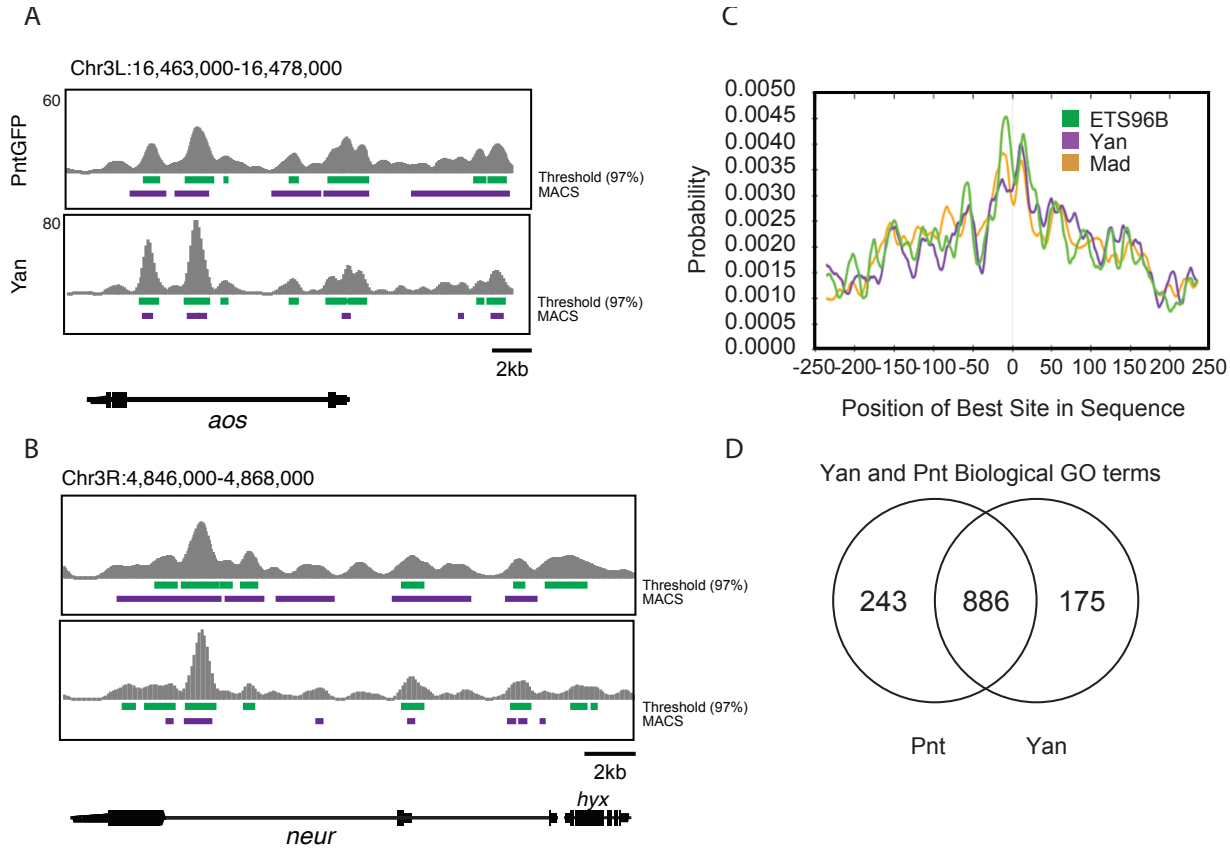

Fig. S1. Yan and Pnt ChIP-seq bound regions are highly overlapping.

A,B) Comparison of ChIP-seq read density for Pnt-GFP (Pnt) and Yan across the *argos* and *neuralized* loci. The RefGene gene track is shown below the profiles. Green boxes depict peaks called using the Integrated Genome Browser (IGB; Affymetrix) using a cut-off of the top 3% of bound regions (97% threshold) as previously described in Webber et al. 2013a. Purple boxes depict peaks called by MACS. C) Analysis of transcription factor PWMs revealed central enrichment for Mad and ETS motifs in the top 100 sequences bound by Pnt. D) Venn Diagram depicting GO terms significantly overrepresented in Yan and Pnt datasets.

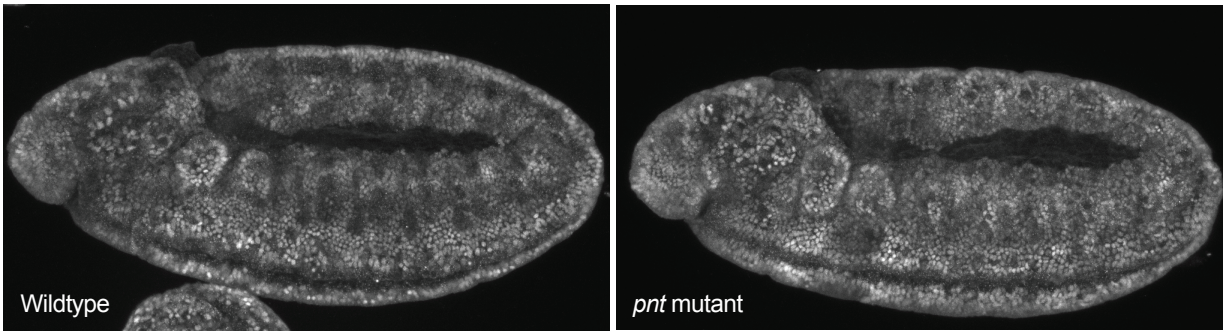

Fig. S2. Yan protein levels in wildtype and *pnt* null embryos.

Embryos were fixed, stained and imaged in parallel with identical confocal settings.

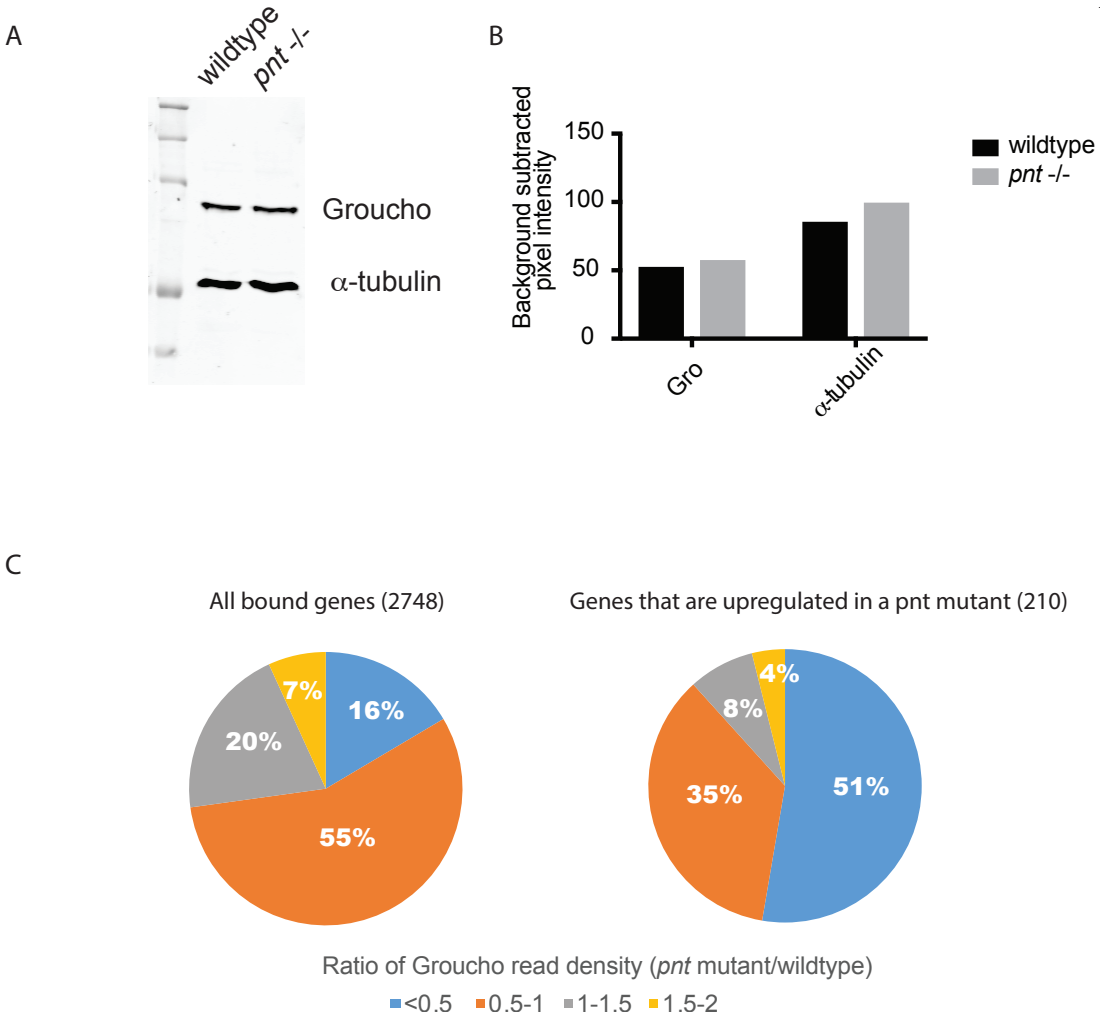

Fig. S3. Gro occupancy is reduced in *pnt* mutants, while Gro protein levels do not change. A) Gro protein level does not change in *pnt* null embryos. B) Pixel intensity ratios of *pnt* mutant to wildtype for Gro and Tubulin. C) The proportion of genes associated with Gro occupancy loss (<0.5), no change (0.5-1.5) and occupancy gain (>1.5) were determined for i) all bound genes and ii) bound genes associated with upregulated expression in a *pnt* mutant. Genes that are upregulated in *pnt* mutants are more frequently associated with Gro peak loss relative to all Gro bound regions.

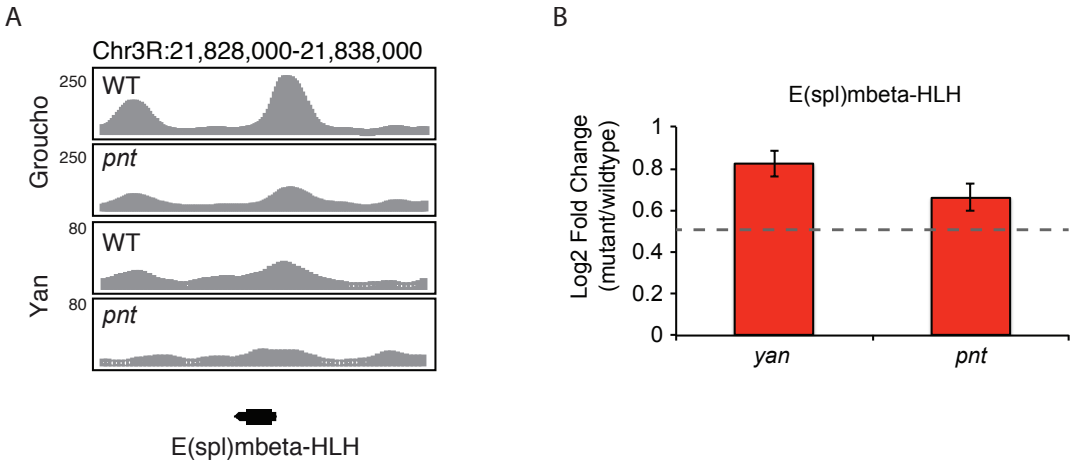

Fig. S4. In the absence of *pnt*, Yan and Gro occupancy is reduced at the *E(spl)mbeta-HLH* gene locus and unregulated expression is observed. A) Comparison of ChIP-seq read density for Groucho and Yan across the *E(spl)mbeta* locus. The RefGene track is shown below the profiles. B) Bar charts showing the average upregulation of E(spl)mbeta-HLH probes from *yan* or *pnt* mutants relative to WT. Error bars represent standard deviation of 4 probe sets.

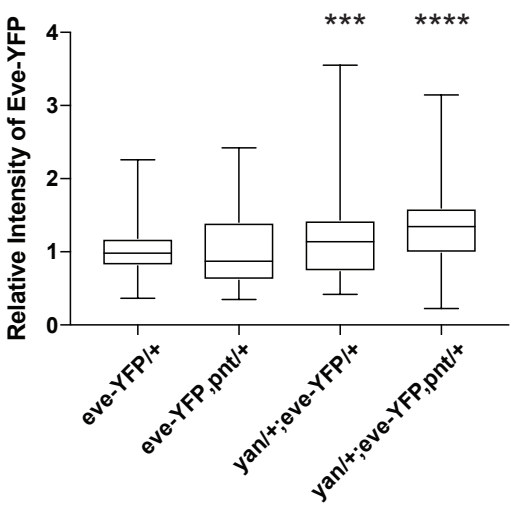

Fig. S5. *yan* and *pnt* collaborate to fine-tune eve expression. For each genotype, the average pixel intensity of Eve-YFP expression was measured per cluster of Eve-positive cells and the background pixel intensity was subtracted. Measurements were normalized to the average wildtype cluster intensity. Animals doubly heterozygous for *yan* and *pnt* were associated with increased cluster intensity relative to the wildtype control and single heterozygotes (n = at least 4 embryos per genotype; P<0.0005 *yan*/+;*eve*-YFP/+ and P<0.0001 *yan*/+;YFP,*pnt*/+, ANOVA).

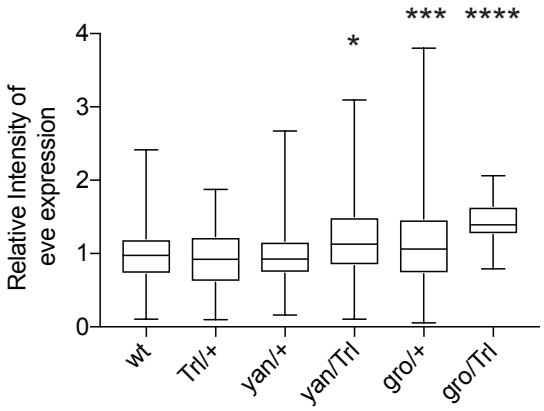

Fig. S6. *yan* and *gro* genetically interact with *Trl*. For each given genotype, Eve expression was measured in each cluster of Eve-positive mesodermal cells. An increase in Eve intensity was measured in animals heterozygous for both *Trl* and either *yan* or *gro* relative to the single heterozygotes (n = at least 5 embryos; P <0.05 *yan*/*Trl*; P<0.0005 *gro*/+ and P<0.0001 *gro*/*Trl*, ANOVA).

Table S1. List of Pnt and Yan bound regions and associated genes. UCSC Table browser (Karolchik et al. 2004) was used to assign each Pnt-GFP and Yan bound region to the nearest gene.

Click here to Download Table S1

Table S2. MAnorm comparison of Pnt datasets. M = $\log_2$ (Pnt-GFP read density wildtype/ Pnt-GFP read density *yan* mutant) and A = 0.5 x $\log_2$ (Pnt-GFP read density wildtype/Pnt-GFP read density *yan* mutant). A linear regression model was applied using peaks common to both wildtype and *yan* mutant datasets (Shao et al. 2012b). This model was then used to output normalized M value (Column E) and A value (Column F) and associated -$\log_{10}$ *P*-value (Column G). Normalized M values were used as a readout for differential binding with M values of >0.5 or <-0.5 denoting peak loss or gain, respectively. Using this cut-off, all differentially bound regions have *P*-values <0.05.

Click here to Download Table S2

Table S3. Panther enrichment analysis of genes that are upregulated or downregulated in *pnt* mutant embryos. Significant GO terms describing each gene set (upregulated or downregulated genes) are listed, along with the fold enrichment over background and associated P-value. Significant GO terms for each dataset were compared to identify overlapping GO terms or uniquely enriched GO terms.

Click here to Download Table S3

Table S4. List of all probes with fold change (mutant/wildtype) and *P*-value. Probes were assigned to genes using probe annotations from affymetrix and Flybase (Gramates et al. 2017). A gene was considered as differentially expressed if it had an assigned probe with a P-value <0.05 irrespective of fold change. Where two or more probes correspond to the same gene, the highest absolute value (maximizing) was utilized for differential expression analyses.

Click here to Download Table S4

Table S5. MAnorm comparison of Yan datasets. M = $\log_2$ (Yan read density wildtype/Yan read density *pnt* mutant) and A = 0.5 x $\log_2$ (Yan read density wildtype/Yan read density *pnt* mutant). A linear regression model was applied using peaks common to both wildtype and *pnt* mutant datasets (Shao et al. 2012b). This model was then used to output normalized M value (Column E) and A value (Column F) and associated -$\log_{10}$ *P*-value (Column G). Normalized M values were used as a readout for differential binding with M values of >0.5 or <-0.5 denoting peak loss or gain, respectively. Using this cut-off, all differentially bound regions have *P*-values <0.05.

Click here to Download Table S5

Supplemental Methods

Chromatin Immunoprecipitation (ChIP) and downstream analysis

Stage 11 embryos (approximately 5-7 h after egg lay) were dechorionated in 50% bleach and cross-linked with 1.8% formaldehyde for 15 min. Cross-linking was stopped by washing embryos with 125mM Glycine. Fixed embryos were washed twice with PBS/T (PBS, 0.1% Triton X-100) and Wash Solution (10mM HEPES, 10mM EDTA, 0.5mM EGTA, 0.25% Triton X-100), and then homogenized in ChIP lysis buffer (50mM HEPES, 140mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, protease inhibitor tablet (Roche)). Lysates were sonicated (11 cycles at 15% amplitude for 15 sec (0.9 sec on/0.1 sec off)) using a Fisher Scientific Sonic Dismembrator sonicator. Clarified lysates were incubated with the previously validated antibodies: guinea pig anti-Yan (1:500, in-house, (Webber et al. 2013a)), rabbit anti-GFP (1:100, A6455, Invitrogen, Lot# 1603336, (Yamaguchi et al. 2009)) or anti-Gro (1:200; (Nègre et al. 2011)) overnight at 4°C. Gamma-bind sepharose beads (GE Healthcare) were added to the lysates and samples were incubated for 4 hours, rotating at 4°C. The beads were washed thrice in ChIP lysis buffer, once in high-salt ChIP lysis buffer and once in TE. ChIPed material was eluted by a 15 min incubation at 65°C in TE/1% SDS with regular vortexing. Chromatin was reverse cross-linked by incubation overnight at 65°C. DNA was purified using the QIAquick PCR Purification Kit (Qiagen).

For ChIP experiments performed in a *pnt* mutant background, batches of 400 stage 11 GFP negative *pnt* null embryos were hand selected from embryo collections of either pnt^{AF397}/*TM3, twist-Gal4,UAS-GFP (TTG)* or $pnt^{\Delta 88}$/*TTG* animals. Around 1600 pnt^{AF397} embryos were used for each ChIP-seq replicate of Gro, and ~4000 $pnt^{\Delta 88}$ embryos for each replicate of Yan.

ChIP-qPCR was performed using the QuantiTech SYBR Green PCR Kit (Qiagen). Briefly, the relative amounts of input and immunoprecipitated DNA were determined based on standard curves generated for each primer pair (sequences previously published in (Webber et al. 2013a, 2013b)), and the ChIP signals were calculated as % input. ChIP signals were normalized to a negative control region (NC1) to account for any differences in starting material amount.

ChIP-seq visualization and thresholding

Integrated genome browser was used to visualize wig files for each ChIP-seq dataset. Thresholding of the top 3% of bound regions (97% threshold) was used to separate the genome into bound and un-bound regions for each transcription factor.

Motif analysis

Centrimo analysis (Bailey and Machanick 2012) was performed on the top-scoring 100 Pnt ChIP-seq peaks identified by MACS. Each region was trimmed to 500bp around the MACS defined summit. Sequences were scanned against a set of 1419 DNA motifs from a combined database of Drosophila TF DNA binding sites (OnTheFly (Shazman et al. 2014), Fly Factor Survey (Zhu et al. 2011; http://pgfe.umassmed.edu/TFDBS), FLYREG (Bergman et al. 2005), iDMMPMM and DMMPMM (Kulakovskiy and Makeev 2009; Kulakovskiy et al. 2009)) with motif sites reported only when enriched.

GO analysis

MACS defined bound regions were assigned to the nearest TSS using the UCSC genome browser. Genes were functionally classified with Gene Ontology terms using GO (Ashburner et al. 2000; The Gene Ontology Consortium, 2017).

Tag density Analysis

Assigning Yan/Gro peak loss to Pnt bound regions

To assess whether Yan and Gro peak loss occurs preferentially at Pnt bound regions, the read density of Pnt at the midpoint of Yan MACS defined peaks was first determined (see Materials and Methods for description of read density analysis; Table S5). Using a read density threshold of 97% (see Fig. S1), these peaks were categorized according to the absence or presence of Pnt (Fig. 1H; Yan only or Yan and Pnt bound). The *pnt* mutant to wildtype Yan density ratio was calculated and plotted for each "Yan only" bound region or each "Yan and Pnt" bound region. An identical approach was taken to assess *pnt* mutant to wildtype Gro density ratios at "Gro only" or "Pnt and Gro" bound regions in Fig. 3B, and at "Gro only" or "Yan and Gro" bound regions in Fig. 3F.

Aggregate Profiles of Trl and Bab1

Intersectbed was used to overlay Yan, Pnt and Gro bound regions (called using our 97% thresholded method). Sorted bed files for the Trl and Bab1 datasets were produced using the BEDOPS wig2bed script (Neph et al. 2012). A Visual Basic Application (VBA) for Microsoft Excel was then used to generate matrices of read density data spanning 2kb either side of the midpoint of peaks identified as Yan/Pnt and Gro bound (available on request). These matrices were then aggregated to produce average read density profiles.

Microarray validation and analysis

We have considered a probe to be differentially expressed if its expression changes between two treatments, regardless of fold change. This approach was taken to

maximize the degree of overlap between the ChIP-seq and microarray datasets, ensuring adequate sample size for the downstream analyses presented in Fig. 3G and Fig. S3C. Thus, a *P*-value of the expression change between wildtype and *yan* or *pnt* mutants was calculated for each probe. Genes were considered to be differentially expressed if they were associated with probes with P-values less than 0.05.

RT-PCR

Total RNA was isolated using TRIzol (Invitrogen, 15596018) according to the standard protocol (http://www.flychip.org.uk/protocols/gene_expression/standard_extraction.pdf). Purified RNA was resuspended in 16µl of RNAse-free dH_20 and subjected to a 1 hour DNase I (Invitrogen, 18068015) treatment at 37°C. 2µg of RNA was reverse transcribed using the Promega Reverse Transcription System (A3500) in 20µl using oligo-dT primer. Real-time PCR was performed using a 1:5 -1:10 dilution of cDNA with the StepOnePlus Real-Time PCR system (Applied Biosystems). The following primers were utilized: aos 5' GCATCCTCTACCAAGTGGGG and 3' GCGATTCGATTCAGGACAACG; mae 5' TATCAAATGCTGGACAAGTG and 3' TCAGTCGATTGTTATTGTCG; eve 5' CCTCTTGGCCACCCAGTA and 3' CGGACTGGATAGGCATTC; aop 5' CCAGCAACGAGGACTGTTATCC and 3' AAGCGGCTACCTGGTGTT; pnr 5' AGAAAACGGGAAGTGGTTCG and 3' CTGAGCGAGGGTTTGAGATC; tinc 5' ATCTGCATCTGAACTCGCTATG and 3' TCCAGGGTATCAAAGAGCATCC; Adhr 5' ACAAGAACGTGATTTTCGTTGC and 3' ACGGTCACCTTTGGATTGATTG; IA2 5' GCACTCCGAGGTCTGCTAC and 3' CTTCTCAATGTCCTCAACGTC and RpS17 5' 5' CGAACCAAGACGGTGAAGAAG and 3' CCTGCAACTTGATGGAGATACC.

Determining whether Yan/Gro loss is associated with differential gene expression

For Fig. 3G, MAnorm was used to first define regions of Yan loss. Regions with a normalized M value >0.5 were selected as high-confidence peaks showing reduced Yan occupancy. The read density of Gro was then determined at the midpoint of each of these regions in WT and mutant datasets, and a ratio calculated. Regions were selected as having reduced Yan and Gro occupancy if the M value was >0.5 and Gro ratio was <0.5. These regions were then assigned to the nearest gene and cross-referenced to genes in the microarray.

To assess Gro occupancy change at differentially expressed genes (Fig. S3C), the genome coordinates of either i) genes associated with upregulated probes (shaded red in Fig. 2A) or ii) all genes bound by Groucho were determined and cross-referenced with read densities of Groucho occupancy in both wildtype and *pnt* mutant datasets. Ratios of Groucho occupancy in *pnt* mutants relative to the wildtype control were calculated and used to bin genes into distinct categories of ratios of <0.5, 0.5-1, 1-1.5, or 1.5-2. For genes that were associated with multiple Gro peaks, the most substantial wildtype peak was identified and used to assess the ratio of Gro read density in the *pnt* mutant vs wildtype control.

Supplemental References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene Ontology: tool for the unification of biology. *Nat Genet*. https://www.nature.com/articles/ng0500_25 (Accessed January 22, 2018).

Bailey TL, Machanick P. 2012. Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res* **40**: e128–e128.

Bergman CM, Carlson JW, Celniker SE. 2005. Drosophila DNase I footprint database: a systematic genome annotation of transcription factor binding sites in the fruitfly, Drosophila melanogaster. *Bioinformatics* **21**: 1747–1749.

Kulakovskiy IV, Favorov AV, Makeev VJ. 2009. Motif discovery and motif finding from genome-mapped DNase footprint data. *Bioinformatics* **25**: 2318–2325.

Kulakovskiy IV, Makeev VJ. 2009. Discovery of DNA motifs recognized by transcription factors through integration of different experimental sources. *Biophysics* **54**: 667–674.

Nègre N, Brown CD, Ma L, Bristow CA, Miller SW, Wagner U, Kheradpour P, Eaton ML, Loriaux P, Sealfon R, et al. 2011. A Cis-Regulatory Map of the Drosophila Genome. *Nature* **471**: 527–531.

Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, et al. 2012. BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**: 1919–1920.

Shazman S, Lee H, Socol Y, Mann RS, Honig B. 2014. OnTheFly: a database of Drosophila melanogaster transcription factors and their binding sites. *Nucleic Acids Res* **42**: D167–D171.

Webber JL, Zhang J, Cote L, Vivekanand P, Ni X, Zhou J, Nègre N, Carthew RW, White KP, Rebay I. 2013a. The Relationship Between Long-Range Chromatin Occupancy and Polymerization of the Drosophila ETS Family Transcriptional Repressor Yan. *Genetics* **193**: 633–649.

Webber JL, Zhang J, Mitchell-Dick A, Rebay I. 2013b. 3D chromatin interactions organize Yan chromatin occupancy and repression at the even-skipped locus. *Genes Dev* **27**: 2293–2298.

Yamaguchi A, Wu M-F, Yang L, Wu G, Poethig RS, Wagner D. 2009. The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL, and APETALA1. *Dev Cell* **17**: 268–278.

Zhu LJ, Christensen RG, Kazemian M, Hull CJ, Enuameh MS, Basciotta MD, Brasefield JA, Zhu C, Asriyan Y, Lapointe DS, et al. 2011. FlyFactorSurvey: a database of Drosophila transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res* **39**: D111–D117.