAGTGGTTT
GGATTTGGATTCGTGCTTGGTTTCGGGGAATTGCCGTTAACTCCTTTCGCGGCCCTTTGTC

linker DNA

TCCAGTGTGGTGGAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCC
CTTCGAA

basal promoter: Hs43

AAGAGCGCCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGC
AAAGTGAACACGTCGCTAAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAAC
AATCTGCAG

basal promoter: DSCP

GAGCTCGCCCGGGGATCGAGCGCAGCGGTATAAAAGGGCGCGGGGGTGGCTGAGAGCATCAG
TTGTGAATGAATGTTCGAGCCGAGCAGACGTGCCGCTGCCTTCGTTAATATCCTTTGAATAAGC
CAACTTTGAATCACAAGACGCATACCAAAC

nuclear tag: Unc84-3xFLAG

ATGGCTCCCGCAACGGAAGCCGACAACAACTTCGACACCCATGAATGGAAATCGGAATTCGCA
TCCACACGCTCTGGACGCAATTCTCCAAACATTTTTGCAAAAGTTCGCCGGAAGCTTCTCCTGA
CTCCACCAGTTCGAAACGCCAGATCGCCACGTCTTACCGAAGAAGAGCTGGATGCTTTGACAG
GCGACTTACCATACGCAACCAACTACACATACGCATACAGCAAAATCTACGATCCATCCTTGCC
GGACCACTGGGAAGTGCCAAACCTTGGTGGTACTACTTCAGGATCACTCTCTGAGCAGGAGCA
CTGGTCAGCGGCCAGTCTCAGCAGACAGCTTCTCTATATCCTCCGTTTCCCCGTCTACCTTGTT
CTTCACGTCATCACCTACATTTTGGAAGCTTTCTACCACGTCATCAAGATCACTAGCTTCACCAT
CTGGGACTACCTGTTGTATTTGGTGAAACTCGCGAAAACTCGTTACTACGCCTACCAAGATCAT
CGTCGCCGTACAGCTCTCATTCGCAACCGGCAAGAGCCATTCTCCACTAAGGCTGCTCGTTCT
ATTCGTCGATTCTTTGAGATCCTTGTCTACGTCGTGCTTACTCCTTACAGAATGCTCACAAGAA
GTAACAATGGCGTGGAACAGTACCAGTACCGTTCGATCAAGGATCAATTGGAAAATGAGAGAG
CTAGCAGAATGACGACAAGATCCCAAACATTGGAAAGAAGCCGCAAGTTTGATGGATTATCGA
AATCACCAGCACGCCGAGCAGCTCCAGCCTTTGTGAAGACTAGTACAATTACCAGAATCACTG
CCAAGGTGCTCTCGAGCTCTCCATTCGGAGAAGGAACGTCCGAAAATATAACCCCGACTGTTG
TGACTACTAGAACAGTGAAGCAACGCTCAGTTACCCCAAGATTCCGCCAAACCCGTGCCACTC
GTGAAGCTATAACTCGAGCACTCGATACTCCGGAACTCGAAATCGACACACCACTCTCCACAT
ATGGACTTCGAAGCCGAGGACTGAGTCATCTGAATACTCCTGAACCAACTTTTGACATTGGTCA
TGCTGCTGCAACTTCCACGCCTTTGTTCCCACAAGAAACTTACAATTATCAATACGAAGAAGCG
ACAGGAAATAAGATTAAAACTGCATTCACTTGGCTAGGTTACTTGATATTGTTCCCGTTCTTTGC
GGCACGACATGTATGGTATACGTTCTACGATTATGGAAAGAGTGCCTACATGAAGCTGACCAAT
TATCAGCAAGCGCCAATGGAGACTATTCATGTCAGAGATATCAACGAACCGGCACCAAGTTCA
TCAGATGTTCATGATGCTGTTGGTGTTTCTTGGAGAATTCGAATTGCCGATTTCTTGAGCTCATT
CGTAGCAACAATCGTTGAAGCGCATCAAGTGGTATTTGCAATGTTCAAAGGAGGAATTGTCGA
GACAGTTTCCTATTTTGGAGGACTATTTGCTGGTCCTACCGATAAGAAATCATCAAAGTTCTCG
TGGTGTCAAATTCTCGGTCTACTTCTGGCTCTTCTCTTCGCCATCTTTCTCCTTGGATTCCTGAC
ATCTGACAACACAGCAATAAGAGTTAAAGAAATTACCAAAGATAAGAATGCATCTAAGAAGTCG
GAAGGATCCCTCCCAGCTGTGCCAATCTGGATTTCAGCTGCAAATCACGTTAAACATTACACAT
GGATGGTGAAGGAATTTGTTGTAGATATTGCATTTGACACGTACAACTATGGAAAGTCGACGAT
TGGTAGACTTGGCACTACTCCACGTTATGCTTGGGACCTGATTGCAAGCGGATGTGGCGCTGT
TGGAAATGGCTTAAAATCTGTGCTCTCATCGAGTTTTCGATTCATCGATTTTTGTGCTGGAAAG
CTATTTTACTATGGCTCAGATGGGTTCTTGTCAGCAAACAAGTCTATCGGAACCTTTTTCAACG
GTTGCTACGAAACCTTGTACAACGGATGCACAGCAATTGTTGGCCATACAAAGAGCTTCATCTA

CAATGCTTCAAATGCTGTTTACAACTTTTTCTCAACTATCTTTGCCGGCCTCTTAAACTTTTCTAC
TTCTTCCCAAAACTCCATTCTTTCTCTTCTCAAGTCATTTGGCACCGGAATCACTAACATTTTTTA
TAACTTCATTTATGCACCAATCGCTGGAGTGTTCAACTTTGCTGGTGATAACTACATGTATTTCT
TCAATGAGGTAGCGGCAGTCTTTGGAAAAGTGTACAACTCCGTGGTTTCCGTGCTCAAAACTG
TAATTAACTGGATTCTCTTCCTCATTGCCTACCCATTCAGTTTGTGCACTCGTGCTTGGATTCGC
ATCAGCCAATATGCTCCAGAAGATGTTGTTCAAGTGATTCCAATTCCACAAGCTATTACCCCAA
CTCCGGATGTGGAGCGTATTGTTGAAGAGCCACTGAGAAAAGTCACCGATGTGGAGGACGAA
GAACTAGTGATAATTCCCGCCCCCGCACCTAAACCTATCCCAGTCCCAGCGCCAACTCCGGCC
CCAGTAATTATCCATCAGACTAACGTTGTTGAGACTGTTGACAAAGATGCCATCATTAAGGAGG
TAACGGAGAAGCTTCGCGCCGAGTTGTCCGCCCAATTCCAGCAAGAGCTTAGCGCAAAGTTTG
AGCAAAACTACAACACAATTATTGAGCAACTGAAAATGGAAAACACCAACATTCAATATGATAAG
AATCATTTGGAAGCTATCATCCGTCAAATGATCTACGAGTATGACACGGATAAAACTGGGAAAG
TTGACTATGCCCTGGAGAGCTCAGGTGGAGCTGTTGTGTCAACAAGATGCTCGGAGACGTACA
AAAGCTACACAAGGCTGGAAAAGTTTTGGGATATCCCAATCTACTATTTCCATTACTCTCCAAG
AGTTGTCATTCAGAGAAATTCCAAATCCCTGTTTCCTGGGGAATGCTGGTGCTTCAAAGAATCC
CGTGGCTACATTGCTGTCGAGCTGTCTCATTTCATTGATGTTTCTAGCATCAGCTATGAGCACA
TTGGATCAGAAGTTGCTCCAGAAGGGAACCGGTCGAGTGCTCCAAAGGGAGTCCTCGTTTGG
GCTTACAAGCAGATTGACGACCTGAACTCGAGAGTTTTGATTGGCGACTACACTTATGATCTTG
ATGGCCCGCCACTTCAATTCTTCCTTGCCAAGCACAAACCCGATTTTCCTGTCAAGTTTGTGGA
GCTCGAGGTGACAAGCAATTACGGAGCTCCGTTCACATGTCTCTACCGCCTTCGTGTTCATGG
AAAAGTTGTTCAAGTTCTGGCCGAGGCCGCCGCCAAGGAGGCCGCCGCCAAGGAGGCCGCC
GCCAAGGAGGCCGCCGCCAAGGCCGCCGCCGACTACAAAGACCATGACGGTGATTATAAAGA
TCATGACATCGATTACAAGGATGACGATGACAAGTGA

fw primer: hb_anterior_actv	AACAATTGCAACAGGCATTAGT
rv primer: hb_anterior_actv	TAGCACAGCACCATCCTCCC
fw primer: eve_stripe2	AATATAACCCAATAATTTGAAGTAACT
rv primer: eve_stripe2	TCGTGGGGTCCACCCACAACTA
fw primer: eve_1_ru	GGCCTAATCACTTCCCTGAA
rv primer: eve_1_ru	CTTTGCGGCCTTGGGATATC
fw primer: D_(+4)	CGGGGAGATCCTTAAATACGCAGAA
rv primer: D_(+4)	GCTCTTTGGATGGGGGTCACATT
fw primer: Kr_CD1_ru	AAATCGGGATCCTAAGTTAACT
rv primer: Kr_CD1_ru	TTATTTTTTTGACGATTACCTGCA
fw primer: eve_stripe5	GGGCGGGTGCATCAAATCCTTTCGG
rv primer: eve_stripe5	TGGCAGGCAAGTCACCCACGCACC
fw primer: gt_(-3)	CTGCTCGTGTTTGCCCTCCTCCTT
rv primer: gt_(-3)	GACAAAGGGCCGCGAAAGGAGTTA
fw primer: UNC84 coding sequence	ATGGCTCCCGCAACGGAAG
rv primer: UNC84 coding sequence	AACTTGAACAACTTTTCCATGAACAC