

Supplementary Figure 7

Contents

Dimeric motifs separated from monomeric motifs from the same TF	2
Unique clusters	12
Cleaned unique clusters	13

```
##load the library
library(motifStack)

getMatAlignOut <- function(pcmpath, outpath="output",
                           groupDistance=NA, trim=0.2){
  pcms <- readPCM(pcmpath)
  pfms<-lapply(pcms,pcm2pfm)
  matalign_path <- "./app/matalign-v4a"
  neighbor_path <- "./app/neighbor.app/Contents/MacOS/neighbor"
  system(paste("perl MatAlign2tree.pl --in . --pcmpath", pcmpath,
               "--out", outpath,
               "--matalign", matalign_path,
               "--neighbor", neighbor_path,
               "--tree","UPGMA"))
  newickstrUPGMA <-
    readLines(con=file.path(outpath, "NJ.matalign.distMX.nwk"))
  phylog <- newick2phylog(newickstrUPGMA, FALSE)
  leaves <- names(phylog$leaves)
  motifs <- pfms[leaves]
  if(!is.na(groupDistance)){
    motifSig <-
      motifSignature(motifs, phylog,
                    groupDistance=groupDistance,
                    min.freq=1, trim=trim)
    sig <- signatures(motifSig)
    gpCol <- sigColor(motifSig)
  }else{
    motifSig <- NA
    sig <- NA
    gpCol <- NA
  }

  return(list(phylog=phylog, sig=sig, gpCol=gpCol,
             motifs=DNAmotifAlignment(motifs),
             leaves=leaves,
             unaligned.pfms=motifs))
}

pcmpath <- dir("pcmsUni", include.dirs=TRUE)
matAlignOut <- lapply(file.path("pcmsUni", pcmpath),
                      getMatAlignOut, groupDistance=4)

description <- c("## Dimeric motifs separated from monomeric motifs from the same TF\n\nFor most dimeric
colorSet <- c("Dm"="#00FC00",
              "Mm"="brown", "Ms"="#F69156",
              "Hs"="#D900D9")
```

```
## kexpand is function to extend knitr.
# kexpand<-function(.ele, id, cap, figheight, des){
#   text <- paste("{des}}\n``{r {cap}},fig.cap='{cap}',fig.height={{figheight}},echo=",
#   ifelse(id==1,"TRUE","FALSE"),
#   "}\n
#   leaveNames <- gsub('^(Dm|Mm|Ms|Hs)_', '', .ele$leaves)\n
#   species <- colorSet[gsub("\^(Dm|Mm|Ms|Hs).*$", "\\|1", .ele$leaves)]\n
#   motifPiles(phylog=.ele$phylog, .ele$motifs,
#   .ele$sig,
#   col.tree=species, col.leaves=species,
#   col.pfms2=.ele$gpCol,
#   col.pfms2.width=.01, labels.leaves=leaveNames,
#   plotIndex=c(FALSE,TRUE), IndexCex=1.5,
#   groupDistance=4, clabel.leaves=3)\n
#   ``\n", sep="")
#   cat(knit(text=knit_expand(text=text)))
# }
```

```
sta <- mapply(function(.ele, name, ID, des){
  kexpand(.ele, id=ID, cap=paste0("Supplementary Figure 7.", name),
    figheight=ceiling(.25*length(.ele$leaves)), des)
}, matAlignOut, pcmpath, 1:length(pcmpath), description)
```

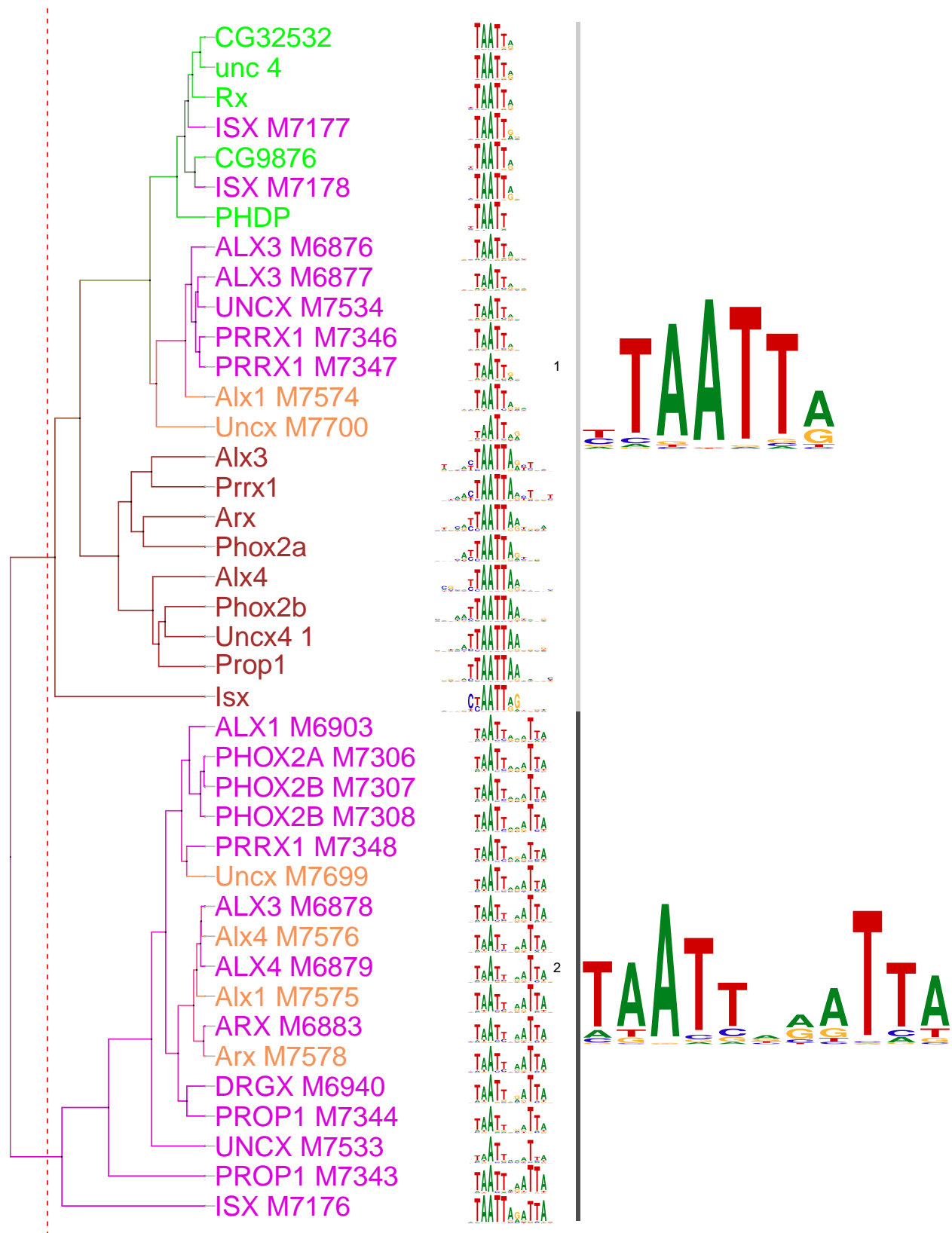
Dimeric motifs separated from monomeric motifs from the same TF

For most dimeric clusters, one or more monomeric motifs cluster with a Drosophilar motif, suggesting that the absence of a fly TF in a dimeric cluster may be a methodological difference rather than a biological one.

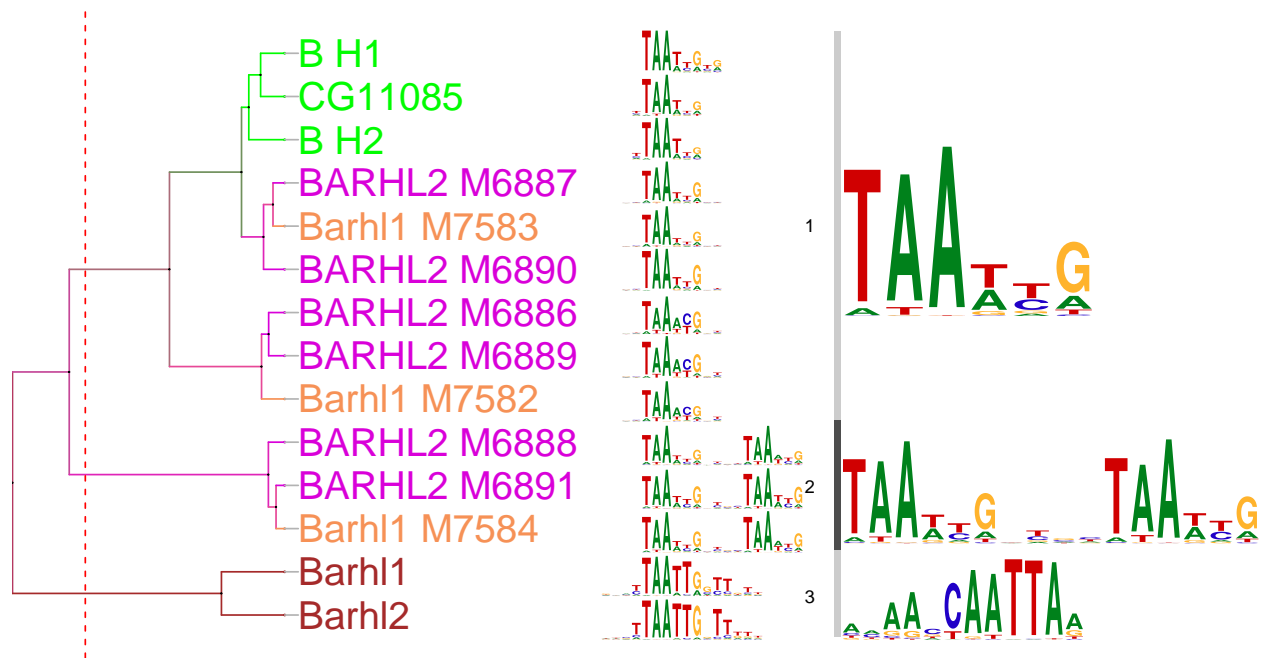
```
leaveNames <- gsub('^(Dm|Mm|Ms|Hs)_', '', .ele$leaves)

species <- colorSet[gsub("\^(Dm|Mm|Ms|Hs).*$", "\\|1", .ele$leaves)]

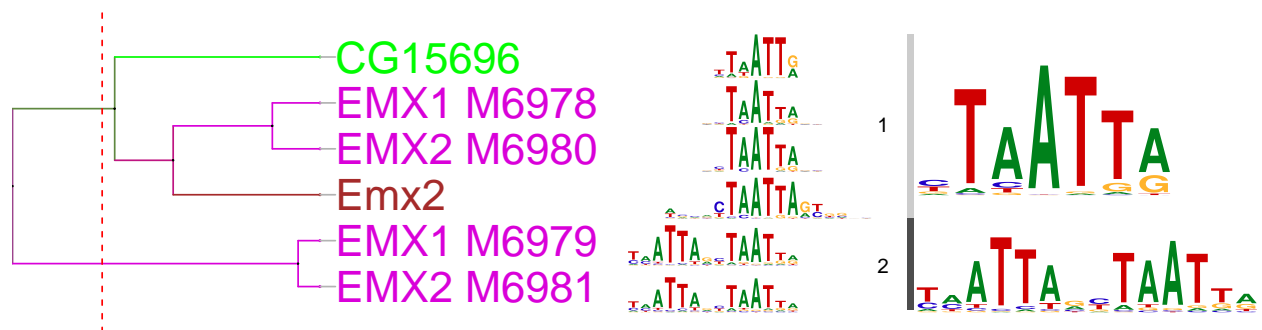
motifPiles(phylog=.ele$phylog, .ele$motifs,
  .ele$sig,
  col.tree=species, col.leaves=species,
  col.pfms2=.ele$gpCol,
  col.pfms2.width=.01, labels.leaves=leaveNames,
  plotIndex=c(FALSE,TRUE), IndexCex=1.5,
  groupDistance=4, clabel.leaves=3)
```



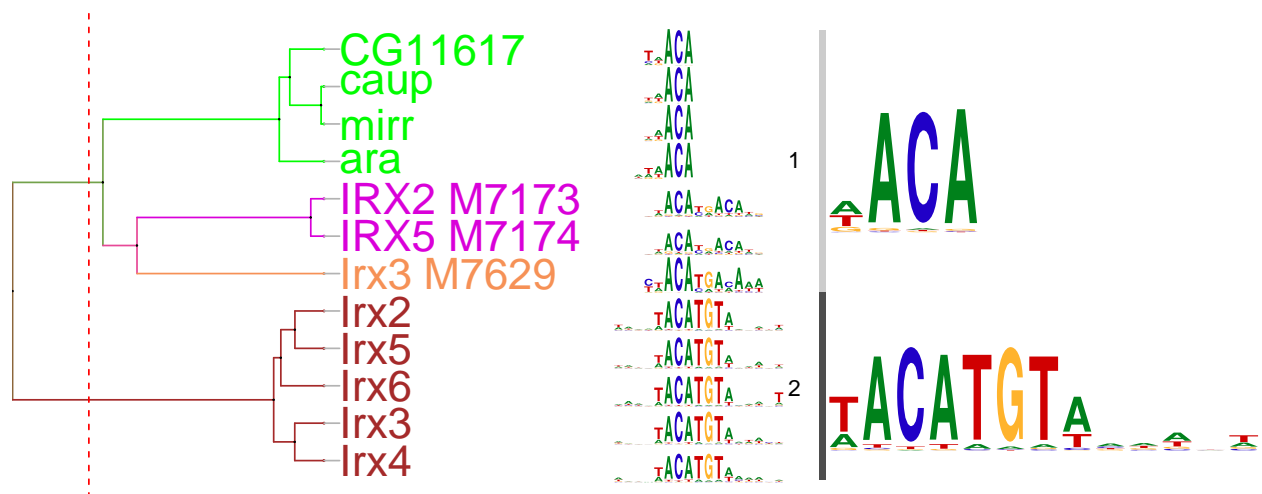
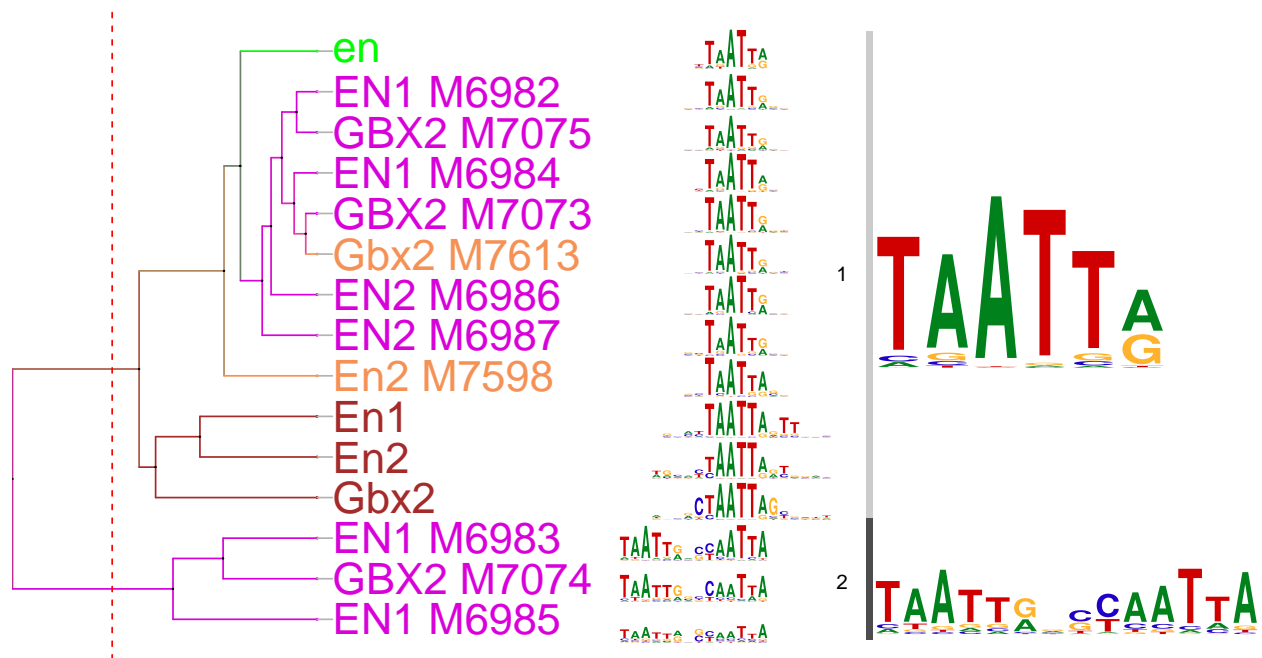
Supplementary Figure 7.01.DimerClusterALX

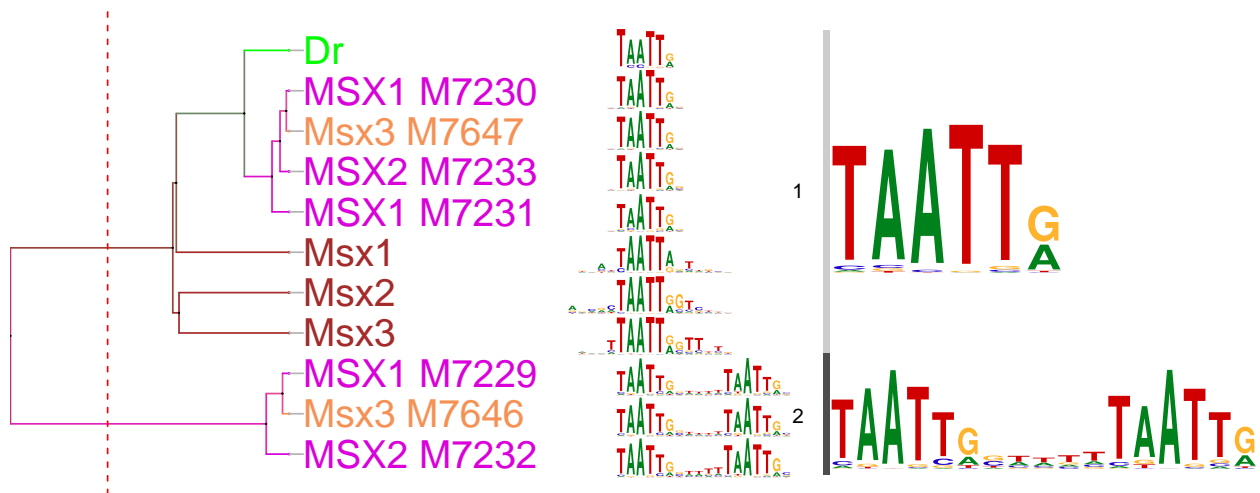


Supplementary Figure 7.02.DimerClusterBARHL



Supplementary Figure 7.03.DimerClusterEMX





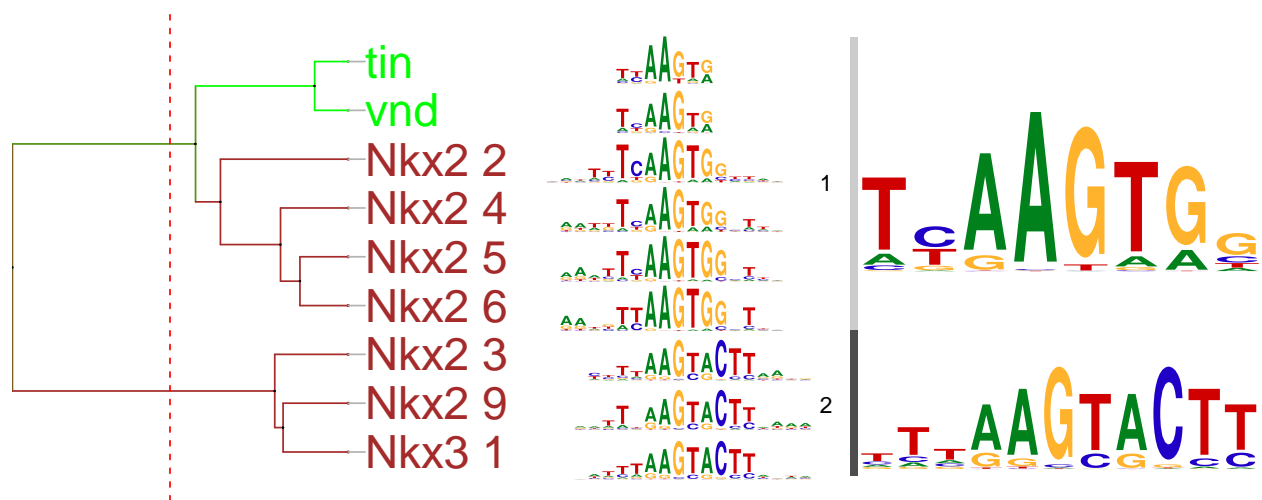
Supplementary Figure 7.06.DimerClusterMSX

There is no monomeric motif in mammals.



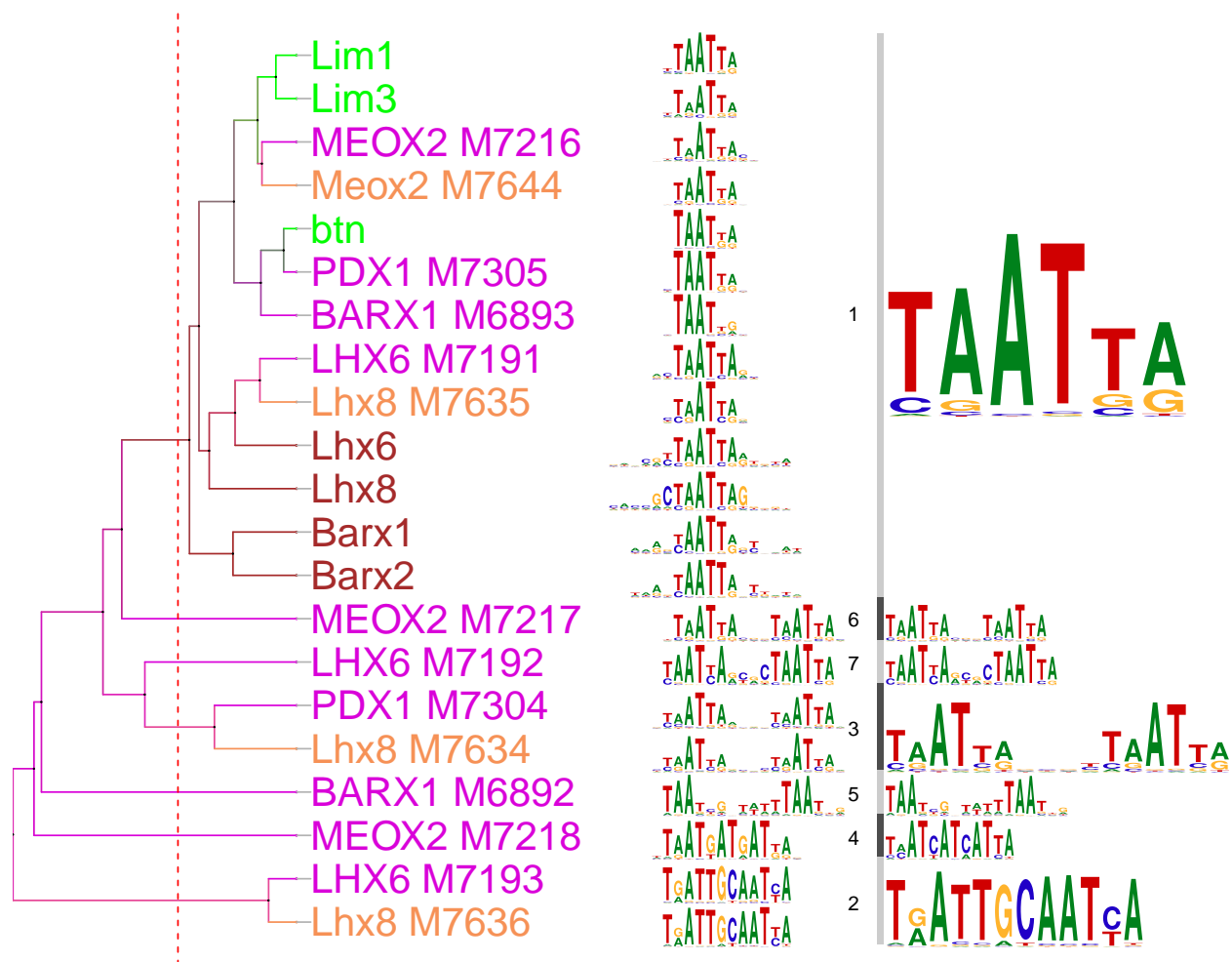
Supplementary Figure 7.07.DimerClusterCphx

The dimeric motif cluster is different from the above clusters. There are overlaps for the subunits of monomeric motif.



Supplementary Figure 7.08.DimerClusterNkx

For MEOX/LHX cluster, there are some exceptions like LHX6_M7193 and Lhx_M7636. The subunit of both has strong G after TAATT comparing to other monomeric motifs with strong A after TAATT.



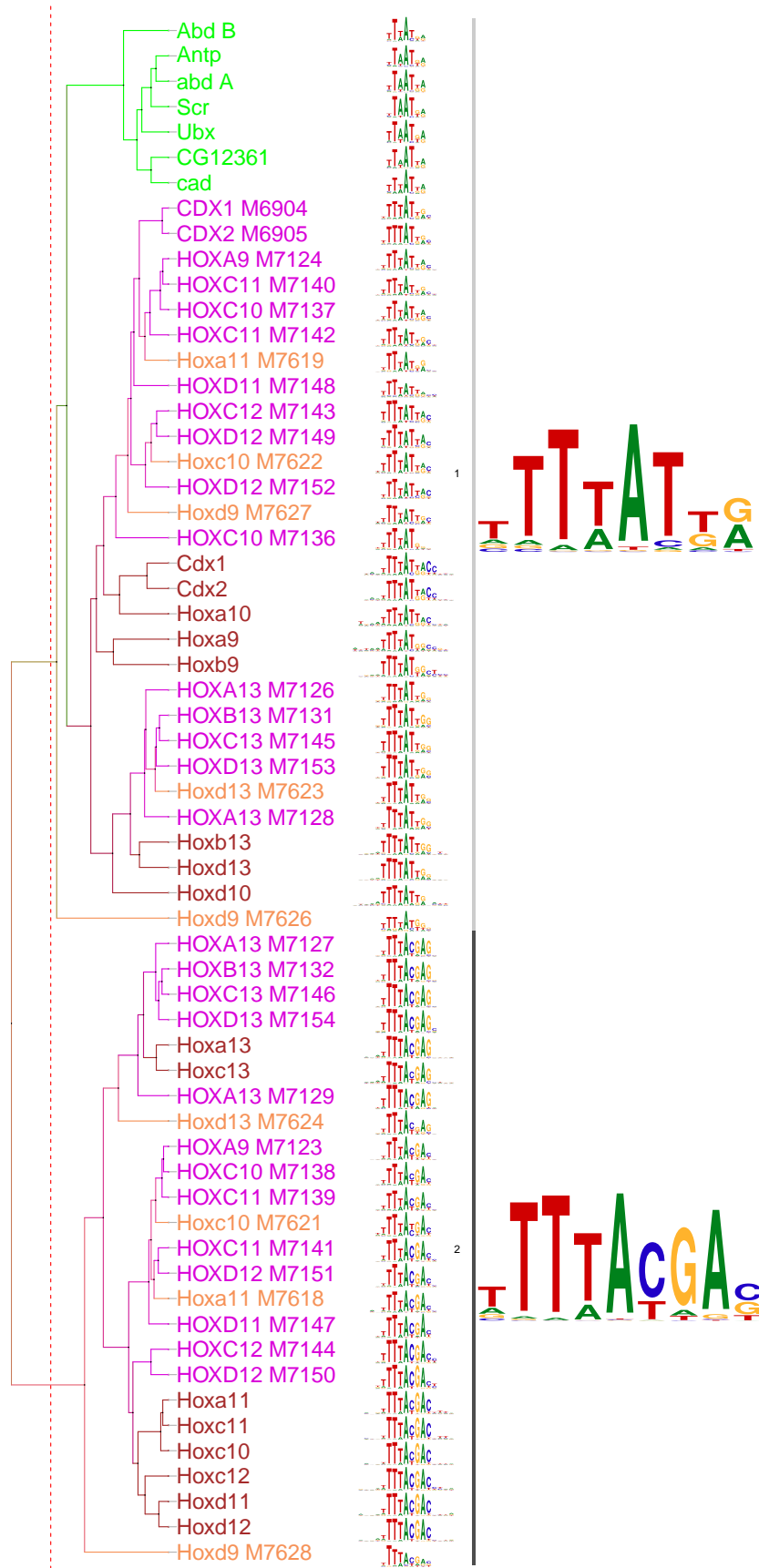
Supplementary Figure 7.09.DimerClusterLHX

Difference by platform.

The absence of a fly TF in a dimeric cluster may be a methodological difference rather than a biological one.

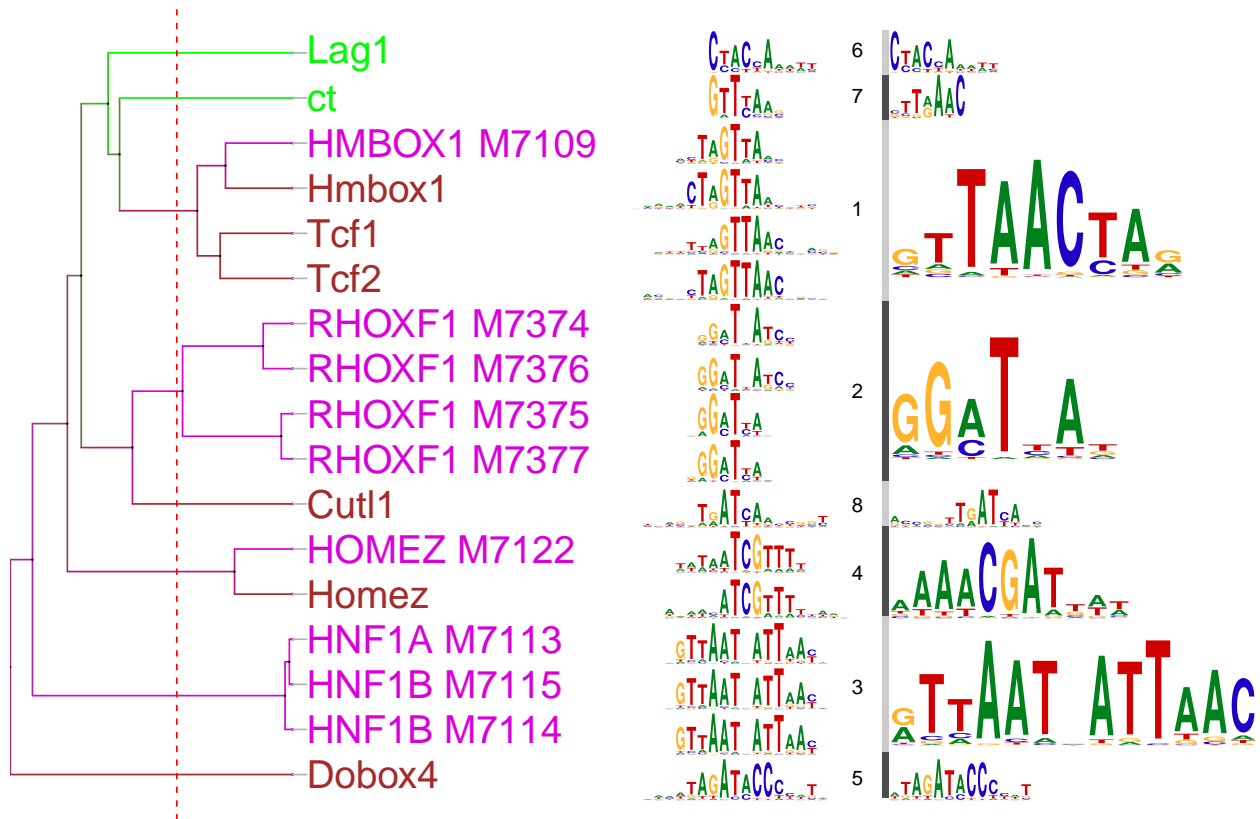


Supplementary Figure 7.11.PlatformClusterHmxCluster



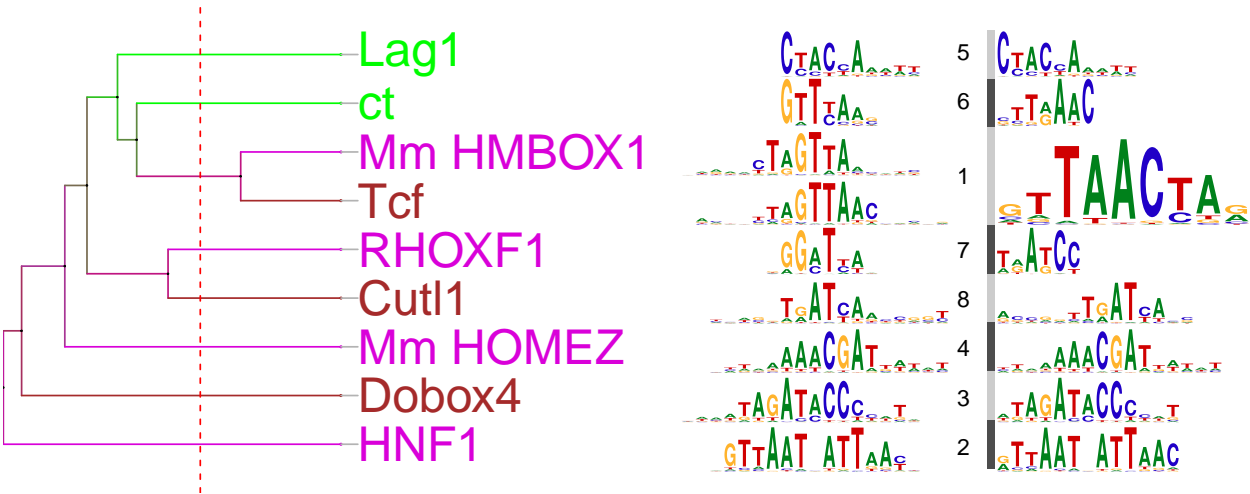
Supplementary Figure 7.12.PlatformClusterHoxCluster

Unique clusters



Supplementary Figure 7.13.UniqueCluster

Cleaned unique clusters



Supplementary Figure 7.14.UniqueCleanCluster