

WholeGenomeAnalyses.NoOverlap.AGSkewAndGradient.R

popadin

2019-09-13

```
#####  
#####  
  
rm(list=ls(all=TRUE))  
  
##### Syn mut  
unzip("../Body/3Results/AllGenesCodonUsageNoOverlap.txt.zip")  
SynNuc = read.table("../Body/3Results/AllGenesCodonUsageNoOverlap.txt", header = TRUE, sep = '\t')  
if (file.exists("../Body/3Results/AllGenesCodonUsageNoOverlap.txt")){file.remove("../Body/3Results/AllGenesCodonUsageNoOverlap.txt")}  
  
## [1] TRUE  
names(SynNuc)  
  
## [1] "Species" "Gene"  
## [3] "GeneStart" "GeneEnd"  
## [5] "CodonsToDeleteInTheBeginning" "CodonsToDeleteAtTheEnd"  
## [7] "AminoNoOverlap" "CodonsNoOverlap"  
## [9] "AAA" "AAC"  
## [11] "AAG" "AAT"  
## [13] "ACA" "ACC"  
## [15] "ACG" "ACT"  
## [17] "AGA" "AGC"  
## [19] "AGG" "AGT"  
## [21] "ATA" "ATC"  
## [23] "ATG" "ATT"  
## [25] "CAA" "CAC"  
## [27] "CAG" "CAT"  
## [29] "CCA" "CCC"  
## [31] "CCG" "CCT"  
## [33] "CGA" "CGC"  
## [35] "CGG" "CGT"  
## [37] "CTA" "CTC"  
## [39] "CTG" "CTT"  
## [41] "GAA" "GAC"  
## [43] "GAG" "GAT"  
## [45] "GCA" "GCC"  
## [47] "GCG" "GCT"  
## [49] "GGA" "GGC"  
## [51] "GGG" "GGT"  
## [53] "GTA" "GTC"  
## [55] "GTG" "GTT"  
## [57] "TAA" "TAC"  
## [59] "TAG" "TAT"  
## [61] "TCA" "TCC"  
## [63] "TCG" "TCT"  
## [65] "TGA" "TGC"  
## [67] "TGG" "TGT"
```

```
## [69] "TTA"          "TTC"
## [71] "TTG"          "TTT"
## [73] "NeutralA"     "NeutralT"
## [75] "NeutralG"     "NeutralC"
## [77] "Taxonomy"     "Class"
```

```
### make ND6 complementary:
```

```
NotND6 = SynNuc[SynNuc$Gene != 'ND6',]
ND6 = SynNuc[SynNuc$Gene == 'ND6',]
A = ND6$NeutralT
T = ND6$NeutralA
G = ND6$NeutralC
C = ND6$NeutralG
ND6$NeutralA = A
ND6$NeutralT = T
ND6$NeutralG = G
ND6$NeutralC = C
SynNuc = rbind(NotND6,ND6)
```

```
VecOfTaxa = unique(SynNuc$Class)
table(SynNuc$Class)/13
```

```
##
## Actinopterygii      Amphibia      AncientFish      Aves      Mammalia
##           1770           205           126           432           788
##           Reptilia
##           269
```

```
##### GENOME WIDE SKEW FOR EACH SPECIES
```

```
AGG = aggregate(list(SynNuc$NeutralA,SynNuc$NeutralT,SynNuc$NeutralG,SynNuc$NeutralC), by = list(SynNuc$Species,
names(AGG) = c('Species','Class','NeutralA','NeutralT','NeutralG','NeutralC'))
```

```
## all six different skews
```

```
AGG$CTSkew = (AGG$NeutralC - AGG$NeutralT)/(AGG$NeutralC + AGG$NeutralT); summary(AGG$CTSkew) # GA on h
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.8802  0.1313  0.2506  0.2362  0.3546  0.7942
```

```
AGG$CGSkew = (AGG$NeutralC - AGG$NeutralG)/(AGG$NeutralC + AGG$NeutralG); summary(AGG$CGSkew) #
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.7073  0.6126  0.6987  0.6806  0.7674  0.9429
```

```
AGG$CASKew = (AGG$NeutralC - AGG$NeutralA)/(AGG$NeutralC + AGG$NeutralA); summary(AGG$CASKew) #
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.780731 -0.244339 -0.125410 -0.124205  0.001386  0.345272
```

```
AGG$TGSkew = (AGG$NeutralT - AGG$NeutralG)/(AGG$NeutralT + AGG$NeutralG); summary(AGG$TGSkew) #
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.1010  0.4111  0.5385  0.5237  0.6497  0.9429
```

```
AGG$TASkew = (AGG$NeutralT - AGG$NeutralA)/(AGG$NeutralT + AGG$NeutralA); summary(AGG$TASkew) #
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.7908 -0.4785 -0.3615 -0.3522 -0.2386  0.3338
```

```

AGG$GASkeW = (AGG$NeutralG - AGG$NeutralA)/(AGG$NeutralG + AGG$NeutralA); summary(AGG$GASkeW) #

##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.97260 -0.82974 -0.75867 -0.73567 -0.66395  0.04819

AGG$TCSkeW = (AGG$NeutralT - AGG$NeutralC)/(AGG$NeutralT + AGG$NeutralC); summary(AGG$TCSkeW) # AG on h

##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.8802  0.1313  0.2506  0.2362  0.3546  0.7942

GT = read.table("../Body/1Raw/GenerationLengthforMammals.xlsx.txt", header = TRUE, sep = '\t')
GT$Species = gsub(' ','_',GT$Scientific_name)
length(unique(GT$Species))

## [1] 5424

summary(GT$AdultBodyMass_g)      # 2      21      71      136058      614 154321304

##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
##      2      21      71      136058      614 154321304

summary(GT$GenerationLength_d) # 129.0  624.4 1101.1 1578.8 2064.3 18980.0 # max = 18980 days => 5.

##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
## 129.0  624.4 1101.1 1578.8 2064.3 18980.0

GT = GT[,c(11,13)]
summary(GT$GenerationLength_d)

##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
## 129.0  624.4 1101.1 1578.8 2064.3 18980.0

Mam = merge(AGG,GT, by = 'Species')

##### generation length versus all six skews: 6 rank correlations and one multiple model:
##### pairwise rank corr:
cor.test(log2(Mam$GenerationLength_d),Mam$CTSkew, method = 'spearman') # rho = 0.4117201; p-value < 2..

## Warning in cor.test.default(log2(Mam$GenerationLength_d), Mam$CTSkew,
## method = "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data: log2(Mam$GenerationLength_d) and Mam$CTSkew
## S = 26926000, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.4117201

cor.test(log2(Mam$GenerationLength_d),Mam$CGSkew, method = 'spearman') # rho = 0.05774663, p = 0.1414

## Warning in cor.test.default(log2(Mam$GenerationLength_d), Mam$CGSkew,
## method = "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data: log2(Mam$GenerationLength_d) and Mam$CGSkew

```

```

## S = 43128000, p-value = 0.1414
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.05774663

cor.test(log2(Mam$GenerationLength_d),Mam$CASKew, method = 'spearman') # rho = 0.4691996, p-value < 2

## Warning in cor.test.default(log2(Mam$GenerationLength_d), Mam$CASKew,
## method = "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data: log2(Mam$GenerationLength_d) and Mam$CASKew
## S = 24295000, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.4691996

cor.test(log2(Mam$GenerationLength_d),Mam$TGSkew, method = 'spearman') # rho = -0.24, p-value = 5.183

## Warning in cor.test.default(log2(Mam$GenerationLength_d), Mam$TGSkew,
## method = "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data: log2(Mam$GenerationLength_d) and Mam$TGSkew
## S = 56783000, p-value = 5.183e-10
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## -0.2405869

cor.test(log2(Mam$GenerationLength_d),Mam$TASkew, method = 'spearman') # rho = -0.1183936, p-value =

## Warning in cor.test.default(log2(Mam$GenerationLength_d), Mam$TASkew,
## method = "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data: log2(Mam$GenerationLength_d) and Mam$TASkew
## S = 51190000, p-value = 0.0025
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## -0.1183936

cor.test(log2(Mam$GenerationLength_d),Mam$GASKew, method = 'spearman') # rho = 0.2176591, p-value = 2

## Warning in cor.test.default(log2(Mam$GenerationLength_d), Mam$GASKew,
## method = "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##

```

```

## data: log2(Mam$GenerationLength_d) and Mam$GASkew
## S = 35808000, p-value = 2.072e-08
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.2176591

##### multiple Linear Model (it is reasonable to add only 4 significant parameters - without CG and
A<-lm(log2(Mam$GenerationLength_d) ~ Mam$CTSkew + Mam$CGSkew + Mam$CASkew + Mam$TGSkew + Mam$TASkew + M

##
## Call:
## lm(formula = log2(Mam$GenerationLength_d) ~ Mam$CTSkew + Mam$CGSkew +
##      Mam$CASkew + Mam$TGSkew + Mam$TASkew + Mam$GASkew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8126 -0.4573  0.1034  0.5898  2.4283
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   10.394      1.192   8.722 < 2e-16 ***
## Mam$CTSkew     2.076      3.054   0.680  0.4969
## Mam$CGSkew    -10.172     2.391  -4.254 2.42e-05 ***
## Mam$CASkew     7.169      3.450   2.078  0.0381 *
## Mam$TGSkew     3.531      1.567   2.253  0.0246 *
## Mam$TASkew    -1.858      3.502  -0.531  0.5958
## Mam$GASkew    -7.859      4.083  -1.925  0.0547 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9697 on 643 degrees of freedom
## Multiple R-squared:  0.2644, Adjusted R-squared:  0.2575
## F-statistic: 38.52 on 6 and 643 DF,  p-value: < 2.2e-16

# (Intercept)   10.394      1.192   8.722 < 2e-16 ***
# Mam$CTSkew     2.076      3.054   0.680  0.4969
# Mam$CGSkew    -10.172     2.391  -4.254 2.42e-05 ***
# Mam$CASkew     7.169      3.450   2.078  0.0381 *
# Mam$TGSkew     3.531      1.567   2.253  0.0246 *
# Mam$TASkew    -1.858      3.502  -0.531  0.5958
# Mam$GASkew    -7.859      4.083  -1.925  0.0547 .

A<-lm(log2(Mam$GenerationLength_d) ~ Mam$CTSkew + Mam$CASkew + Mam$TGSkew + Mam$TASkew + Mam$GASkew);

##
## Call:
## lm(formula = log2(Mam$GenerationLength_d) ~ Mam$CTSkew + Mam$CASkew +
##      Mam$TGSkew + Mam$TASkew + Mam$GASkew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.82243 -0.50679  0.09509  0.60718  2.50097
##
## Coefficients:

```

```
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.9167      1.0473 12.333 <2e-16 ***
## Mam$CTSkew  4.2264      3.0513  1.385  0.1665
## Mam$CAskew  0.2277      3.0799  0.074  0.9411
## Mam$TGSkew  2.6255      1.5732  1.669  0.0956 .
## Mam$TAskew  1.2575      3.4694  0.362  0.7171
## Mam$GAskew  4.5130      2.9037  1.554  0.1206
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9825 on 644 degrees of freedom
## Multiple R-squared:  0.2437, Adjusted R-squared:  0.2378
## F-statistic: 41.5 on 5 and 644 DF, p-value: < 2.2e-16

A<-lm(log2(Mam$GenerationLength_d) ~ Mam$CTSkew + Mam$CAskew + Mam$TGSkew + Mam$GAskew); summary(A)

##
## Call:
## lm(formula = log2(Mam$GenerationLength_d) ~ Mam$CTSkew + Mam$CAskew +
##     Mam$TGSkew + Mam$GAskew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.81901 -0.48267  0.09342  0.61612  2.50418
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.917      1.047 12.342 < 2e-16 ***
## Mam$CTSkew   3.187      1.045  3.049 0.00239 **
## Mam$CAskew   1.268      1.116  1.136 0.25619
## Mam$TGSkew   2.700      1.558  1.733 0.08361 .
## Mam$GAskew   4.658      2.874  1.621 0.10559
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9818 on 645 degrees of freedom
## Multiple R-squared:  0.2435, Adjusted R-squared:  0.2388
## F-statistic: 51.91 on 4 and 645 DF, p-value: < 2.2e-16

# Estimate Std. Error t value Pr(>|t|)
# (Intercept) 12.917      1.047 12.342 < 2e-16 ***
# Mam$CTSkew   3.187      1.045  3.049 0.00239 **
# Mam$CAskew   1.268      1.116  1.136 0.25619
# Mam$TGSkew   2.700      1.558  1.733 0.08361 .
# Mam$GAskew   4.658      2.874  1.621 0.10559
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# Residual standard error: 0.9818 on 645 degrees of freedom
# Multiple R-squared:  0.2435, Adjusted R-squared:  0.2388
# F-statistic: 51.91 on 4 and 645 DF, p-value: < 2.2e-16

##### plot
pdf("../Body/4Figures/WholeGenomeAnalyses.NoOverlap.AGSkew.R.01.pdf", height = 10, width = 20)
```

```
plot(log2(Mam$GenerationLength_d),Mam$CTSkew, main = 'Mammalia', xlab = 'log2(Generation Length)', ylab = 'CTSkew')
a<-lm(Mam$CTSkew ~ log2(Mam$GenerationLength_d)); summary(a)
```

```
##
## Call:
## lm(formula = Mam$CTSkew ~ log2(Mam$GenerationLength_d))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.49663 -0.11229  0.00834  0.10929  0.39780
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.559964    0.060291  -9.288  <2e-16 ***
## log2(Mam$GenerationLength_d)  0.067226    0.005425  12.392  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1555 on 648 degrees of freedom
## Multiple R-squared:  0.1916, Adjusted R-squared:  0.1903
## F-statistic: 153.6 on 1 and 648 DF,  p-value: < 2.2e-16
```

```
abline(a, col = 'red', lwd = 4)
abline(h=0, col = 'grey', lwd = 2, lt = 2)
```

```
##### GENE_SPECIFIC SKEW FOR EACH SPECIES
```

```
AGG = aggregate(list(SynNuc$NeutralA,SynNuc$NeutralT,SynNuc$NeutralG,SynNuc$NeutralC), by = list(SynNuc$Species, SynNuc$Class),
names(AGG) = c('Species', 'Class', 'Gene', 'NeutralA', 'NeutralT', 'NeutralG', 'NeutralC'))
AGG$CTSkew = (AGG$NeutralC - AGG$NeutralT)/(AGG$NeutralC + AGG$NeutralT); summary(AGG$CTSkew) # GA on h
```

```
## Length Class Mode
##      0    NULL  NULL
```

```
##### GENERATION LENGTH FOR MAMMALS
```

```
GT = read.table("../Body/1Raw/GenerationLengthforMammals.xlsx.txt", header = TRUE, sep = '\t')
GT$Species = gsub(' ','_',GT$Scientific_name)
length(unique(GT$Species))
```

```
## [1] 5424
```

```
summary(GT$AdultBodyMass_g)      # 2      21      71      136058      614 154321304
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
##      2         21        71     136058     614 154321304
```

```
summary(GT$GenerationLength_d) # 129.0  624.4 1101.1 1578.8 2064.3 18980.0 # max = 18980 days => 5
```

```
##      Min. 1st Qu.  Median     Mean 3rd Qu.     Max.
##    129.0  624.4 1101.1 1578.8 2064.3 18980.0
```

```
GT = GT[,c(11,13)]
summary(GT$GenerationLength_d)
```

```
##      Min. 1st Qu.  Median     Mean 3rd Qu.     Max.
##    129.0  624.4 1101.1 1578.8 2064.3 18980.0
```

```

M = merge(AGG,GT, by = 'Species')
summary(M$GenerationLength_d) # median = 2190.0

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    341.3  1458.1  2190.0  2803.9  3650.0 18980.0

ShortLived = unique(M[M$GenerationLength_d <= median(M$GenerationLength_d),]$Species); length(ShortLived)

## [1] 327

LongLived = unique(M[M$GenerationLength_d > median(M$GenerationLength_d),]$Species); length(LongLived)

## [1] 322

MShort = M[M$Species %in% ShortLived,]; MShort$GT = 'short'; MLong = M[M$Species %in% LongLived,]; MLong$GT = 'long'
M = rbind(MShort,MLong)

M$Gene = ordered(M$Gene, levels = c('COX1','COX2','ATP8','ATP6','COX3','ND3','ND4L','ND4','ND5','ND6',
# M$Gene = ordered(M$Gene, levels = c('COX1','COX2','ATP8','ATP6','COX3','ND3','ND4L','ND4','ND5','CytB',
M = M[order(M$Gene),]

# par(mfrow=c(2,1), oma = c(3, 1, 1, 1), cex = 2)
boxplot(CTSkew ~ GT*Gene, data = M, notch = TRUE, outline = FALSE, las = 2, col = c('red','green'), main = c('red','green'))

M = M[!M$Gene %in% c('ND6','ND1','ND2'),]
M$Gene = ordered(M$Gene, levels = c('COX1','COX2','ATP8','ATP6','COX3','ND3','ND4L','ND4','ND5','CytB',
boxplot(CTSkew ~ GT*Gene, data = M, notch = TRUE, outline = FALSE, las = 2, col = c('red','green'), main = c('red','green'))

##### naive multiple linear model, assigning numbers to order of genes. We can improve it substituting
FromGenesToNumbers = data.frame(c('COX1','COX2','ATP8','ATP6','COX3','ND3','ND4L','ND4','ND5','CytB'),stringsAsFactors = FALSE)
M=merge(M,FromGenesToNumbers)

A<-lm(M$CTSkew ~ log2(M$GenerationLength_d) + M$TSSS); summary(A)

##
## Call:
## lm(formula = M$CTSkew ~ log2(M$GenerationLength_d) + M$TSSS)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13370 -0.14668  0.00373  0.15000  0.94851
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.8477990   0.0283319  -29.92  <2e-16 ***
## log2(M$GenerationLength_d)  0.0722798   0.0025027   28.88  <2e-16 ***
## M$TSSS         0.0387429   0.0009798   39.54  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2269 on 6497 degrees of freedom
## Multiple R-squared:  0.2696, Adjusted R-squared:  0.2693
## F-statistic: 1199 on 2 and 6497 DF, p-value: < 2.2e-16

# (Intercept)           -0.8477990   0.0283319  -29.92  <2e-16 ***
# log2(M$GenerationLength_d)  0.0722798   0.0025027   28.88  <2e-16 ***

```



```
# M$TSSS                                0.0387429  0.0009798  39.54  <2e-16 ***
A<-lm(M$CTSkew ~ log2(M$GenerationLength_d) + scale(M$TSSS)); summary(A) # unique(scale(M$TSSS))
```

```
##
## Call:
## lm(formula = M$CTSkew ~ log2(M$GenerationLength_d) + scale(M$TSSS))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13370 -0.14668  0.00373  0.15000  0.94851
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.634713   0.027815  -22.82  <2e-16 ***
## log2(M$GenerationLength_d)  0.072280   0.002503   28.88  <2e-16 ***
## scale(M$TSSS)     0.111289   0.002815   39.54  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2269 on 6497 degrees of freedom
## Multiple R-squared:  0.2696, Adjusted R-squared:  0.2693
## F-statistic: 1199 on 2 and 6497 DF, p-value: < 2.2e-16
```

```
# (Intercept)    -0.634713   0.027815  -22.82  <2e-16 ***
# log2(M$GenerationLength_d)  0.072280   0.002503   28.88  <2e-16 ***
# scale(M$TSSS)     0.111289   0.002815   39.54  <2e-16 ***
```

```
A<-lm(M$CTSkew ~ scale(M$GenerationLength_d) + scale(M$TSSS)); summary(A)
```

```
##
## Call:
## lm(formula = M$CTSkew ~ scale(M$GenerationLength_d) + scale(M$TSSS))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.10702 -0.15167  0.00573  0.15595  0.96423
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.164476   0.002865   57.41  <2e-16 ***
## scale(M$GenerationLength_d) 0.068882   0.002865   24.04  <2e-16 ***
## scale(M$TSSS)     0.111289   0.002865   38.84  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.231 on 6497 degrees of freedom
## Multiple R-squared:  0.2431, Adjusted R-squared:  0.2429
## F-statistic: 1043 on 2 and 6497 DF, p-value: < 2.2e-16
```

```
# (Intercept)     0.164476   0.002865   57.41  <2e-16 ***
# scale(M$GenerationLength_d) 0.068882   0.002865   24.04  <2e-16 ***
# scale(M$TSSS)     0.111289   0.002865   38.84  <2e-16 ***
```

```
A<-lm(M$CTSkew ~ scale(M$GenerationLength_d)*scale(M$TSSS)); summary(A) # interaction is not significant
```

```
##
## Call:
## lm(formula = M$CTSkew ~ scale(M$GenerationLength_d) * scale(M$TSSS))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.10831 -0.15167  0.00645  0.15521  0.96527
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)      0.164476   0.002865  57.412
## scale(M$GenerationLength_d)  0.068882   0.002865  24.042
## scale(M$TSSS)      0.111289   0.002865  38.844
## scale(M$GenerationLength_d):scale(M$TSSS) -0.002585   0.002865  -0.902
##              Pr(>|t|)
## (Intercept)      <2e-16 ***
## scale(M$GenerationLength_d) <2e-16 ***
## scale(M$TSSS)      <2e-16 ***
## scale(M$GenerationLength_d):scale(M$TSSS)    0.367
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.231 on 6496 degrees of freedom
## Multiple R-squared:  0.2432, Adjusted R-squared:  0.2429
## F-statistic: 695.9 on 3 and 6496 DF,  p-value: < 2.2e-16
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.164476	0.002865	57.412	<2e-16 ***
scale(M\$GenerationLength_d)	0.068882	0.002865	24.042	<2e-16 ***
scale(M\$TSSS)	0.111289	0.002865	38.844	<2e-16 ***
scale(M\$GenerationLength_d):scale(M\$TSSS)	-0.002585	0.002865	-0.902	0.367

```
dev.off()

## pdf
## 2
```