

Supplementary Figure 4

Contents

Supplementary Figure 4A. alignment of <i>Drosophila</i> motifs by matAlign	2
Supplementary Figure 4B. alignment of <i>Drosophila</i> motifs by MotIV	3
Supplementary Figure 4C. fly vs mouse, by matAlign	3
Supplementary Figure 4D. fly vs human, by matAlign	6
Supplementary Figure 4E. mouse selex vs mouse pbm/sw vs mouse pbm dream5	7
Supplementary Figure 4F. mouse selex vs mouse pbm/sw vs mouse pbm dream5, by motIV	10
Supplementary Figure 4G. Engrailed orthologs, by matAlign	12
Supplementary Figure 4H. Engrailed orthologs, by motIV	15
Supplementary Figure 4I. fly vs ms vs hum matAlign	16
Supplementary Figure 4J. fly vs ms vs hum motIV	17

```
##load the library
library(motifStack)
colorSet <- c("Dm"="#00FC00", "b1h"="#00FC00", "sw"="#008080", "bml"="darkgreen",
            "Mm"="brown", "MmDREAM"="blue", "Ms"="#F69156",
            "Hs"="#D900D9")

## function to read example data
getMatAlignOut <- function(pcopath, outpath="output", groupDistance=2.5, trim=0.2){
  pcms <- readPCM(pcopath)
  pfms<-lapply(pcms,pcm2pfm)
  matalign_path <- "./app/matalign-v4a"
  neighbor_path <- "./app/neighbor.app/Contents/MacOS/neighbor"
  system(paste("perl MatAlign2tree.pl --in . --pcopath", pcopath, "--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))
}

pfmList2matrixList <- function(pfms){
  m <- lapply(pfms, function(.ele) as(.ele, "matrix"))
  names(m) <- unlist(lapply(pfms, function(.ele) .ele@name))
  m
}

getMotIVOut <- function(pcopath, cc, align, groupDistance=.005, trim=0.2){
  pcms <- readPCM(pcopath)
  pfms<-lapply(pcms,pcm2pfm)
  jaspar.scores <-
    MotIV::readDBScores(
      file.path(".", "app", "scores",
                  paste("JaspRand_", cc, "_", align, ".scores", sep="")))
  d <- MotIV::motifDistances(pfmList2matrixList(pfms), cc=cc, align=align)
  hc <- MotIV::motifHclust(d)
  phylog <- hclust2phylog(hc)
  pfms <- pfms[[hc$order]]
  aligned.pfms <- DNAmotifAlignment(pfms)
  leaveNames <-names(phylog$leaves)
  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))
}

```

Supplementary Figure 4A. alignment of *Drosophila* motifs by matAlign

```

gpDis <- 2
Fly_B1H <- getMatAlignOut("pcmsDatasetFly", groupDistance=gpDis)
attach(Fly_B1H)

leaveNames <- gsub("^Dm_ ", "", leaves)

motifPiles(phylog=phylog, motifs, sig,
           col.pfms=gpCol, col.pfms.width=.01,

```

```

col.pfms2=gpCol, col.pfms2.width=.01,
labels.leaves=leaveNames,
plotIndex=c(FALSE, TRUE), IndexCex=1,
groupDistance=gpDis, clabel.leaves=1)

detach(Fly_B1H)

```

Supplementary Figure 4B. alignment of *Drosophila* motifs by MotIV

```

gpDis <- 0.005
Fly_B1H.MotIV <- getMotIVOut("pcmsDatasetFly",
                               cc="ALLR", align="SWU",
                               groupDistance=gpDis)
attach(Fly_B1H.MotIV)

leaveNames <- gsub("^Dm_ ", "", leaves)
motifPiles(phylog=phylog, motifs, sig,
           col.pfms=gpCol, col.pfms.width=.01,
           col.pfms2=gpCol, col.pfms2.width=.01,
           labels.leaves=leaveNames,
           plotIndex=c(FALSE, TRUE), IndexCex=1,
           groupDistance=gpDis, clabel.leaves=1)

detach(Fly_B1H.MotIV)

```

Supplementary Figure 4C. fly vs mouse, by matAlign

```

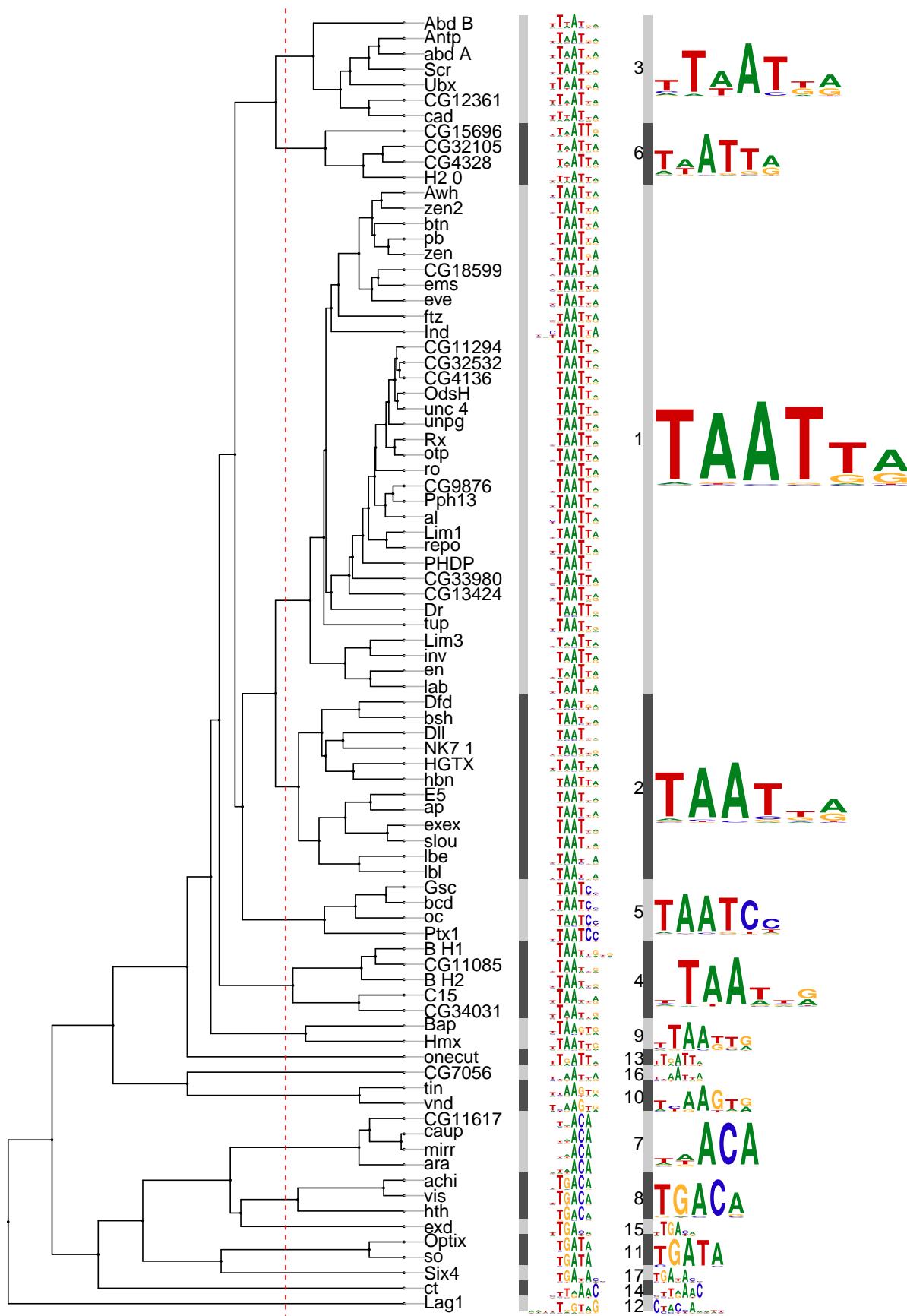
gpDis <- 2
FlyMousePBMSAW <- getMatAlignOut("pcmsDatasetFlyMousePBM.SAW",
                                    groupDistance=gpDis, trim=0.2)
attach(FlyMousePBMSAW)

##set methods color
species <- gsub("(Dm|Mm)_.*$)", "\\\1", leaves)
species <- colorSet[species]##c("Dm/FFS/B1H", "Mm/Uniprobe/PBM/Seed-And-Wobble")

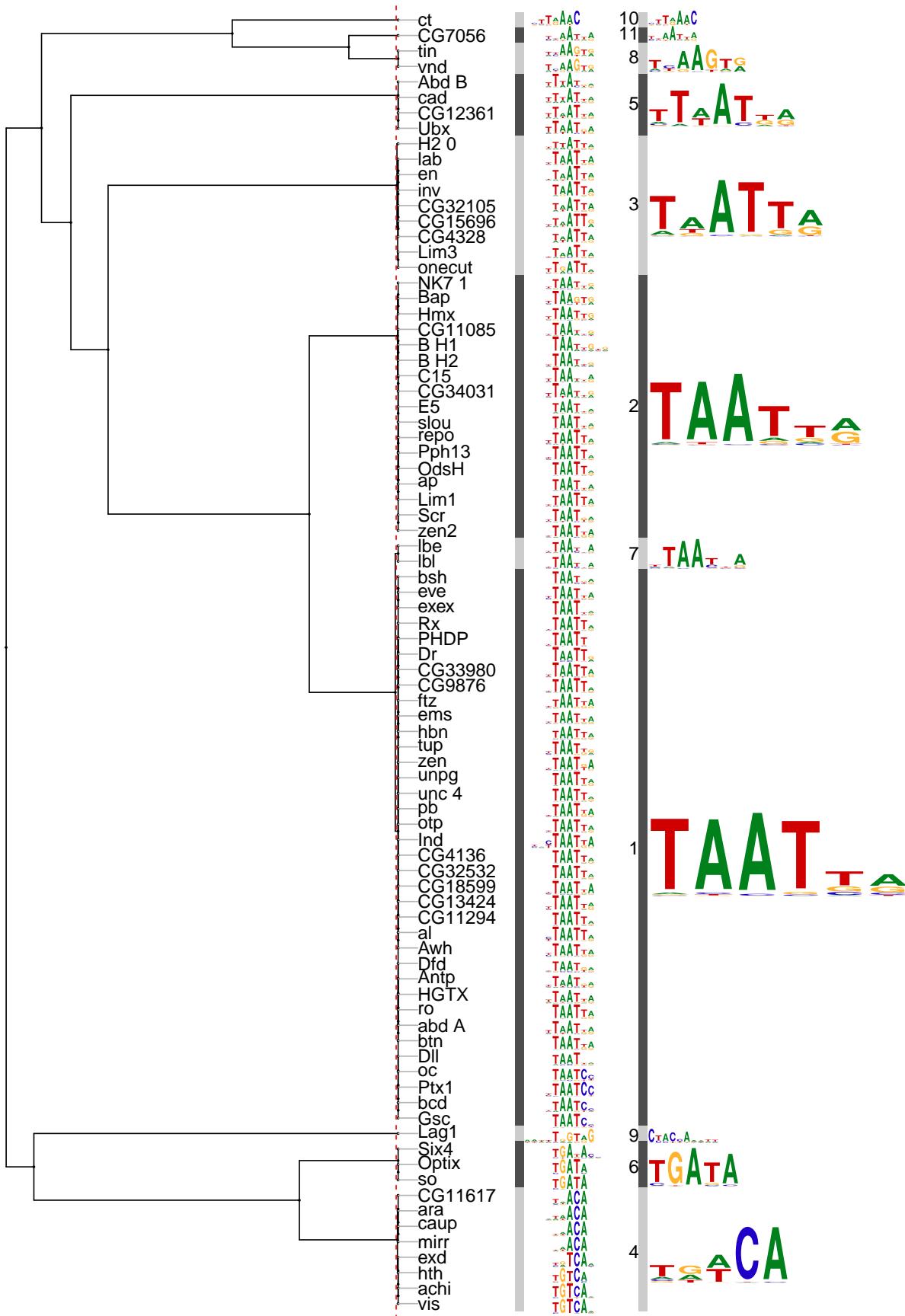
leaveNames <- gsub("(Dm|Mm)_", "", leaves)

plotMotifStackWithRadialPhylog(phylog=phylog, pfms=sig,
                               col.bg=species, col.bg.alpha=.3,
                               col.inner.label.circle=gpCol,
                               inner.label.circle.width=0.02,
                               labels.leaves=leaveNames,
                               cleaves=.1, circle=1.35, circle.motif=1.6,
                               clabel.leaves=.5, motifScale="logarithmic",
                               angle=358, plotIndex=TRUE, IndexCex=.6,
                               groupDistance=gpDis)

```



Supplementary Figure 4A: by matAlign

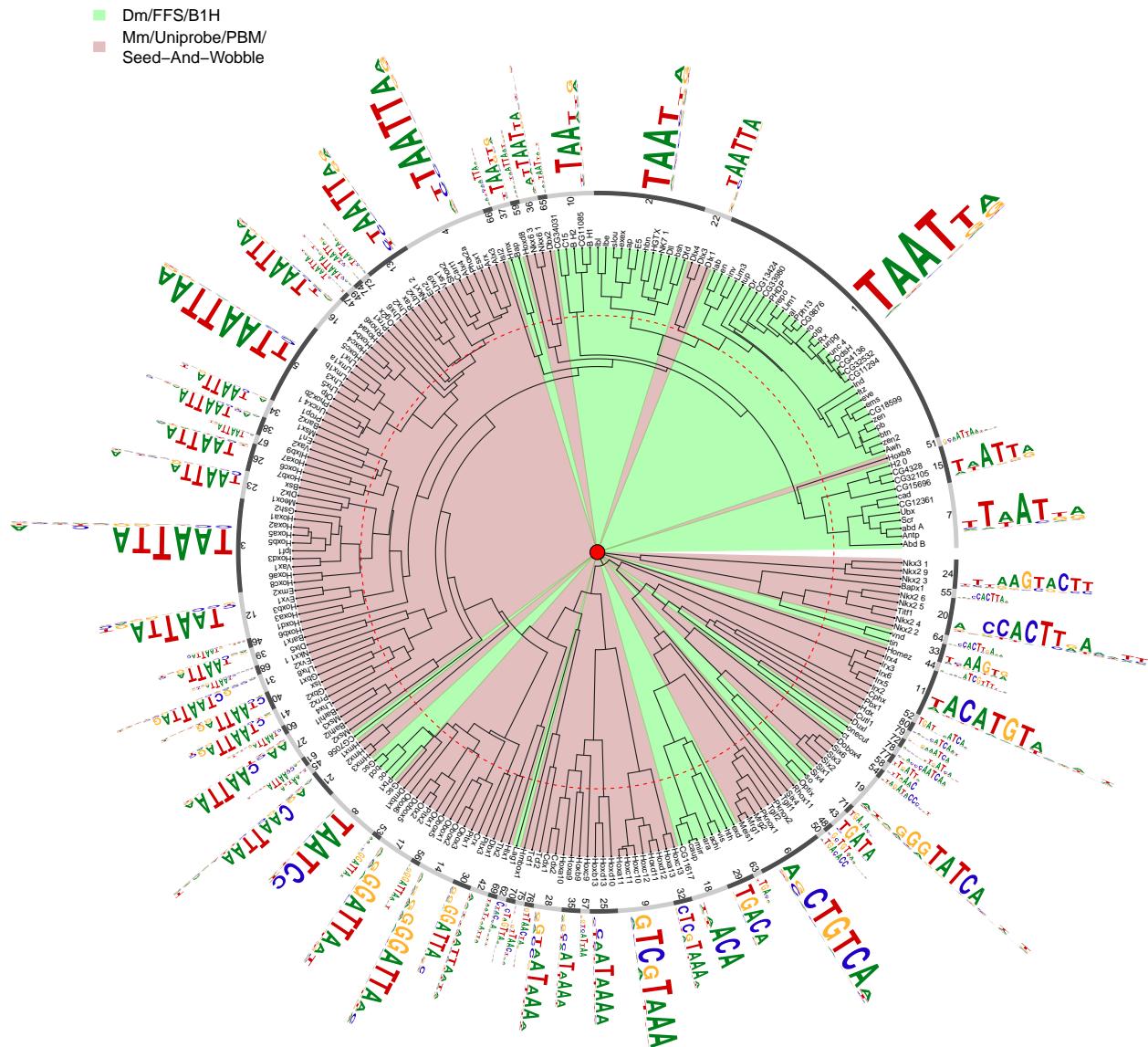


Supplementary Figure 4B: by MotIV

```

legend(-2.2, 2.4, legend=c("Dm/FFS/B1H", "Mm/Uniprobe/PBM/\nSeed-And-Wobble"),
      fill= highlightCol(colorSet[c("Dm", "Mm")], alpha=.3),
      border="white", lty=NULL, bty = "n")

```



Supplementary Figure 4C: Fly vs Mouse

```

detach(FlyMousePBMSAW)

```

Supplementary Figure 4D. fly vs human, by matAlign

```

gpDis <- 2
FlyHuman <- getMatAlignOut("pcmsDatasetFlyHuman", groupDistance=gpDis)
attach(FlyHuman)

```

```

##set methods color
species <- gsub("^(Dm|Hs)_.*$", "\\\1", leaves)
species <- colorSet[species] ##c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX")

leaveNames <- gsub("^(Dm|Hs)_", "", leaves)

plotMotifStackWithRadialPhylog(phylog=phylog, pfms=sig, col.bg=species,
                               col.bg.alpha=.3, col.inner.label.circle=gpCol,
                               inner.label.circle.width=0.04,
                               labels.leaves=leaveNames,
                               cleaves=.1, circle=1.35, circle.motif=1.8,
                               clabel.leaves=.5, motifScale="logarithmic",
                               angle=358, plotIndex=TRUE, IndexCex=.6,
                               groupDistance=gpDis)

legend(-2, 2.2, legend=c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX"),
       fill=highlightCol(colorSet[c("Dm", "Hs")], alpha=.3),
       border="white", lty=NULL, bty = "n")

detach(FlyHuman)

```

Supplementary Figure 4E. mouse selex vs mouse pbm/sw vs mouse pbm dream5

```

gpDis <- 2
MmPBM_SELEX <- getMatAlignOut("pcmsDatasetMmPBM_SELEX", groupDistance=gpDis)
attach(MmPBM_SELEX)

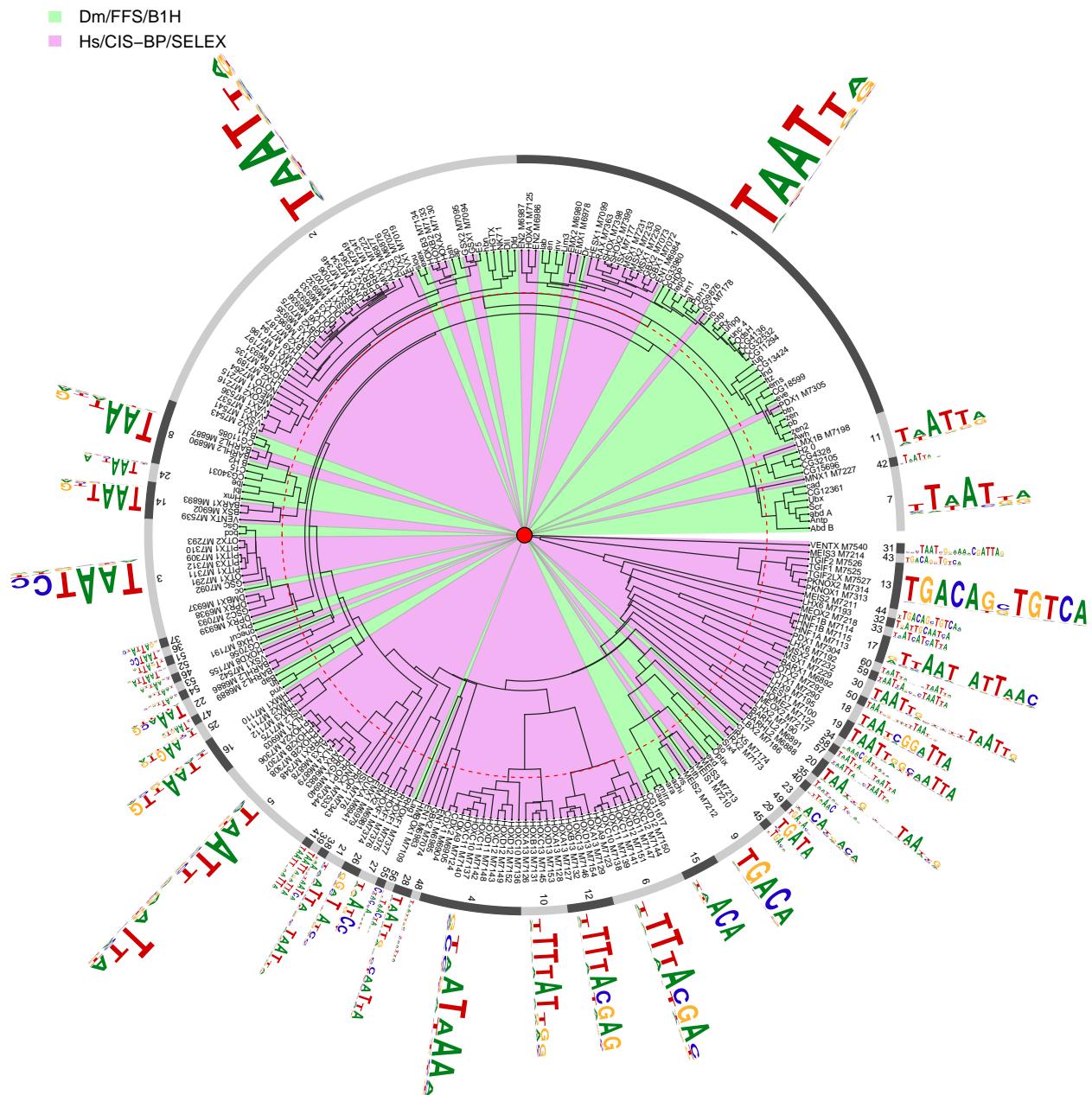
##set methods color
methods <- factor(grep("_SELEX", leaves) + grep("_M\d{4}", leaves))
levels(methods) <- c("Mm", "MmDREAM", "Ms")
##c("PBM/Uniprobe/Seed-And-Wobble", "PBM/CIS-BP/DREAM5", "SELEX/CIS-BP")
levels(methods) <- colorSet[levels(methods)]
methods <- as.character(methods)

leaveNames <- gsub("_SELEX|PBM", "", gsub("_Jolma|Berger", "", leaves))

## calculate average of top 8 position information content for each motif
icgp <- sapply(sapply(unaligned.pfms, getIC), function(.ele)
  mean(sort(.ele, decreasing=TRUE)[1:min(8, length(.ele))]))
icgp.ranges <- range(icgp)
icgp <- cut(icgp, 10, labels=colorRampPalette(c("green", "black", "red"))(10))
icgp.image <- as.raster(matrix(rev(levels(icgp)), ncol=1))
icgp <- as.character(icgp)

plotMotifStackWithRadialPhylog(phylog=phylog, pfms=sig,
                               col.bg=methods, col.bg.alpha=.3,
                               col.inner.label.circle=gpCol,
                               inner.label.circle.width=0.02,
                               labels.leaves=leaveNames, cleaves=.1,
                               circle=.95, circle.motif=1.8, clabel.leaves=.8,

```

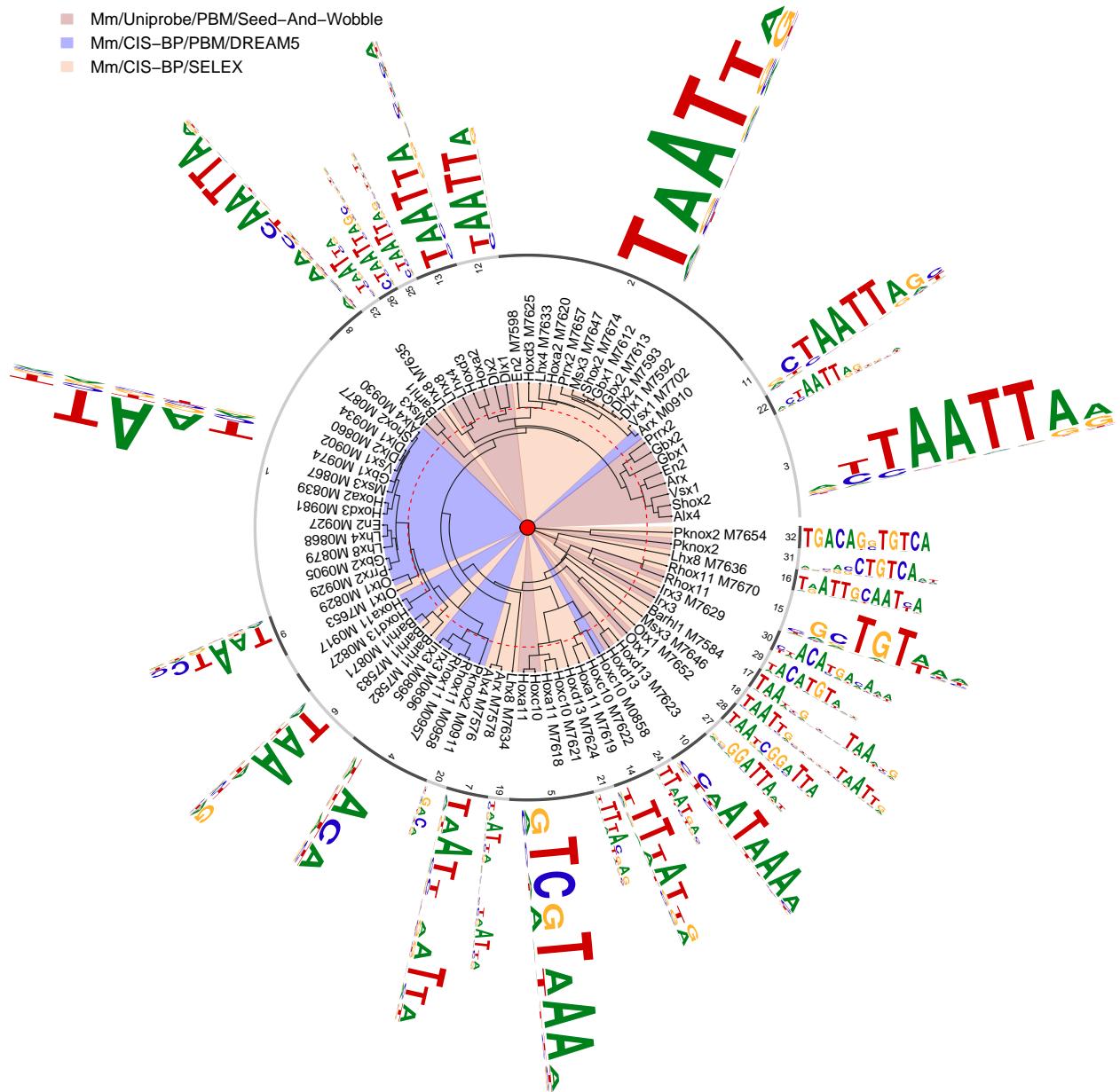


```

motifScale="logarithmic", angle=358,
plotIndex=TRUE, IndexCex=.6, groupDistance=gpDis)

legend(-2, 2.2, legend=c("Mm/Uniprobe/PBM/Seed-And-Wobble",
                        "Mm/CIS-BP/PBM/DREAM5", "Mm/CIS-BP/SELEX"),
       fill=highlightCol(colorSet[c("Mm", "MmDREAM", "Ms")], .3),
       border="white", lty=NULL, bty = "n")

```



Supplementary Figure 4E.a: Mouse data from different methods

```
motifCircos(phylog=phylog, pfms=motifs, col.tree.bg=methods,  
            col.tree.bg.alpha=.3, r.rings=c(.1, .1),  
            col.rings=list(icgp, rep("white", length(icgp))),  
            col.inner.label.circle=gpCol, inner.label.circle.width=0.02,
```

```

    labels.leaves=leaveNames, cleaves=.1, r.tree=1, r.leaves=.5,
    clabel.leaves=.9, motifScale="logarithmic", angle=358,
    plotIndex=TRUE, IndexCex=.6, groupDistance=gpDis)

legend(-2.4, 2.4, legend=c("Mm/Uniprobe/PBM/Seed-And-Wobble",
                           "Mm/CIS-BP/PBM/DREAM5", "Mm/CIS-BP/SELEX"),
       fill=highlightCol(colorSet[c("Mm", "MmDREAM", "Ms")], .3),
       border="white", lty=NULL, bty = "n")

text(2.05, 2.3, labels="Information Content")
rasterImage(icgp.image, 2, 1.8, 2.04, 2.2, interpolate = FALSE)
text(c(2.05, 2.05), c(1.8, 2.2), labels=format(icgp.ranges, digits=3), adj=0)

detach(MmPBM_SELEX)

```

Supplementary Figure 4F. mouse selex vs mouse pbm/sw vs mouse pbm dream5, by motIV

```

gpDis <- 0.005
MmPBM_SELEX.MotIV <- getMotIVOut("pcmsDatasetMmPBM_SELEX",
                                      cc="ALLR", align="SWU",
                                      groupDistance=gpDis)
attach(MmPBM_SELEX.MotIV)

##set methods color
methods <- factor(grep("_SELEX", leaves) + grep("_M\\d{4}", leaves))
levels(methods) <- c("Mm", "MmDREAM", "Ms")
##c("PBM/Uniprobe/Seed-And-Wobble", "PBM/CIS-BP/DREAM5", "SELEX/CIS-BP")
levels(methods) <- colorSet[levels(methods)]
methods <- as.character(methods)

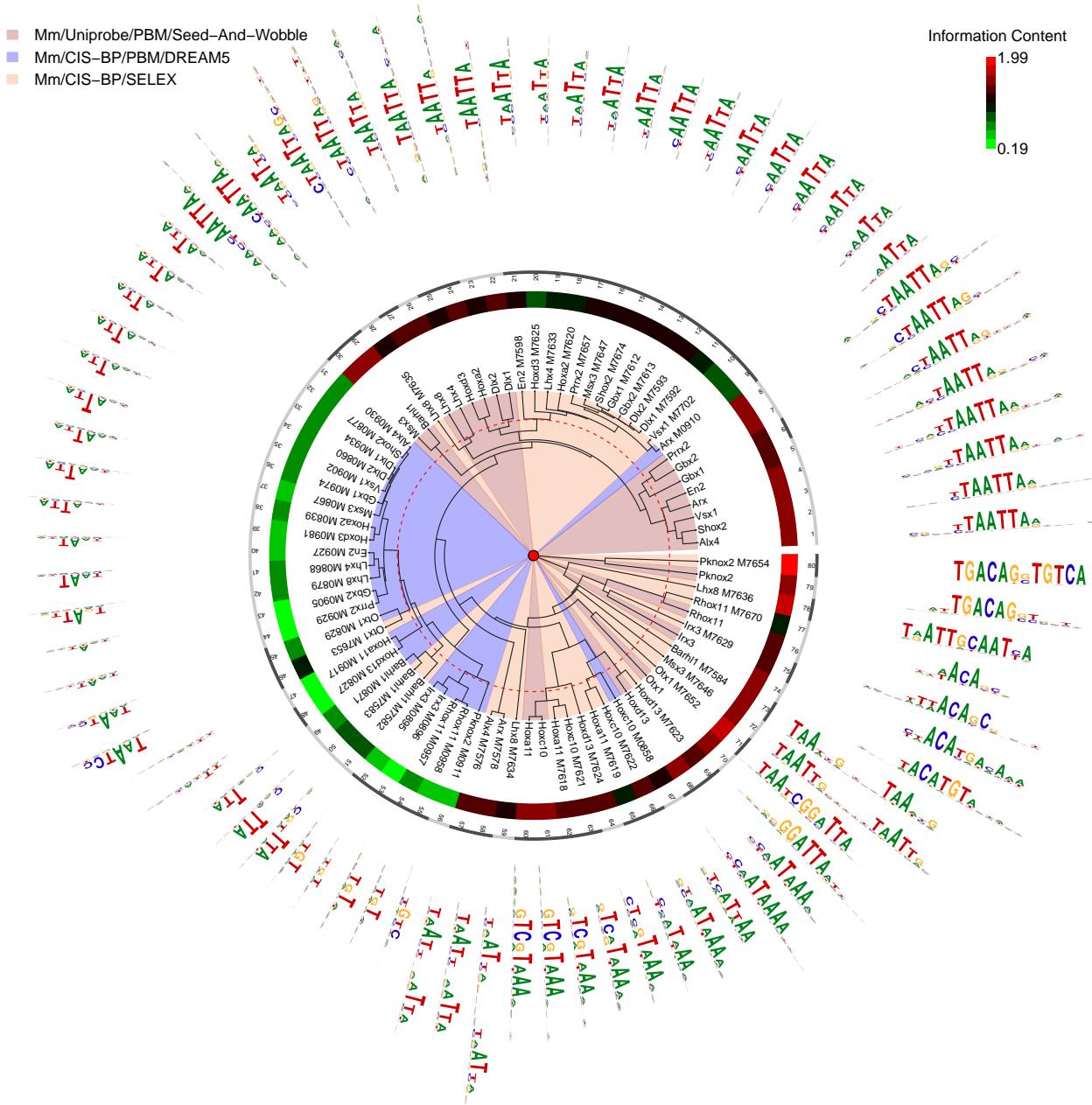
leaveNames <- gsub("_SELEX|PBM)", "", gsub("_Jolma|Berger)", "", leaves))

## calculate average of top 8 position information content for each motif
icgp <- sapply(sapply(unaligned.pfms, getIC), function(.ele)
               mean(sort(.ele, decreasing=TRUE)[1:min(8, length(.ele))]))
icgp.ranges <- range(icgp)
icgp <- cut(icgp, 10, labels=colorRampPalette(c("green", "black", "red"))(10))
icgp.image <- as.raster(matrix(rev(levels(icgp)), ncol=1))
icgp <- as.character(icgp)

plotMotifStackWithRadialPhylog(phylog=phylog, pfms=sig,
                               col.bg=methods, col.bg.alpha=.3,
                               col.inner.label.circle=gpCol,
                               inner.label.circle.width=0.02,
                               labels.leaves=leaveNames, cleaves=.1,
                               circle=.95, circle.motif=1.8, clabel.leaves=.8,
                               motifScale="logarithmic", angle=358,
                               plotIndex=TRUE, IndexCex=.6, groupDistance=gpDis)

legend(-2, 2.2, legend=c("Mm/Uniprobe/PBM/Seed-And-Wobble",

```



Supplementary Figure 4E.b: Mouse data from different methods

```

    "Mm/CIS-BP/PBM/DREAM5", "Mm/CIS-BP/SELEX"),
fill=highlightCol(colorSet[c("Mm", "MmDREAM", "Ms")], .3),
border="white", lty=NULL, bty = "n")

motifCircos(phylog=phylog, pfms=motifs, col.tree.bg=methods,
            col.tree.bg.alpha=.3, r.rings=c(.1, .1),
            col.rings=list(icgp, rep("white", length(icgp))),
            col.inner.label.circle=gpCol, inner.label.circle.width=0.02,
            labels.leaves=leaveNames, cleaves=.1, r.tree=1, r.leaves=.5,
            clabel.leaves=.9, motifScale="logarithmic", angle=358,
            plotIndex=TRUE, IndexCex=.6, groupDistance=gpDis)

legend(-2.4, 2.4, legend=c("Mm/Uniprobe/PBM/Seed-And-Wobble",
                           "Mm/CIS-BP/PBM/DREAM5", "Mm/CIS-BP/SELEX"),
       fill=highlightCol(colorSet[c("Mm", "MmDREAM", "Ms")], .3),
       border="white", lty=NULL, bty = "n")

text(2.05, 2.3, labels="Information Content")
rasterImage(icgp.image, 2, 1.8, 2.04, 2.2, interpolate = FALSE)
text(c(2.05, 2.05), c(1.8, 2.2), labels=format(icgp.ranges, digits=3), adj=0)

detach(MmPBM_SELEX.MotIV)

```

Supplementary Figure 4G. Engrailed orthologs, by matAlign

```

gpDis <- 2
EN <- getMatAlignOut("pcmsDatasetEn", groupDistance=gpDis)
attach(EN)

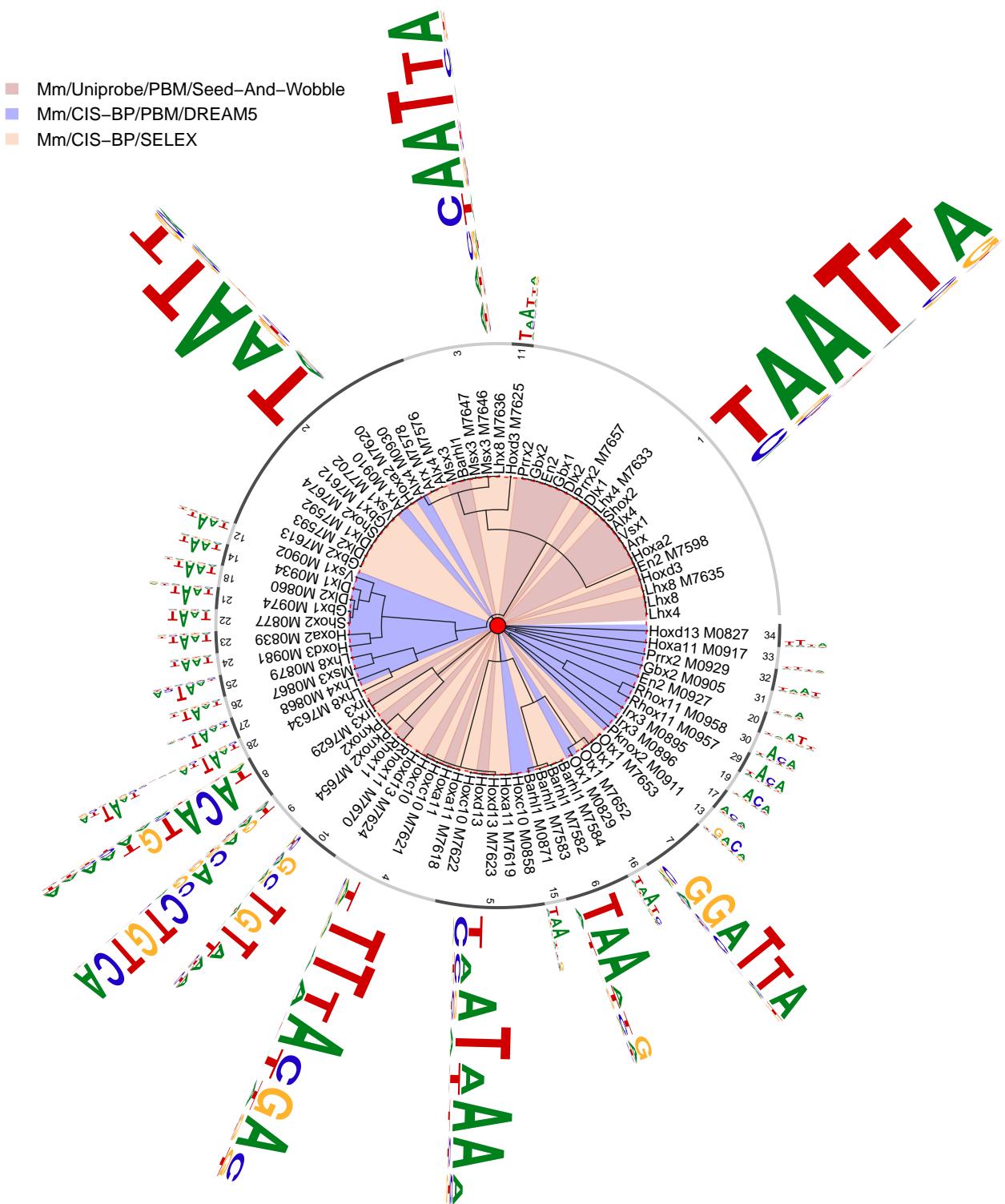
##set methods color
species <- gsub("(Dm|Hs|Mm).*?$", "\\\1", leaves)
species[grepl("_M\\d+", leaves) & species=="Mm"] <- "MmDREAM"
species <- colorSet[species] ##c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX", "Mm/Uniprobe/PBM/\nSeed-And-Wobble")

leaveNames <- gsub("(Dm|Mm|Hs)_", "", leaves)

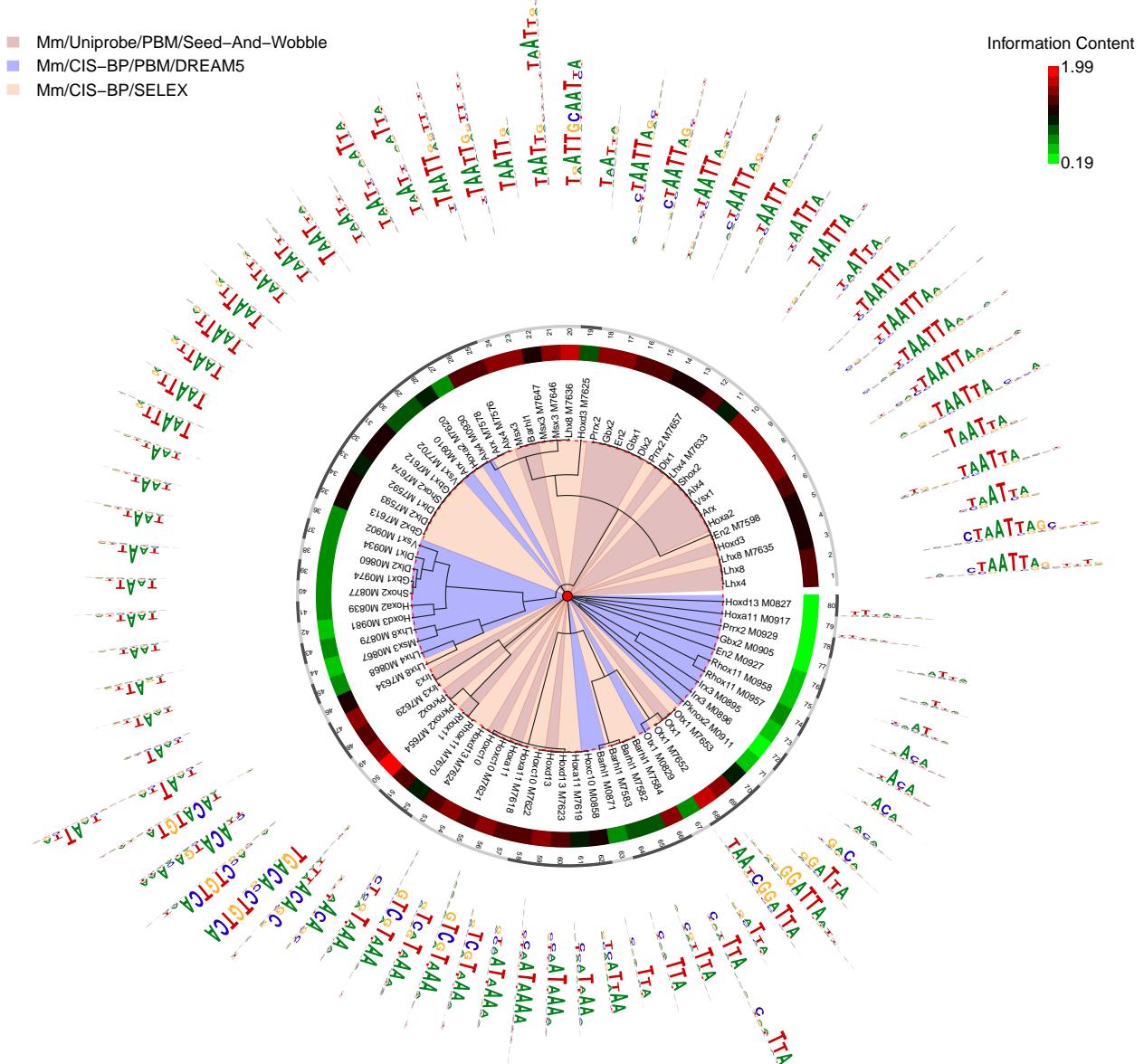
motifPiles(phylog=phylog, motifs, r.tree=.5, sig,
           col.tree=species, col.leaves=species,
           col.pfms2=gpCol, col.pfms2.width=.01,
           labels.leaves=leaveNames, plotIndex=TRUE,
           IndexCex=2, groupDistance=gpDis, clabel.leaves=2)
legend("topleft", legend=c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX",
                           "Mm/Uniprobe/PBM/\nSeed-And-Wobble",
                           "Mm/CIS-BP/PBM/Dream5"),
      fill=colorSet[c("Dm", "Hs", "Mm", "MmDREAM")],
      border="white", lty=NULL, bty = "n", cex=.5)

detach(EN)

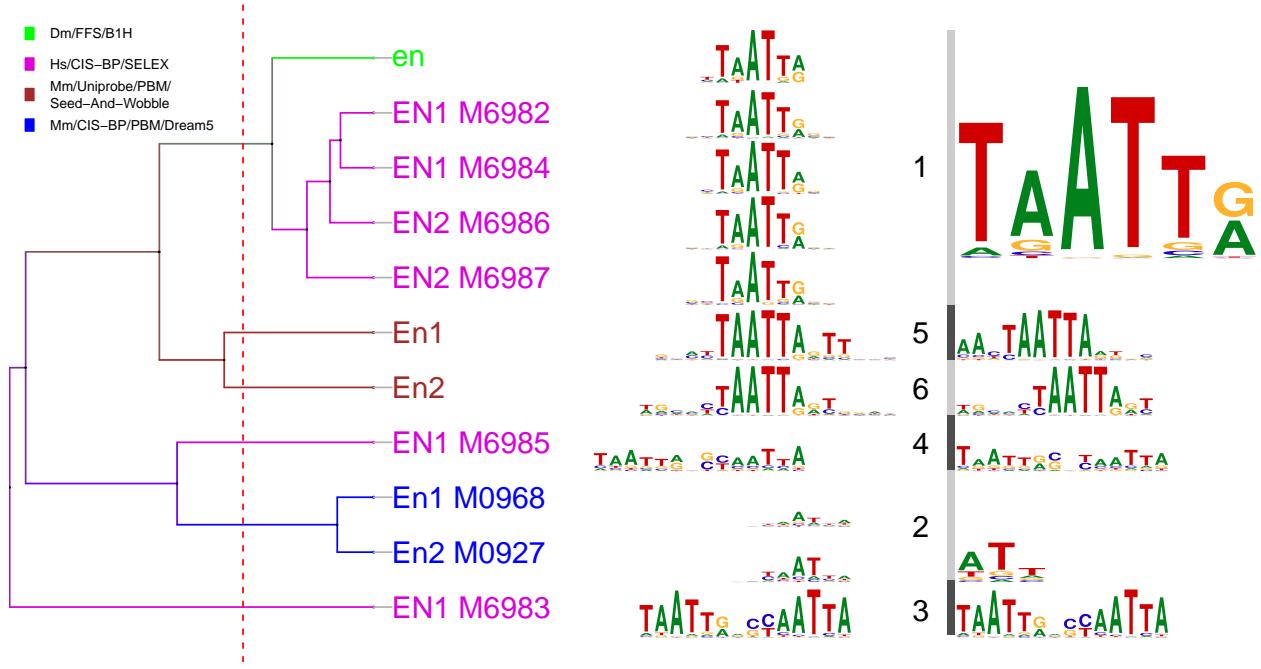
```



Supplementary Figure 4F.a: by MotIV



Supplementary Figure 4F.b: Mouse data from different methods



Supplementary Figure 4G: En

Supplementary Figure 4H. Engrailed orthologs, by motIV

```

gpDis <- 0.005
EN.MotIV <- getMotIVOut("pcmsDatasetEn",
                           cc="ALLR", align="SWU",
                           groupDistance=gpDis)
attach(EN.MotIV)

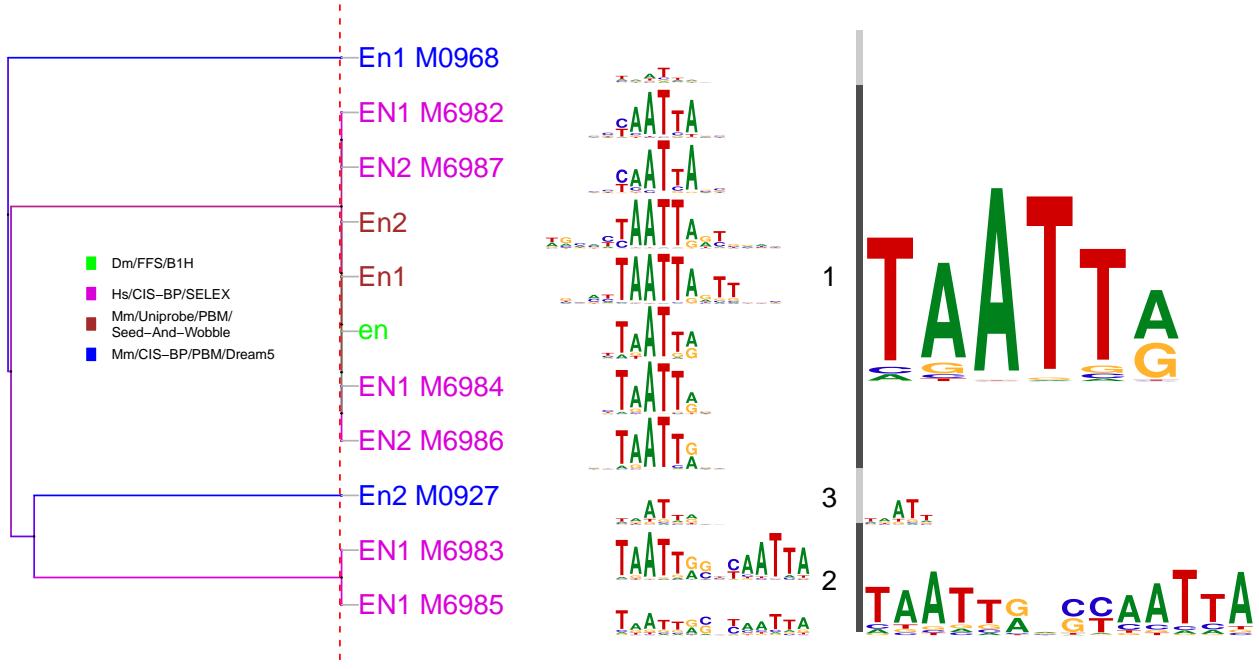
##set methods color
species <- gsub("(Dm|Hs|Mm).*?$", "\\\1", leaves)
species[grep("M\d+", leaves) & species=="Mm"] <- "MmDREAM"
##c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX", "Mm/Uniprobe/PBM/\nSeed-And-Wobble")
species <- colorSet[species]

leaveNames <- gsub("(Dm|Mm|Hs)_", "", leaves)

motifPiles(phylog=phylog, motifs, r.tree=.5, sig,
           col.tree=species, col.leaves=species,
           col.pfms2=gpCol, col.pfms2.width=.01,
           labels.leaves=leaveNames, plotIndex=TRUE,
           IndexCex=2, groupDistance=gpDis, clabel.leaves=2)
legend(0.1, 0.65, legend=c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX",
                           "Mm/Uniprobe/PBM/\nSeed-And-Wobble",
                           "Mm/CIS-BP/PBM/Dream5"),
      fill=colorSet[c("Dm", "Hs", "Mm", "MmDREAM")],
      border="white", lty=NULL, bty = "n", cex=.5)

detach(EN.MotIV)

```



Supplementary Figure 4H: En.MotIV

Supplementary Figure 4I. fly vs ms vs hum matAlign

```

gpDis <- 2.5
FlyMousePBMHuman <- getMatAlignOut("pcmsDatasetFlyMousePBMHuman",
                                      groupDistance=gpDis)
attach(FlyMousePBMHuman)

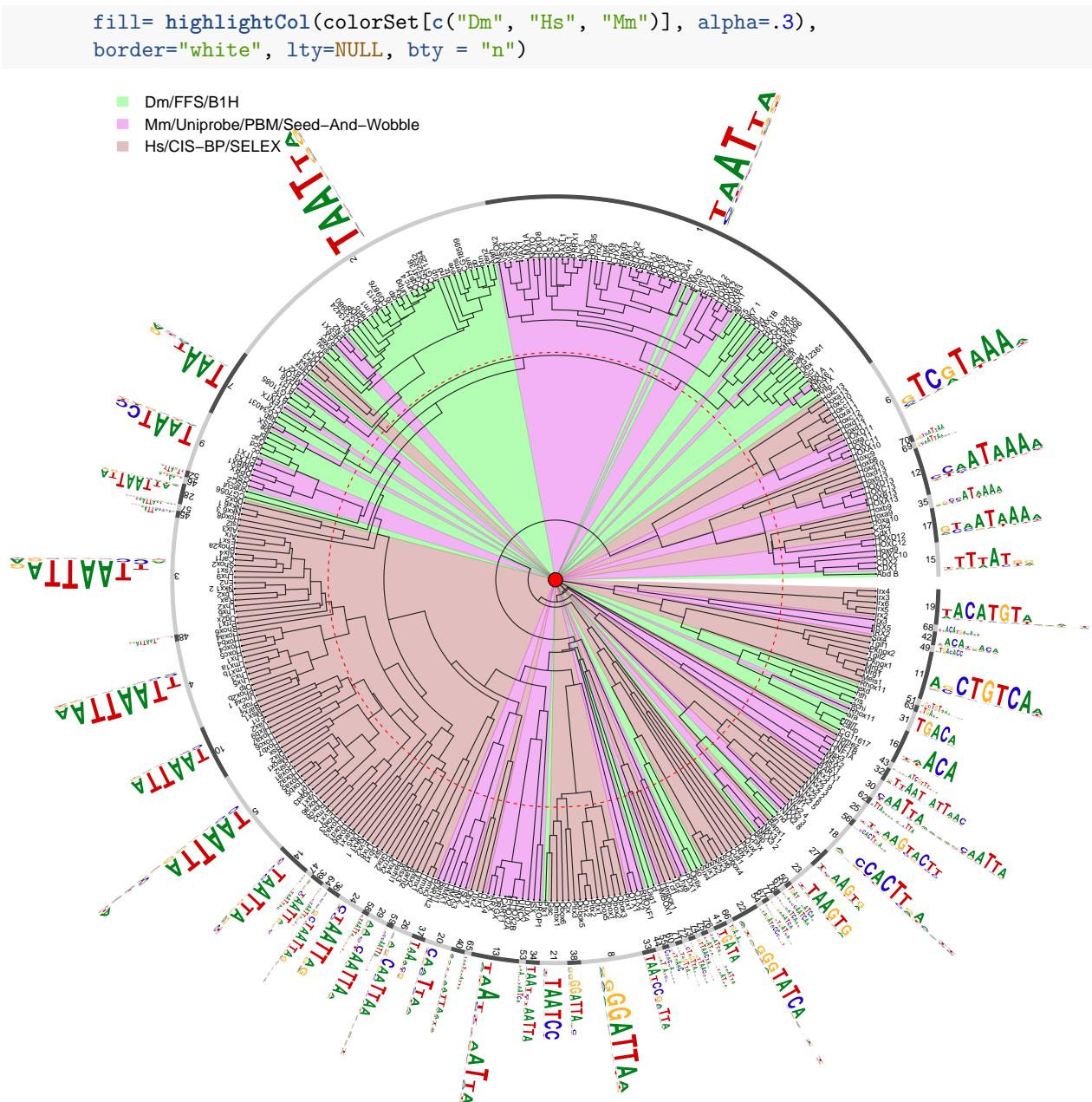
##set methods color
species <- factor(gsub("(Dm|Mm|Hs).*$", "\\\1", leaves))
##c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX", "Mm/Uniprobe/PBM/Seed-And-Wobble")
levels(species) <- colorSet[c("Dm", "Hs", "Mm")]
species <- as.character(species)

leaveNames <- gsub("(Dm|Mm|Hs)_", "", leaves)

plotMotifStackWithRadialPhylog(phylog=phylog, pfms=sig,
                                col.bg=species, col.bg.alpha=.3,
                                col.inner.label.circle=gpCol,
                                inner.label.circle.width=0.02,
                                labels.leaves=leaveNames,
                                cleaves=.1, circle=1.5, circle.motif=1.8,
                                clabel.leaves=.5, motifScale="logarithmic",
                                angle=358, plotIndex=TRUE, IndexCex=.6,
                                groupDistance=gpDis)

legend(-2, 2.2,
       legend=c("Dm/FFS/B1H",
               "Mm/Uniprobe/PBM/Seed-And-Wobble",
               "Hs/CIS-BP/SELEX"),

```



Supplementary Figure 4I: Fly vs Mouse vs Human

```
detach(FlyMousePBMHuman)
```

Supplementary Figure 4J. fly vs ms vs hum motIV

```

gpDis <- 0.005
FlyMousePBMHuman.MotIV <- getMotIVOut("pcmsDatasetFlyMousePBMHuman",
                                         cc="ALLR", align="SWU",
                                         groupDistance=gpDis)
attach(FlyMousePBMHuman.MotIV)

```

```

##set methods color
species <- factor(gsub("(Dm|Mm|Hs).*$", "\\\1", leaves))
##c("Dm/FFS/B1H", "Hs/CIS-BP/SELEX", "Mm/Uniprobe/PBM/Seed-And-Wobble")
levels(species) <- colorSet[c("Dm", "Hs", "Mm")]
species <- as.character(species)

leaveNames <- gsub("(Dm|Mm|Hs)_", "", leaves)

plotMotifStackWithRadialPhylog(phylog=phylog, pfms=sig,
                                col.bg=species, col.bg.alpha=.3,
                                col.inner.label.circle=gpCol,
                                inner.label.circle.width=0.02,
                                labels.leaves=leaveNames,
                                cleaves=.1, circle=1.5, circle.motif=1.8,
                                clabel.leaves=.5, motifScale="logarithmic",
                                angle=358, plotIndex=TRUE, IndexCex=.6,
                                groupDistance=gpDis)

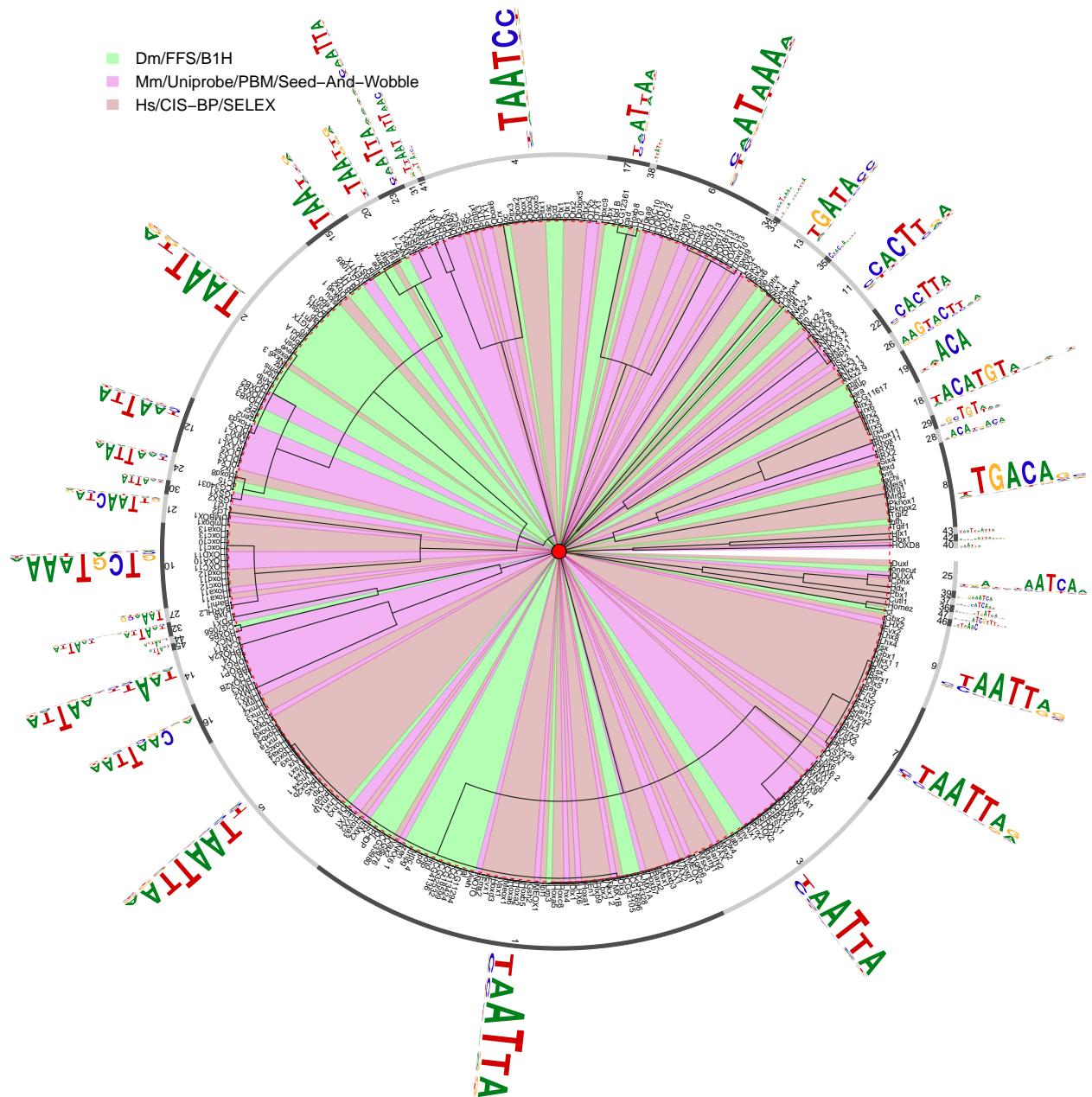
legend(-2, 2.2,
       legend=c("Dm/FFS/B1H",
               "Mm/Uniprobe/PBM/Seed-And-Wobble",
               "Hs/CIS-BP/SELEX"),
       fill= highlightCol(colorSet[c("Dm", "Hs", "Mm")], alpha=.3),
       border="white", lty=NULL, bty = "n")

detach(FlyMousePBMHuman.MotIV)

sessionInfo()

## R version 3.4.1 (2017-06-30)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: OS X El Capitan 10.11.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats4      parallel    grid        stats       graphics   grDevices utils
## [8] datasets    methods     base
##
## other attached packages:
## [1] motifStack_1.20.1   Biostrings_2.44.2   XVector_0.16.0
## [4] IRanges_2.10.2     S4Vectors_0.14.3   ade4_1.7-6
## [7] MotIV_1.32.0      BiocGenerics_0.22.0 grImport_0.9-0
## [10] XML_3.98-1.9
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.12.12          highr_0.6
## [3] plyr_1.8.4            compiler_3.4.1
## [5] GenomeInfoDb_1.12.2   bitops_1.0-6

```



Supplementary Figure 4J: Fly vs Mouse vs Human by MotIV

```
## [7] tools_3.4.1                  zlibbioc_1.22.0
## [9] digest_0.6.12                 evaluate_0.10.1
## [11] lattice_0.20-35              BSgenome_1.44.0
## [13] Matrix_1.2-10                DelayedArray_0.2.7
## [15] yaml_2.1.14                  seqLogo_1.42.0
## [17] GenomeInfoDbData_0.99.0     rtracklayer_1.36.4
## [19] stringr_1.2.0                knitr_1.16
## [21] htmlwidgets_0.9               rprojroot_1.2
## [23] Biobase_2.36.2              BiocParallel_1.10.1
## [25] rGADEM_2.24.0               rmarkdown_1.6
## [27] magrittr_1.5                 scales_0.4.1
## [29] backports_1.1.0              Rsamtools_1.28.0
## [31] htmltools_0.3.6              matrixStats_0.52.2
## [33] GenomicRanges_1.28.4         GenomicAlignments_1.12.1
## [35] SummarizedExperiment_1.6.3   colorspace_1.3-2
## [37] stringi_1.1.5               munsell_0.4.3
## [39] RCurl_1.95-4.8
```