CIF_Analysis_TryTrypDB

David H. Ardell

6/11/2019

This is the CIF Analysis Workflow for Kelly et al. (2019) using the old software for this (logofun and logotax by Eva Freyhult).

Function Logos

```
logofun --states ACGT -f 1 -b 1 -o HOMO HOMO_/clustalW/HOMO_*.aln
logofun --states ACGT -f 1 -b 1 -o Leishmania1 Leishmania1_*.aln
logofun --states ACGT -f 1 -b 1 -o Leishmania2 Leishmania2_*.aln
logofun --states ACGT -f 1 -b 1 -o Leishmania3 Leishmania3_*.aln
logofun --states ACGT -f 1 -b 1 -o Leishmania4 Leishmania4_*.aln
logofun --states ACGT -f 1 -b 1 -o Leishmania5 Leishmania5_*.aln
logofun --states ACGT -f 1 -b 1 -o AfricanTrypanosome AfricanTrypanosome_*.aln
logofun --states ACGT -f 1 -b 1 -o AmericanTrypanosome AmericanTrypanosome_*.aln
logofun --states ACGT -f 1 -b 1 -o LenriettiComplex LenriettiComplex_*.aln
```

ID Logos

```
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_Leishmania1_ID 'HOMO_*.aln' 'Leishmania1_*.aln
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_Leishmania2_ID 'HOMO_*.aln' 'Leishmania2_*.aln
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_Leishmania3_ID 'HOMO_*.aln' 'Leishmania3_*.aln
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_Leishmania4_ID 'HOMO_*.aln' 'Leishmania4_*.aln
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_Leishmania5_ID 'HOMO_*.aln' 'Leishmania5_*.aln
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_AfricanTrypanosome_ID 'HOMO_*.aln' 'AfricanTry
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_AmericanTrypanosome_ID 'HOMO_*.aln' 'American'
logotax -b 3 -l infodiff -p other --states 'ACGT' -o HOMO_LenriettiComplex_ID 'HOMO_*.aln' 'LenriettiCo
```

KLD Logos

```
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_Leishmania1_KLD 'HOMO_*.aln' 'Leishmania1_*.aln
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_Leishmania2_KLD 'HOMO_*.aln' 'Leishmania2_*.aln
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_Leishmania3_KLD 'HOMO_*.aln' 'Leishmania3_*.aln
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_Leishmania4_KLD 'HOMO_*.aln' 'Leishmania4_*.aln
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_Leishmania5_KLD 'HOMO_*.aln' 'Leishmania5_*.aln
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_AfricanTrypanosome_KLD 'HOMO_*.aln' 'AfricanTry
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_AmericanTrypanosome_KLD 'HOMO_*.aln' 'American'
logotax -b 3 -l KLdiff2 -p other --states 'ACGT' -o HOMO_LenriettiComplex_KLD 'HOMO_*.aln' 'LenriettiCom
```

Convert to Tables

```
./ID_KLD_eps2table Leishmania1 HOMO_Leishmania1_ID_Leishmania1 HOMO_Leishmania1_ID_HOMO HOMO_Leishmania1
./ID_KLD_eps2table Leishmania2 HOMO_Leishmania2_ID_Leishmania2 HOMO_Leishmania2_ID_HOMO HOMO_Leishmania2
./ID_KLD_eps2table Leishmania3 HOMO_Leishmania3_ID_Leishmania3 HOMO_Leishmania3_ID_HOMO HOMO_Leishmania3
./ID_KLD_eps2table Leishmania4 HOMO_Leishmania4_ID_Leishmania4 HOMO_Leishmania4_ID_HOMO HOMO_Leishmania4
./ID_KLD_eps2table Leishmania5 HOMO_Leishmania5_ID_Leishmania5 HOMO_Leishmania5_ID_HOMO HOMO_Leishmania5
./ID_KLD_eps2table AfricanTrypanosome HOMO_AfricanTrypanosome_ID_AfricanTrypanosome HOMO_AfricanTrypanos
./ID_KLD_eps2table AmericanTrypanosome HOMO_AmericanTrypanosome_ID_AmericanTrypanosome HOMO_AmericanTryp
./ID_KLD_eps2table LenriettiComplex HOMO_LenriettiComplex_ID_LenriettiComplex HOMO_LenriettiComplex_ID_L
```

Bubbleplots

```
Leishmania1 <- read.table("Leishmania1_table.txt",header=T);
Leishmania2 <- read.table("Leishmania2_table.txt",header=T);
Leishmania3 <- read.table("Leishmania3_table.txt",header=T);
Leishmania4 <- read.table("Leishmania4_table.txt",header=T);
Leishmania5 <- read.table("Leishmania5_table.txt",header=T);
AfricanTrypanosome <- read.table("AfricanTrypanosome_table.txt",header=T);
AmericanTrypanosome <- read.table("AmericanTrypanosome_table.txt",header=T);
LenriettiComplex <- read.table("LenriettiComplex_table.txt",header=T);
library("heR.Misc"); ## THIS IS REQUIRED FOR THE BUBBLEPLOT FUNCTION AND MUST BE DOWNLOADED FROM
```

```
## Loading required package: lattice


##
## Attaching package: 'heR.Misc'

## The following object is masked from 'package:lattice':
##
##     panel.superpose.2
```

```
## http://exposurescience.org/her.html

# THESE ARE THE COORDINATES FOR THE TRNA STRUCTURE BACKBONE IN THE FIGURES

# THESE ARE THE COORDINATES FOR THE TRNA STRUCTURE BACKBONE IN THE FIGURES

line.x <- c(6.875,6.500,6.125,5.750,5.375,5.000,4.625,4.625,5.000,5.000,2.375
         ,2.750,2.375,2.750,2.500,2.875,3.250,2.875,2.500,2.125,1.750,1.375,1.000
         ,0.625,0.250,0.625,1.000,1.500,1.125,1.500,1.125,1.500,1.125,1.500,1.125
         ,1.500,1.125,0.625,1.000,1.375,1.750,2.125,2.500,2.875,2.375,2.750,2.375
         ,2.750,2.375,4.250,4.250,3.875,4.250,3.875,4.250,4.250,3.875,3.500,3.125
         ,2.750,2.250,1.875,1.500,1.125,1.500,1.875,2.250,2.750,3.125,3.500,3.875
         ,4.250,4.625,5.000,5.375,5.750,6.125,6.500,6.875,7.250,7.625,8.000,8.375);

line.y <- c(8.875,8.500,8.875,8.500,8.875,8.500,8.875,7.375,7.000
         ,3.500,3.500,3.875,4.250,4.625,5.125,5.500,5.875,6.250,6.625
```

```
          ,7.000,7.375,7.000,6.625,6.250,5.875,5.500,5.125,4.625,4.250
          ,3.875,3.500,3.125,2.750,2.375,2.000,1.625,1.250,0.750,0.375
          ,0.000,-0.375,0.000,0.375,0.750,1.250,1.625,2.000,2.375,2.750
          ,2.750,3.875,4.250,4.625,5.000,5.375,8.500,8.875,8.500,8.875
          ,8.500,8.000,8.375,8.750,9.125,9.500,9.875,10.250,9.750,10.125
          ,9.750,10.125,9.750,10.125,9.750,10.125,9.750,10.125,9.750,10.125
          ,9.750,10.125,9.750,10.125);

## THESE ARE FOR THE SPRINZL COORD LABELS
coord.labels <- c("1","5","10","14A","18","20C","25","30","35","40","45","50","55","60","65","70");
coord.labels.x <- c(6.875,5.375,2.375,2.875,1.750,0.625,1.125,1.500,1.750,2.750,3.875,3.875,1.875,2.250
coord.labels.y <- c(8.875,8.875,3.500,5.500,7.375,5.500,3.500,1.625,-0.375,1.625,4.250,8.875,8.375,10.25

xbump <- 0.35;
ybump <- 0.4;

up <- c(5,13,14,15,16,17);
up.coord.labels <- coord.labels[up];
up.coord.labels.x <- coord.labels.x[up];
up.coord.labels.y <- coord.labels.y[up] + ybump;

dn <- c(1,2,9,11,12);
dn.coord.labels <- coord.labels[dn];
dn.coord.labels.x <- coord.labels.x[dn];
dn.coord.labels.y <- coord.labels.y[dn] - ybump;

lt <- c(3,4,7);
lt.coord.labels <- coord.labels[lt];
lt.coord.labels.x <- coord.labels.x[lt] - xbump;
lt.coord.labels.y <- coord.labels.y[lt];

rt <- c(6,8,10);
rt.coord.labels <- coord.labels[rt];
rt.coord.labels.x <- coord.labels.x[rt] + xbump;
rt.coord.labels.y <- coord.labels.y[rt];


all.bubble <- function(df,name="bubble",alpha=0.5,fact=0.5,area=TRUE,legend=FALSE) {
  gains <- (df$gainbits * df$gainfht);
  convs <- (df$convbits * df$convfht);
  map2rgb <- function (c) { rgb(t(col2rgb(c))/255,alpha=alpha);}
  colormap <- function (g,c) {
    y <- rep(0,length(g));
    y[g <  0.48            & c < 0.44]             <- rgb(t(col2rgb("white"))/255,alpha=alpha);
    y[g >= 0.48 & g < 0.95 & c < 0.44]             <- rgb(t(col2rgb("darkred"))/255,alpha=alpha);
    y[g >= 0.95            & c < 0.44]             <- rgb(t(col2rgb("red"))/255,alpha=alpha);
    y[g <  0.48            & c >= 0.44 & c < 0.70] <- rgb(t(col2rgb("darkblue"))/255,alpha=alpha);
    y[g >= 0.48 & g < 0.95 & c >= 0.44 & c < 0.70] <- rgb(t(col2rgb("darkmagenta"))/255,alpha=alpha);
    y[g >= 0.95            & c >= 0.44 & c < 0.70] <- map2rgb("deeppink");
    y[g <  0.48            & c >= 0.70]            <- map2rgb("blue");
    y[g >= 0.48 & g < 0.95 & c >= 0.70]            <- map2rgb("blueviolet");
    y[g >= 0.95            & c >= 0.70]            <- map2rgb("magenta");
    y;
```

```
}
colors <- colormap(gains,convs);

widthmap <- function (g,c) {
   y <- rep(1,length(g));
   y[g < 0.48 & c < 0.44] <- 1;
   y[g >= 0.48 | c >= 0.44] <- 2;
   y[g >= 0.95 | c >= 0.70] <- 3;
   y;
}
widths <- widthmap(gains,convs);

for (class in levels(df$aa)) {
  filenm <- paste(name,"_",class,".pdf",sep="");
  pdf(file=filenm,version="1.4");
  bubbleplot(
          df$x[df$aa == class],
          df$y[df$aa == class],
       (df$fbits[df$aa == class] * df$fht[df$aa == class]),
       fact=fact, #0.265165 = sqrt(2*0.375^2)/2
       area=area,
       fg = rgb(t(col2rgb("black")/255)),
       bg = colors[df$aa == class],
       box=FALSE,
       axes=FALSE,
       lwd=2
       );
  lines(line.x,line.y);
  text(labels=up.coord.labels,x=up.coord.labels.x,y=up.coord.labels.y);
  text(labels=dn.coord.labels,x=dn.coord.labels.x,y=dn.coord.labels.y);
  text(labels=lt.coord.labels,x=lt.coord.labels.x,y=lt.coord.labels.y);
  text(labels=rt.coord.labels,x=rt.coord.labels.x,y=rt.coord.labels.y);
  if (legend) {
    legend.x <- rep(df$x[df$aa == "X" &  df$state == "A" &
                  (df$sprinzl == "68" |  df$sprinzl == "70" |
                   df$sprinzl == "72" )] + c(0.2,0.3,0.4),3);
    legend.y <- c(rep(df$y[df$aa == "X" &  df$state == "A" & df$sprinzl == "31"],3),
                  rep(df$y[df$aa == "X" &  df$state == "A" & df$sprinzl == "27"],3),
                  rep(df$y[df$aa == "X" &  df$state == "A" & df$sprinzl == "23"],3));
    legend.z <- rep(2.2,9);
    legend.c <- colormap(c(0,0,0,0.5,0.5,0.5,1,1,1),c(0,0.5,1,0,0.5,1,0,0.5,1));
    bubbleplot(legend.x, legend.y, legend.z,fact=fact,area=area,bg=legend.c,
        add=TRUE,
        box=FALSE,axes=FALSE,lwd=2);

  }
  prime.x <- df$x[df$aa == "X" &  df$state == "A" &
              (df$sprinzl == "1" |  df$sprinzl == "75")];
  prime.y <- df$y[df$aa == "X" &  df$state == "A" &
              (df$sprinzl == "1" |  df$sprinzl == "75")];
  text(labels=c("5'","3'"),x=prime.x,y=prime.y,adj=c(-0.9,0));
  dev.off();
}
```

```
}

all.bubble(Leishmania1,name="Leishmania1");
```

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

```r
all.bubble(Leishmania2,name="Leishmania2");
```

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

Radius scale factor is 0.5

```r
all.bubble(Leishmania3,name="Leishmania3");
```

Radius scale factor is 0.5

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```r
all.bubble(Leishmania4,name="Leishmania4");
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```r
all.bubble(Leishmania5,name="Leishmania5");
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```r
all.bubble(AfricanTrypanosome,name="AfricanTrypanosome");
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```r
all.bubble(AmericanTrypanosome,name="AmericanTrypanosome");
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```r
all.bubble(LenriettiComplex,name="LenriettiComplex");
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```

```
## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5

## Radius scale factor is 0.5
```