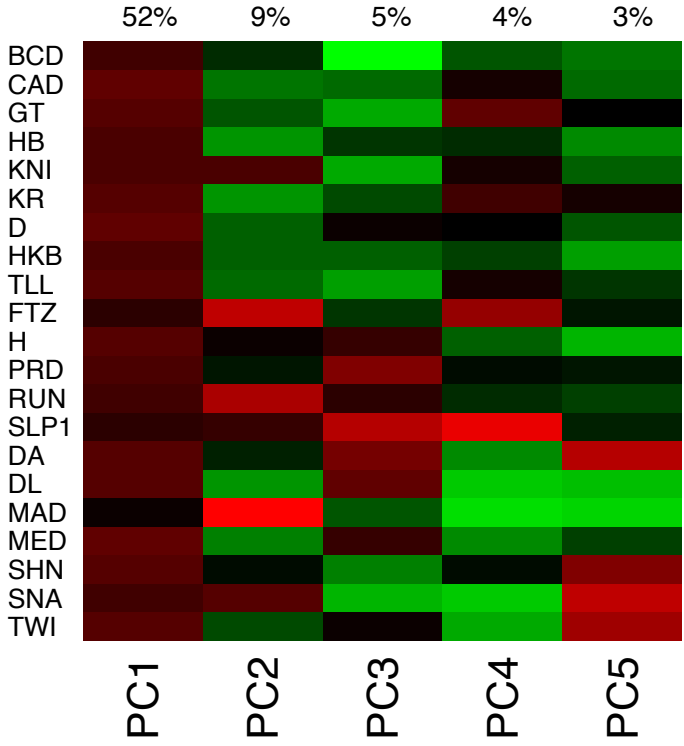
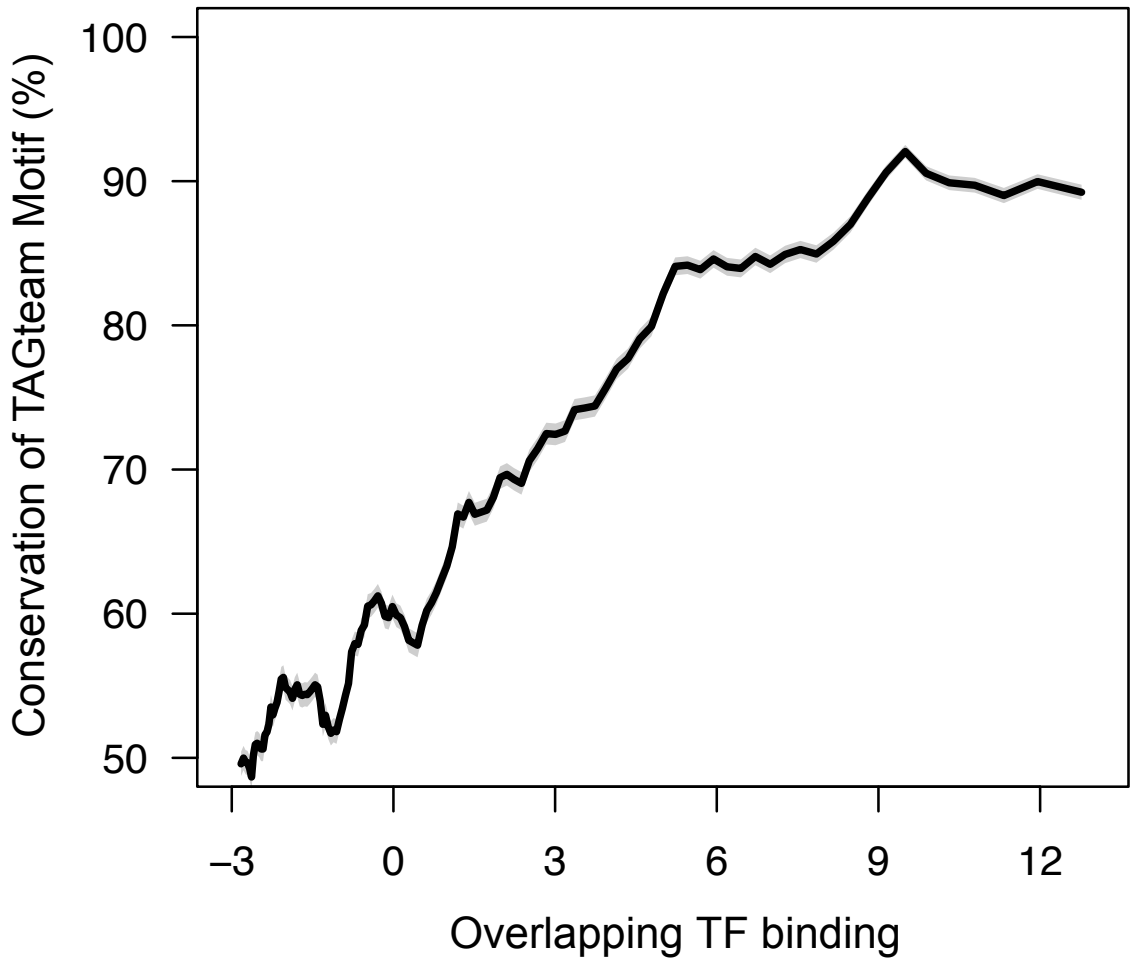


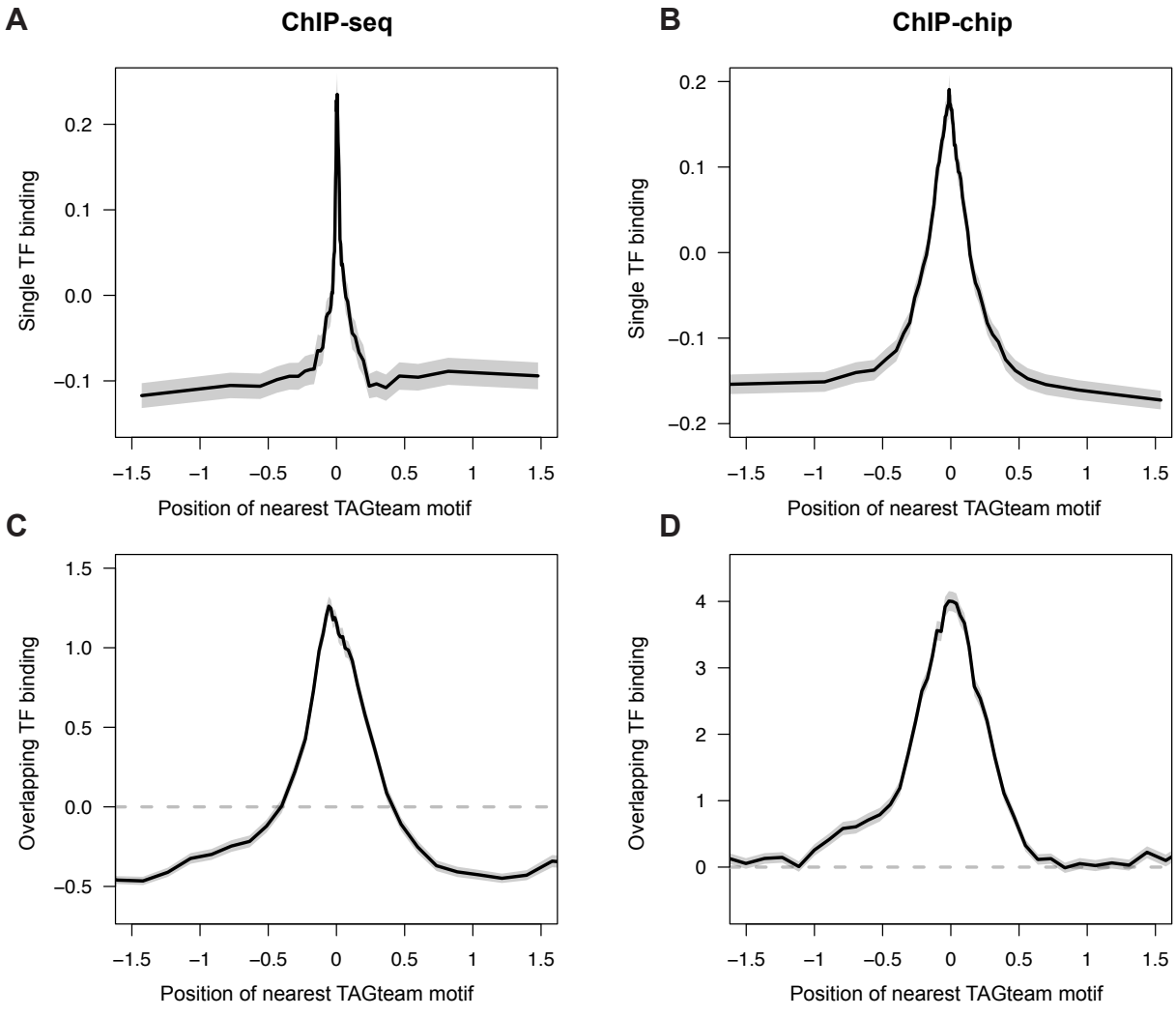
Supplementary Figure 1 Binding of all 21 factors is the primary mode of variation across the genome.

Plot shows the five most important components resulting from principal components analysis (PCA) on the normalized ChIP data for all 21 sequence-specific transcription factors. The color of each cell represents the sign (red positive, green negative) and magnitude of each transcription factor's contribution (row) to each principal component (column); the overall sign of each principal component is arbitrary. Note that each factor's contribution to PC1 has the same sign, indicating the majority (52%) of binding variation can be explained by a common mechanism affecting all of the transcription factors.



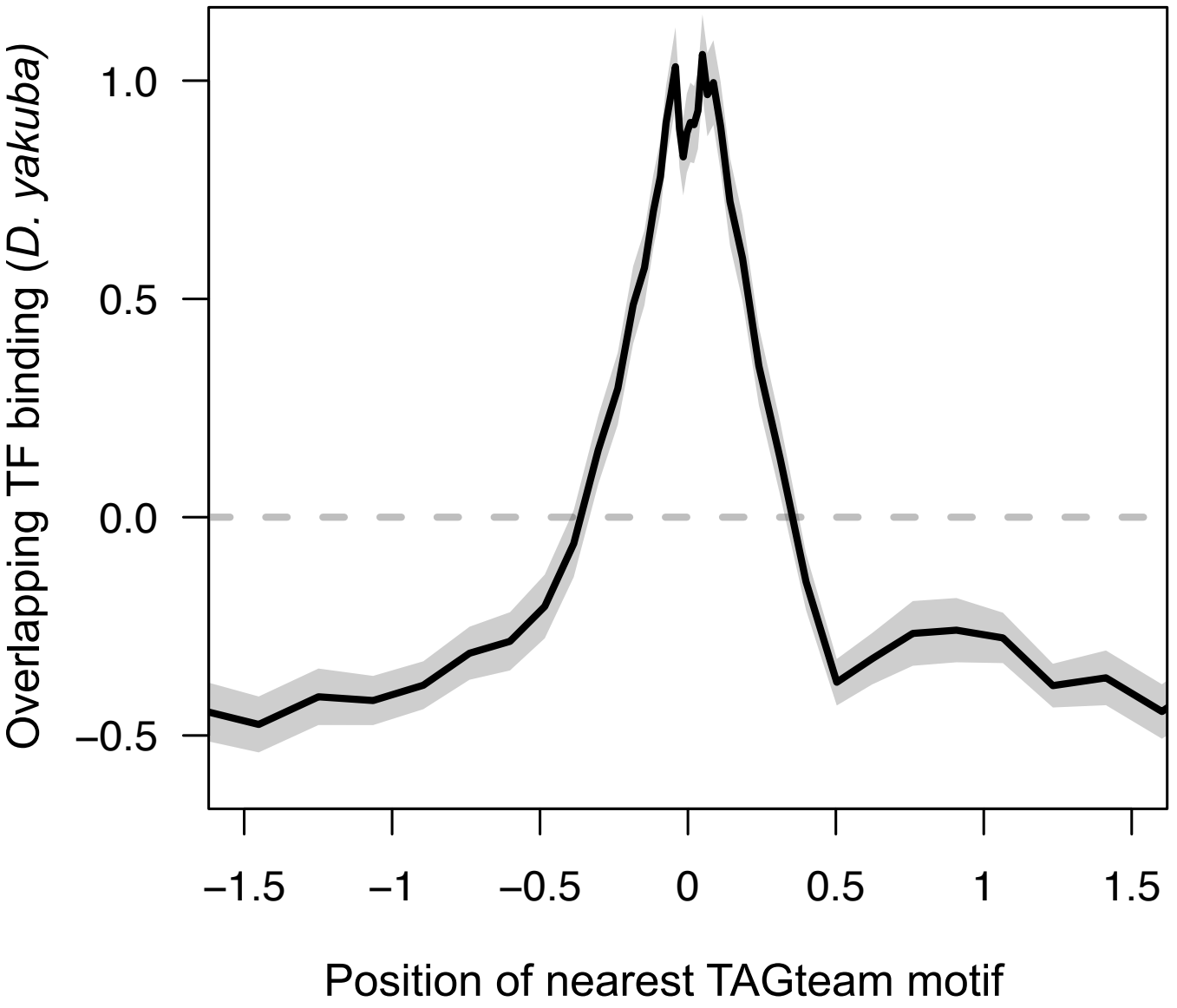
Supplementary Figure 2 Conservation of TAGteam motifs near ChIP peaks increases with overlapping TF binding.

For each ChIP peak in *D. melanogaster* with a nearby (within 500 bp) TAGteam motif, we assessed whether the nearest TAGteam motif was aligned to a TAGteam motif in *D. yakuba*. The three heptamers CAGGTAG, TAGGTAG, and CAGGTAA were considered to be TAGteam motifs, so if, for example, CAGGTAG in *D. melanogaster* was aligned to TAGGTAG in *D. yakuba*, then the motif was considered conserved.



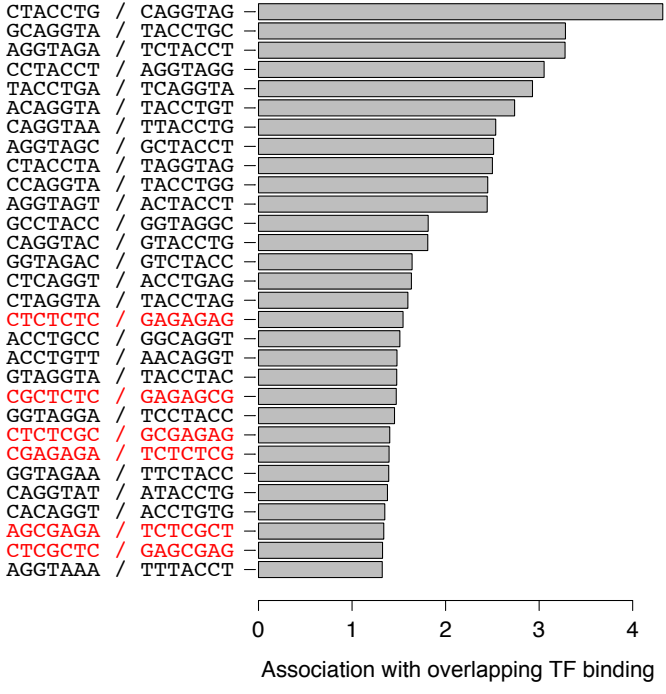
Supplementary Figure 3 ChIP-chip and ChIP-seq data give similar estimates of the TAGteam motif's range of effect.

(A) The ChIP-seq assay has an estimated spatial resolution of 20 bp. Plot indicates the position of the transcription factor motif ($p < 0.0001$) closest to each ChIP peak (x axis) and the corresponding binding strength at the ChIP peak (y axis), averaged over data for the six factors for which ChIP-Seq data was available. The spatial resolution of ChIP-seq was estimated as 20 bp based on the full width at half maximum of the figure. (B) The ChIP-chip assay has an estimated spatial resolution of 110 bp. Computed as (a), but using ChIP-chip data for the same six factors. (C-D) The ChIP-seq and ChIP-chip assays, despite their differing spatial resolutions, support similar estimates of the TAGteam's range of effect. Overlapping binding was computed using the first principal component resulting from PCA on only the ChIP-seq or ChIP-chip binding for these six factors.



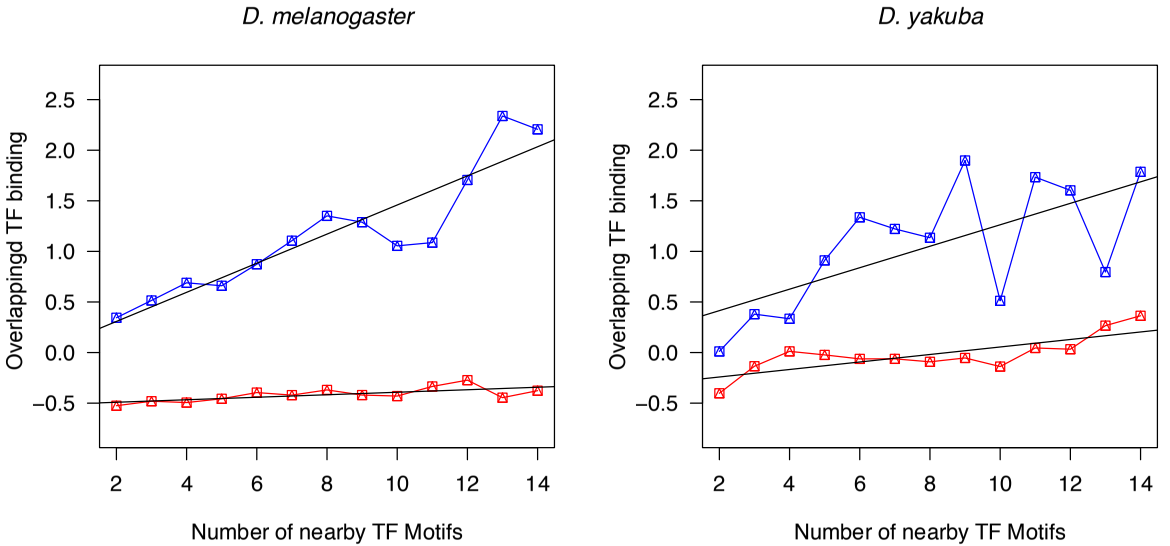
Supplementary Figure 4 ChIP peaks with nearby TAGteam motifs exhibit much higher levels of overlapping binding in *D. yakuba.*

Computed as Fig. 2c, but for ChIP-seq data in *D. yakuba*.



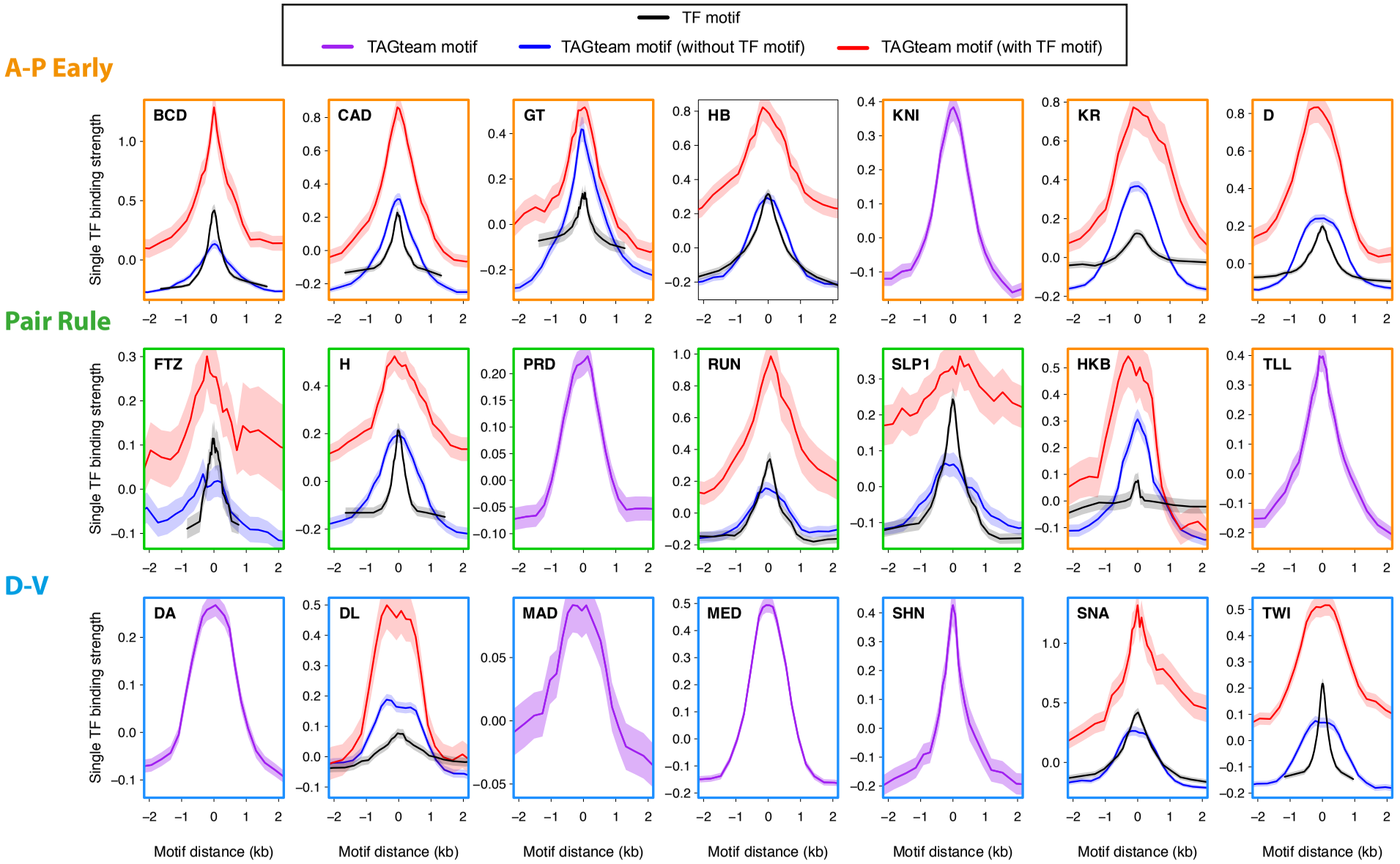
Supplementary Figure 5 TAGteam motifs are uniquely associated with high levels of overlapping binding.

We ranked each heptamer, excluding heptamers with single (five or more) or dinucleotide (two or more) repeats, by its ability to induce overlapping binding, measured as the peak height illustrated in Fig. 3c but restricted to a single heptamer. Of the top 30 heptamers, 24 were closely related to the TAGteam motif, sharing at least 5 bases of overlap with the consensus motif CAGGTAG. The remaining six heptamers, shown in red, are closely related (within 1bp) of a GA dinucleotide repeat (GAGA motif).



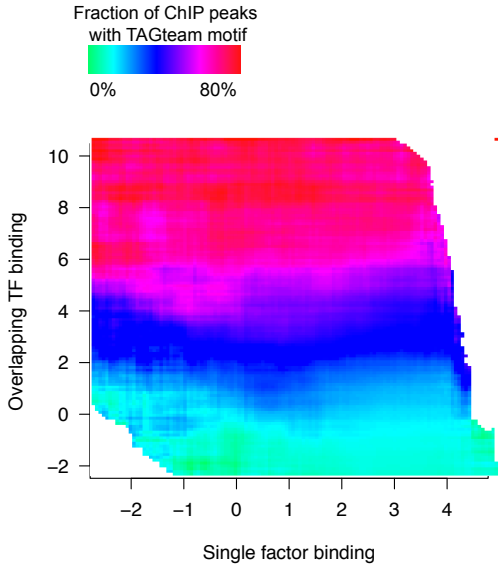
Supplementary Figure 6 TAGteam motifs demarcate functional clusters of TFBS in both *D. melanogaster* and D. *yakuba*

For two sister *Drosophila* species, we calculated the number of TF motifs for five A-P factors with well-defined recognition sequences (bcd, hb, cad, gt, kr) located in the 1000bp surrounding the summits of ChIP-Seq peaks, and searched for TAGteam motifs in the same region. In both species, clustering of recognition sequences is only associated with an increase in overlapping binding strength when a TAGteam motif is present.

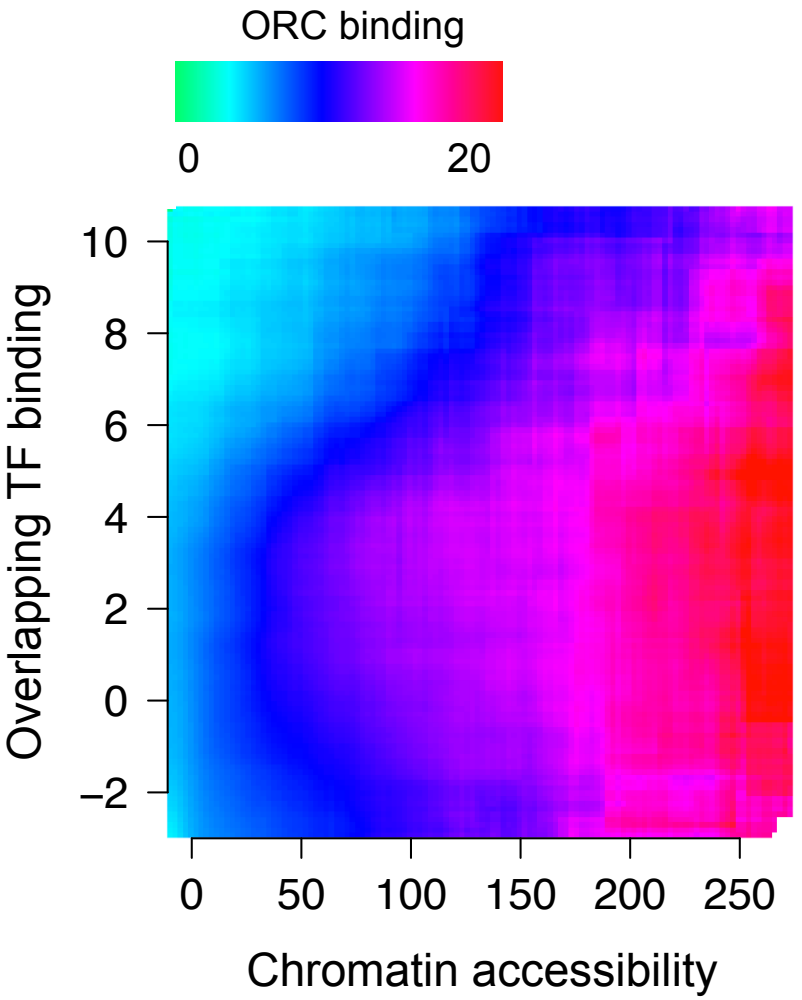


Supplementary Figure 7 TAGteam motifs act additively with additional transcription factor binding sites.

For each of the 15 transcription factors with well-defined binding sites enriched in the ChIP data, we classified its ChIP peaks based on whether they had nearby binding sites ($p < 0.0001$; within 100 bp) and/or TAGteam motifs (within 500 bp). For the remaining 6 factors without well-defined binding sites, we aggregated all peaks. Plot indicates the position of cognate recognition motifs or TAGteam motifs (red, blue, and purple) closest to each ChIP peak (x axis) and the corresponding binding strength at the ChIP peak (y axis).

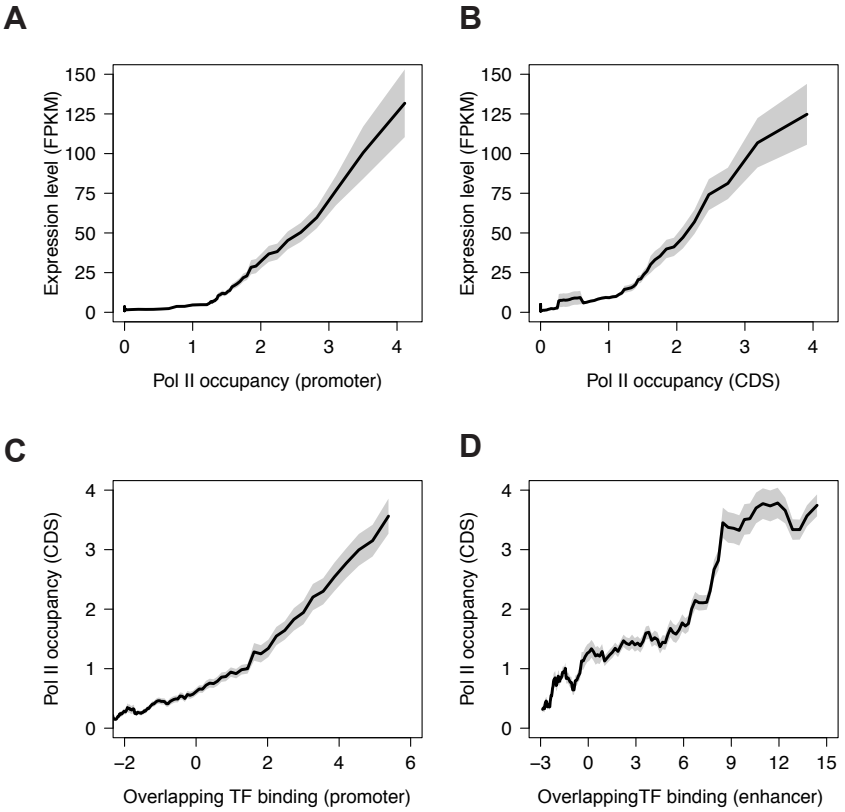


Supplementary Figure 8 Correlations between single factor binding and TAGteam motif occurrence disappear after controlling for overlapping TF binding.



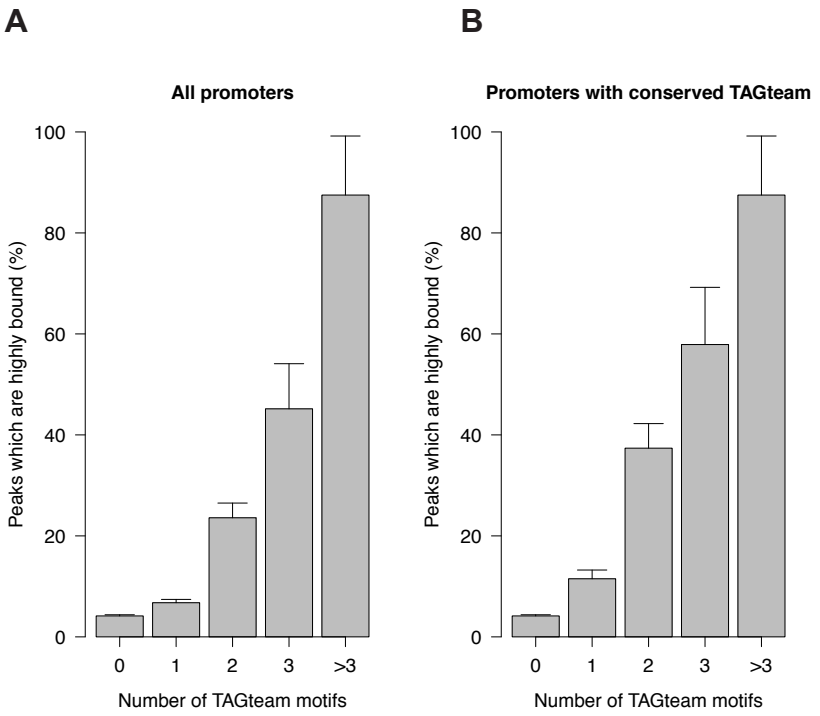
Supplementary Figure 9 ORC localization is primarily determined by chromatin accessibility, not TF binding.

Plot shows ORC binding strength (arbitrary units) for different levels of chromatin accessibility and overlapping binding. After controlling for overlapping binding levels (looking along a row), increases in chromatin accessibility are strongly correlated with increases in ORC binding strength. Conversely, after controlling for chromatin accessibility (looking along a column), increases in overlapping binding strength are associated with a mild reduction in ORC binding.

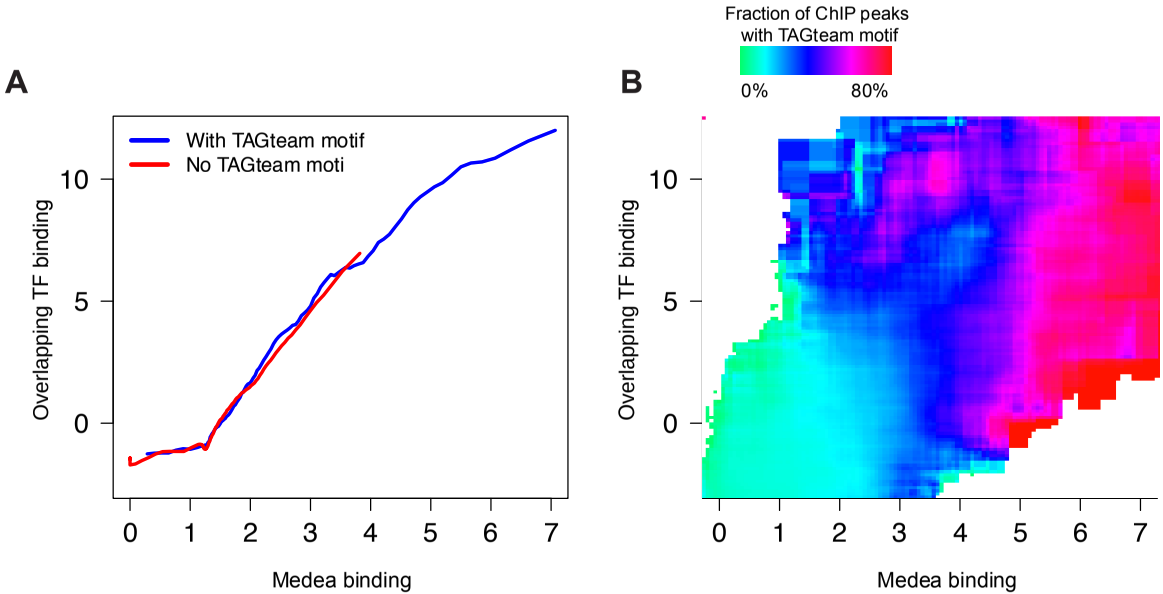


Supplementary Figure 10 Overlapping binding and gene expression correlate with Pol II occupancy at both the promoter and downstream coding sequence.

For all genes identified to be zygotically transcribed by (Lott et al. 2011), we quantified Pol II occupancy at the CDS by calculating the average Pol II occupancy between the translation start and stop sites. CDS occupancy values agreed closely with promoter occupancy values, resulting in highly similar correlations with gene expression values (A, B), as well as overlapping TF binding (C,D compare to Fig. 5C and Fig. 5F).



Supplementary Figure 11 Clusters of TAGteam binding sites identify highly bound promoter regions.



Supplementary Figure 12 Medea is a putative interaction partner for Zelda.

(A) After controlling for the maximum binding strength of Medea near (within 200bp) the peak summit, the presence of TAGteam motifs is not associated with an increase in overlapping transcription factor binding. (B) The fraction of peaks with nearby TAGteam motifs increases with Medea binding strength, even after controlling for levels of overlapping binding. Plot shows the fraction of peaks with nearby (within 500 bp) TAGteam motifs for different levels of Medea binding and overlapping binding. Empty (white) space in the plot indicates levels of Medea binding and overlapping binding for which insufficient data was available.