

Supplementary Figure 3. Alignment and Visualization of Motifs from Fly, Mouse and Human HD Family

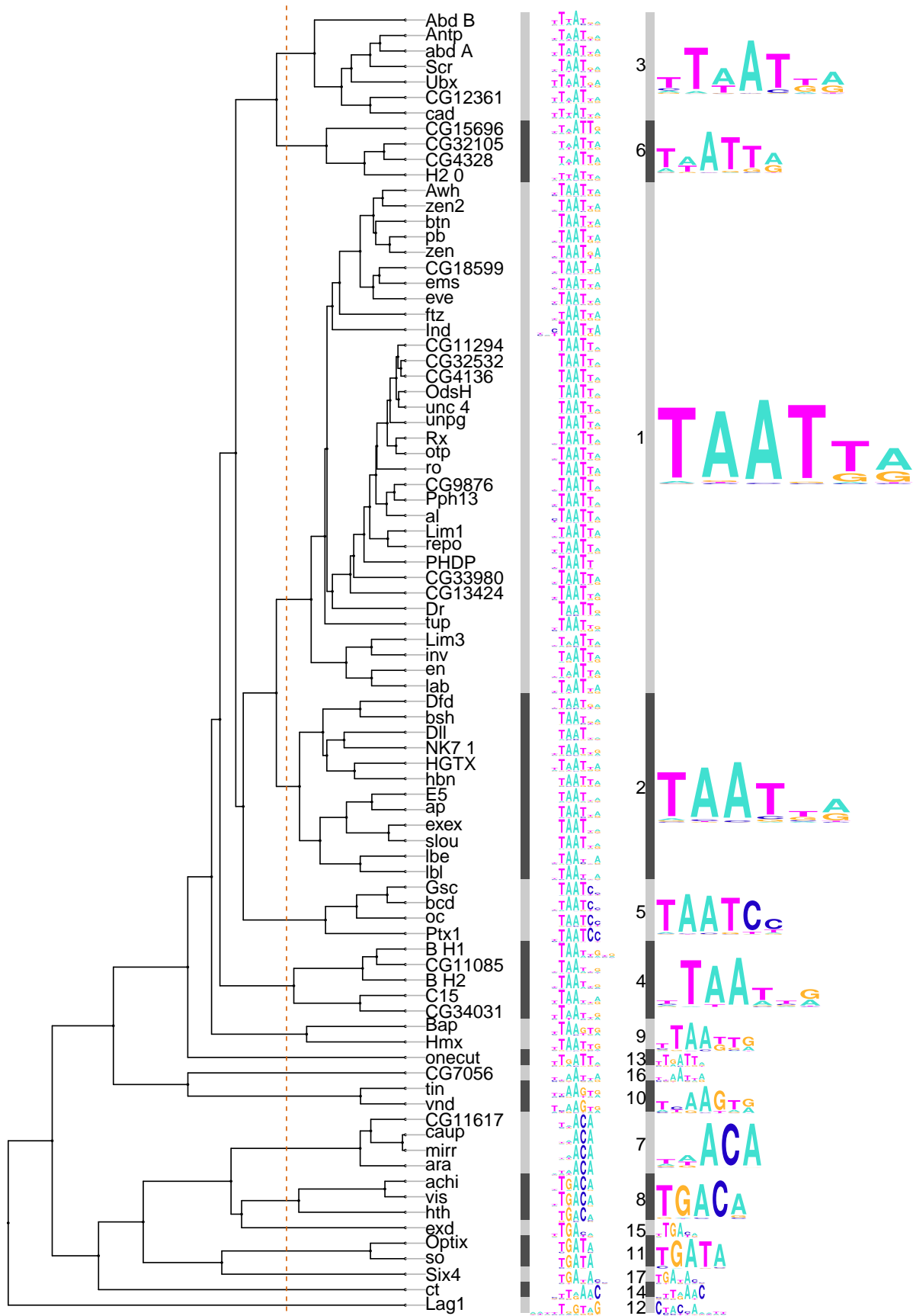
Contents

Supplementary Figure 3A: Linear dendrogram of all *Drosophila* HD motifs. This example depicts the relationships between TF motifs, including aligned motifs for individual TFs (center) and motif signatures for motif clusters (far right), where the logo size scales with the number of members within the cluster. Motif clusters are determined using a distance threshold of 2 (dashed magenta line), and are delineated using alternating light and dark gray bars numbered based on cluster size.

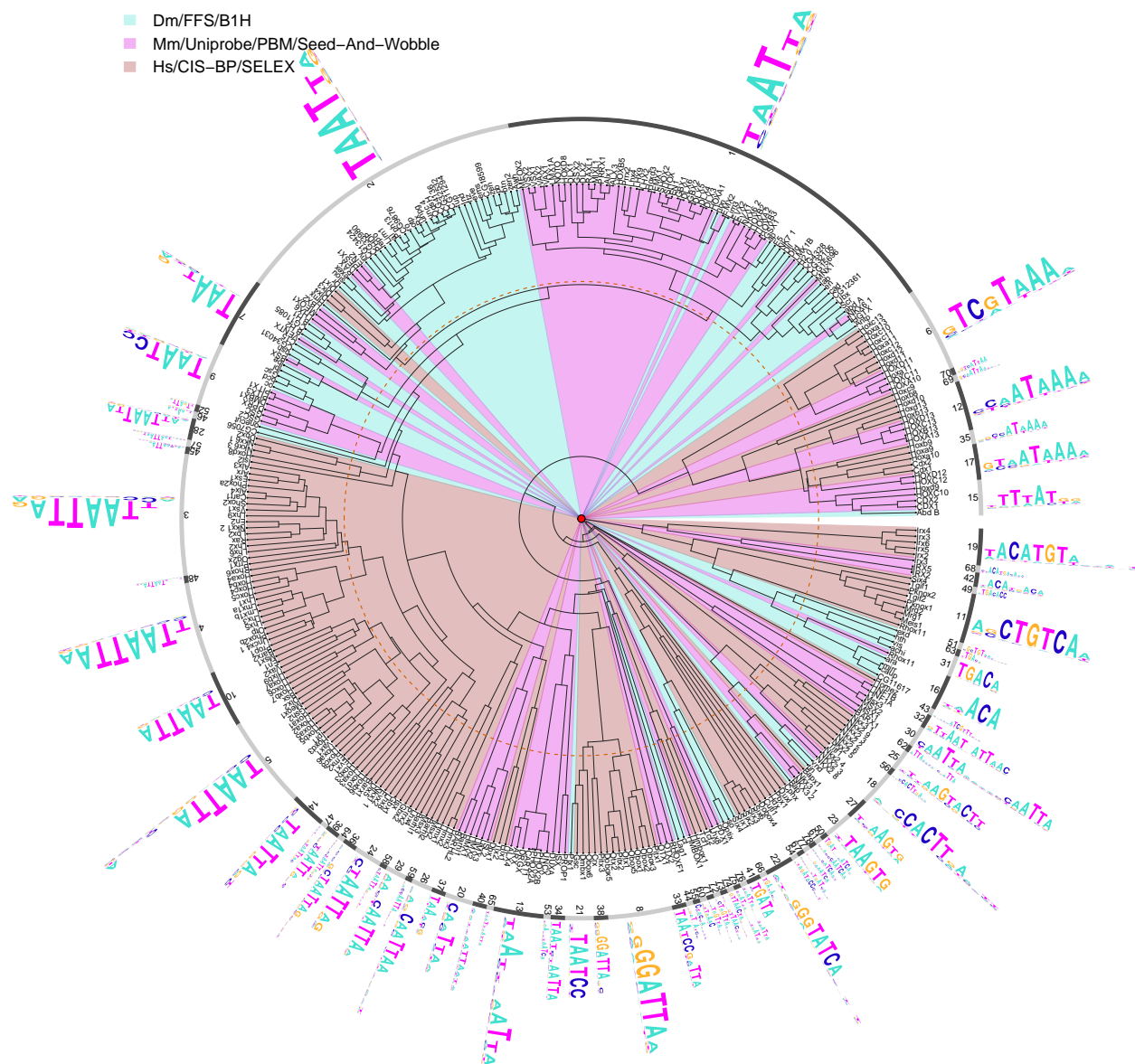
Supplementary Figure 3B: Joint alignment of motifs from fly, mouse and human by MatAlign. MatAlign was used to align and cluster motifs from human (SELEX data, colored in brown), mouse (PBM data, colored in purple) and fly (B1H data, colored in turquoise) HD domain proteins. The length of tree branches reflects motif distances. Motif clusters determined with a distance threshold of 2.5 are indicated by the alternating light and dark gray bars that are numbered by size, with the largest group listed as cluster 1. A single merged motif is shown for each cluster.

Supplementary Figure 3C: Joint alignment of motifs from fly, mouse and human by MotIV. MotIV was used to align and cluster motifs using a distance threshold of 0.005 with the same datasets and display settings as Supplementary Figure 3B.

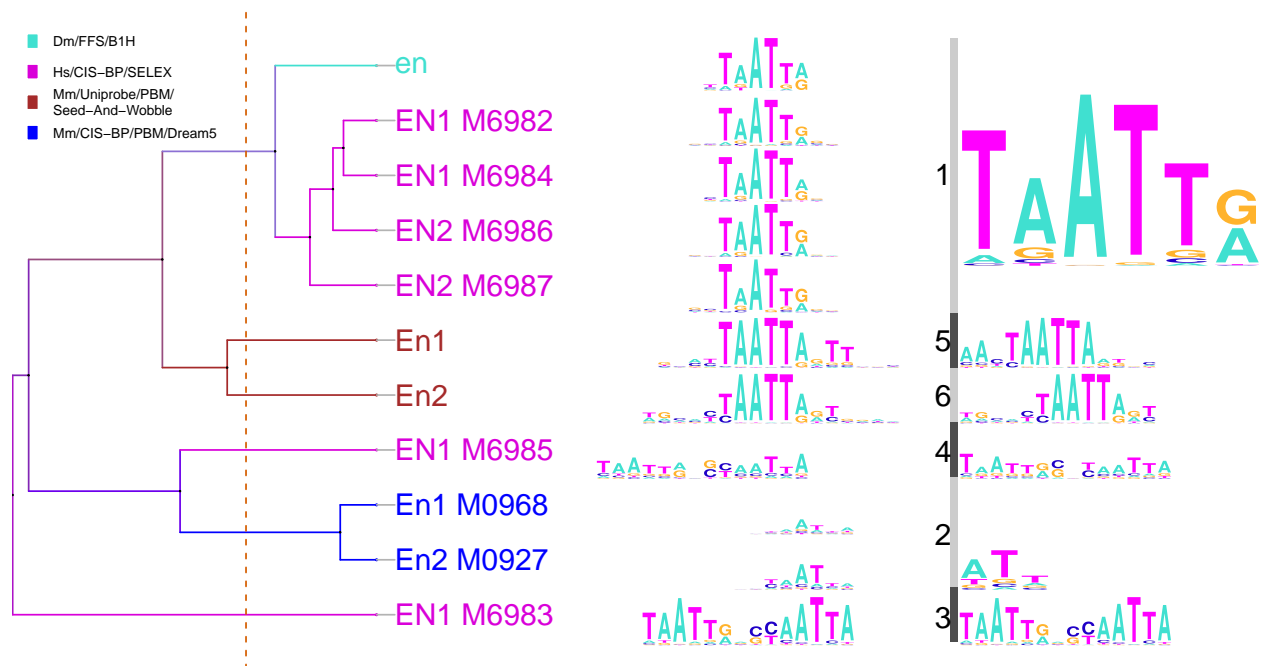
Supplementary Figure 3D: Alignment of HD motifs of EN homologs from *Drosophila*, mouse and human. This example illustrates how different experimental and computational methods influence motif clustering for orthologous proteins. Motif clusters are determined using a distance threshold of 2 (dashed magenta line). Dm/FFS/B1H (colored in turquoise) represents fly B1H binding site data. Hs/CIS-BP/SELEX (colored in purple) represents human SELEX data. Mm/Uniprobe/PBM/Seed-And-Wobble (colored in magenta) represents mouse PBM data analyzed by the Seed and Wobble motif discovery tool. Mm/CIS-BP/PBM/Dream5 (colored in blue) is mouse PBM data analyzed by the Dream5 motif discovery tool.



Supplementary Figure 3A: Linear dendrogram of all Drosophila HD motifs. This example depicts the relationships between TF motifs, including aligned motifs for individual TFs (center) and motif signatures for motif clusters (far right), where the logo size scales with the number of members within the cluster. Motif clusters are determined using a distance threshold of 2 (dashed magenta line), and are delineated using alternating light and dark gray bars numbered based on cluster size.



Supplementary Figure 3B: Joint alignment of motifs from fly, mouse and human by MatAlign. MatAlign was used to align and cluster motifs from human (SELEX data, colored in brown), mouse (PBM data, colored in purple) and fly (B1H data, colored in turquoise) HD domain proteins. The length of tree branches reflects motif distances. Motif clusters determined with a distance threshold of 2.5 are indicated by the alternating light and dark gray bars that are numbered by size, with the largest group listed as cluster 1. A single merged motif is shown for each cluster.



Supplementary Figure 3D: Alignment of HD motifs of EN homologs from Drosophila, mouse and human. This example illustrates how different experimental and computational methods influence motif clustering for orthologous proteins. Motif clusters are determined using a distance threshold of 2 (dashed magenta line). Dm/FFS/B1H (colored in turquoise) represents fly B1H binding site data. Hs/CIS-BP/SELEX (colored in purple) represents human SELEX data. Mm/Uniprobe/PBM/Seed-And-Wobble (colored in magenta) represents mouse PBM data analyzed by the Seed and Wobble motif discovery tool. Mm/CIS-BP/PBM/Dream5 (colored in blue) is mouse PBM data analyzed by the Dream5 motif discovery tool.