LDAvsISSR.R

rob_c

Mon Jun 26 14:59:17 2017

```r
# Simplified/clean theme for plotting
theme_simple <- function (base_size = 12, base_family = "") {
  theme_classic(base_size = base_size, base_family = base_family) %+replace%
    theme(
      axis.text = element_text(colour = "black"),
      axis.title.x = element_text(size=18),
      axis.text.x = element_text(size=12),
      axis.title.y = element_text(size=18,angle=90),
      axis.text.y = element_text(size=12),
      axis.ticks = element_blank(),
      panel.background = element_rect(fill="white"),
      panel.border = element_blank(),
      plot.title=element_text(face="bold", size=24),
      legend.position="none"
    )
}

# Libraries
library(ggplot2) # For plotting
library(proxy) # For similarity matrix calculations using Jaccard's coefficient
```

```
## Warning: package 'proxy' was built under R version 3.3.3
```

```
##
## Attaching package: 'proxy'
```

```
## The following objects are masked from 'package:stats':
##
##     as.dist, dist
```

```
## The following object is masked from 'package:base':
##
##     as.matrix
```

```r
library(reshape) # For reformatting similarity matrices
```

```
## Warning: package 'reshape' was built under R version 3.3.3
```

```
## Import and inspect LDA coordinates and ISSR data
LDA<-read.csv("LDAData.csv",header=T,row.names="IndID")
ISSR<-read.csv("MuhaidatEtAl_ISSRData.csv",header=T,row.names="IndID")

str(LDA)
```

```
## 'data.frame':    57 obs. of  17 variables:
##  $ Loc      : Factor w/ 3 levels "AlYotm","DeadSea",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ BLen     : num  4.8 5 5.4 5.6 5.5 5.2 5.9 5.7 4.9 5.9 ...
##  $ BWdth    : num  0.4 0.5 0.4 0.4 0.4 0.5 0.4 0.5 0.4 0.5 ...
##  $ BVeins   : num  3 5 3.67 3 3 ...
##  $ BSpines  : int  2 2 2 2 2 2 3 2 2 2 ...
##  $ SpLen    : num  1 1 1 1.2 1.1 0.9 0.9 1.3 1 1.5 ...
##  $ NodeLen  : num  1.3 1.6 1.5 1.6 1.3 1.8 1.8 1.6 1.6 2 ...
##  $ LLen     : num  7 7 9 8.5 7.5 10 8 8 9.1 10 ...
##  $ LWdth    : num  1 1 1.1 0.9 0.7 1.2 1 1.2 0.8 1.2 ...
##  $ LTeeth   : num  4 5 4 4 5 ...
##  $ ASFilLen : num  1.3 1.3 1.2 1.3 1.2 1.3 1.3 1.2 1.3 1.3 ...
##  $ ASAnthLen: num  0.5 0.5 0.5 0.5 0.4 0.5 0.5 0.5 0.6 0.5 ...
##  $ ASApenLen: num  0.5 0.5 0.6 0.5 0.5 0.5 0.6 0.5 0.6 0.5 ...
##  $ PSFilLen : num  1.2 1.19 1.1 1.2 1.1 ...
##  $ PSAnthLen: num  0.5 0.5 0.5 0.5 0.5 ...
##  $ LD1      : num  2.37 2.72 4.03 4.41 3.39 ...
##  $ LD2      : num  -0.21 0.514 2.335 -0.588 0.442 ...
```

```
str(ISSR)
```

```
## 'data.frame':    57 obs. of  52 variables:
##  $ X16.280 : int  1 1 0 1 1 1 1 1 1 1 ...
##  $ X16.480 : int  1 1 0 1 1 1 1 1 1 1 ...
##  $ X16.400 : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X16.700 : int  0 0 0 0 0 1 0 0 0 0 ...
##  $ X16.1000: int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X16.1200: int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X14.400 : int  1 1 0 1 1 1 1 1 1 1 ...
##  $ X14.500 : int  0 0 0 0 0 0 0 0 0 1 ...
##  $ X14.800 : int  1 0 0 1 1 0 0 0 0 1 ...
##  $ X14.1000: int  1 0 0 1 1 0 1 0 1 1 ...
##  $ X14.1500: int  1 0 0 0 0 0 0 0 0 1 ...
##  $ X14.600 : int  0 0 0 0 0 0 0 0 0 1 ...
##  $ X11.280 : int  1 0 0 1 1 0 0 0 0 1 ...
##  $ X11.320 : int  1 1 0 1 1 1 1 1 1 1 ...
##  $ X11.370 : int  1 0 0 1 1 0 0 0 0 1 ...
##  $ X11.550 : int  1 0 0 0 1 1 0 0 0 1 ...
##  $ X11.700 : int  1 1 0 1 1 1 0 1 0 1 ...
##  $ X11.1000: int  1 1 0 1 1 1 0 0 0 1 ...
##  $ X11.1200: int  1 0 0 0 0 0 0 0 0 0 ...
##  $ X3.200  : int  0 0 0 1 1 1 1 1 0 0 ...
##  $ X3.300  : int  1 1 1 1 1 1 1 0 0 1 ...
##  $ X3.400  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ X3.500  : int  0 0 0 1 1 1 1 0 0 0 ...
##  $ X3.600  : int  1 1 1 0 0 1 1 1 0 1 ...
##  $ X3.900  : int  1 0 1 0 0 1 1 1 0 1 ...
##  $ X3.950  : int  0 0 0 1 1 0 0 0 0 0 ...
##  $ X3.1050 : int  1 0 1 1 1 1 1 1 0 1 ...
##  $ X9.270  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ X9.310  : int  0 0 1 1 0 1 1 0 1 0 ...
##  $ X9.400  : int  0 0 1 0 0 0 0 0 0 0 ...
##  $ X9.700  : int  0 0 1 0 0 0 0 0 0 0 ...
##  $ X9.1100 : int  0 0 1 0 0 0 0 1 0 0 ...
##  $ X4.600  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X4.500  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X4.400  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X4.350  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X4.300  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X4.250  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X4.200  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X5.200  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ X5.300  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ X5.350  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ X5.550  : int  0 0 0 0 0 1 1 0 0 1 ...
##  $ X5.730  : int  1 1 0 1 1 1 1 1 1 1 ...
##  $ X5.900  : int  0 0 0 0 0 1 1 1 0 1 ...
##  $ X7.300  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X7.350  : int  1 0 1 0 1 0 0 1 1 1 ...
```

```
##  $ X7.370  : int  0 0 0 1 0 0 0 0 0 0 ...
##  $ X7.400  : int  0 1 0 0 1 0 0 1 0 0 ...
##  $ X7.450  : int  1 0 0 0 1 0 0 0 0 0 ...
##  $ X7.500  : int  0 1 1 1 1 0 1 1 1 0 ...
##  $ X7.600  : int  1 0 0 0 0 0 0 0 0 0 ...
```

```
## Nonmetric Multidimensional Scaling (NMDS) of ISSR Data
NMDS<-cmdscale(dist(ISSR, method = "Jaccard",upper=F),k=2)
str(NMDS)
```

```
##  num [1:57, 1:2] -0.1179 0.0617 -0.1436 -0.1183 -0.0809 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : chr [1:57] "D1" "D2" "D3" "D4" ...
##   ..$ : NULL
```

```
#### Merge LDA & NMDS
Coords<-merge(LDA[,c("Loc","LD1","LD2")],NMDS,by="row.names",all=T)
names(Coords)[grep("V[0-9]$",names(Coords))]<-c("NMDS1","NMDS2")
head(Coords)
```

```
##   Row.names    Loc      LD1         LD2       NMDS1       NMDS2
## 1        A1 AlYotm -4.527856  1.24321681 -0.3444122 -0.13830252
## 2       A10 AlYotm -6.014491  0.58384566 -0.2961578  0.11496876
## 3       A11 AlYotm -4.709414 -0.02153503 -0.3092342  0.16413962
## 4       A12 AlYotm -6.373753  1.09055544 -0.2750844 -0.03046156
## 5       A13 AlYotm -7.201540 -0.14078594 -0.2730852  0.13114528
## 6       A14 AlYotm -6.403451 -0.49441686 -0.2413878 -0.06031499
```

```
## Plot NMDS
p<-ggplot(data=Coords,aes(x=NMDS1,y=NMDS2)) +
  stat_ellipse(geom="polygon",aes(colour=Loc),fill=NA,size=1.2,alpha=0.3)+
  stat_ellipse(geom="polygon",aes(fill=Loc,colour=Loc),size=1.2,alpha=0.3)+
  geom_point(aes(colour=Loc),alpha=0.5,size=I(4)) +
  xlab("NMDS Axis1")+ylab("NMDS Axis2")+theme_simple()
print(p)
```
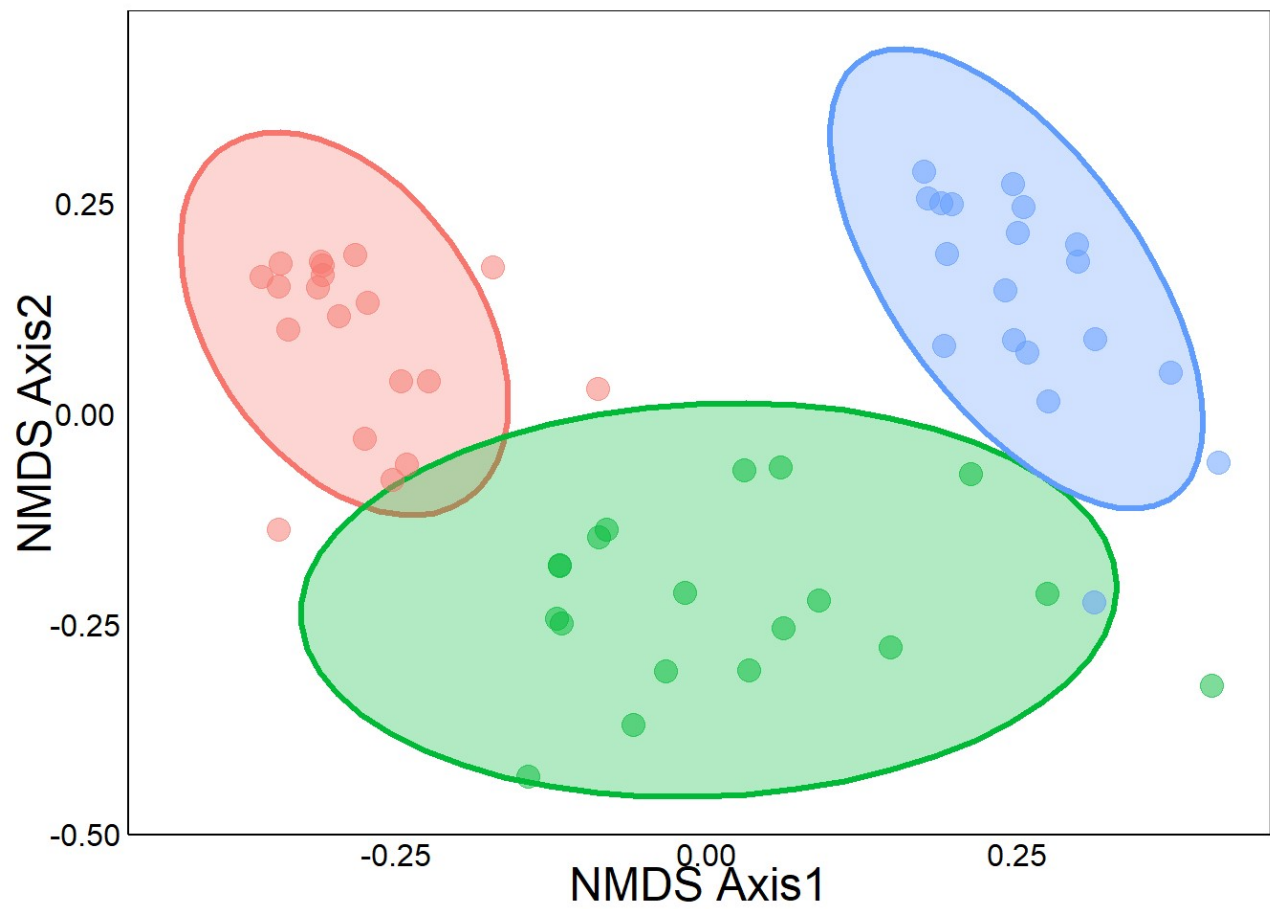
```
### Output as pdf
pdf("NMDSplot.pdf",width=6,height=6)
  print(p)
dev.off()
```

```
## png
##   2
```

```
## Plot LDA
p2<-ggplot(data=Coords,aes(x=LD1,y=LD2)) +
  stat_ellipse(geom="polygon",aes(colour=Loc),fill=NA,size=1.2,alpha=0.3)+
  stat_ellipse(geom="polygon",aes(fill=Loc,colour=Loc),size=1.2,alpha=0.3)+
  geom_point(aes(colour=Loc),alpha=0.5,size=I(4)) +
  xlab("LD Axis1")+ylab("LD Axis2")+theme_simple()
print(p2)
```
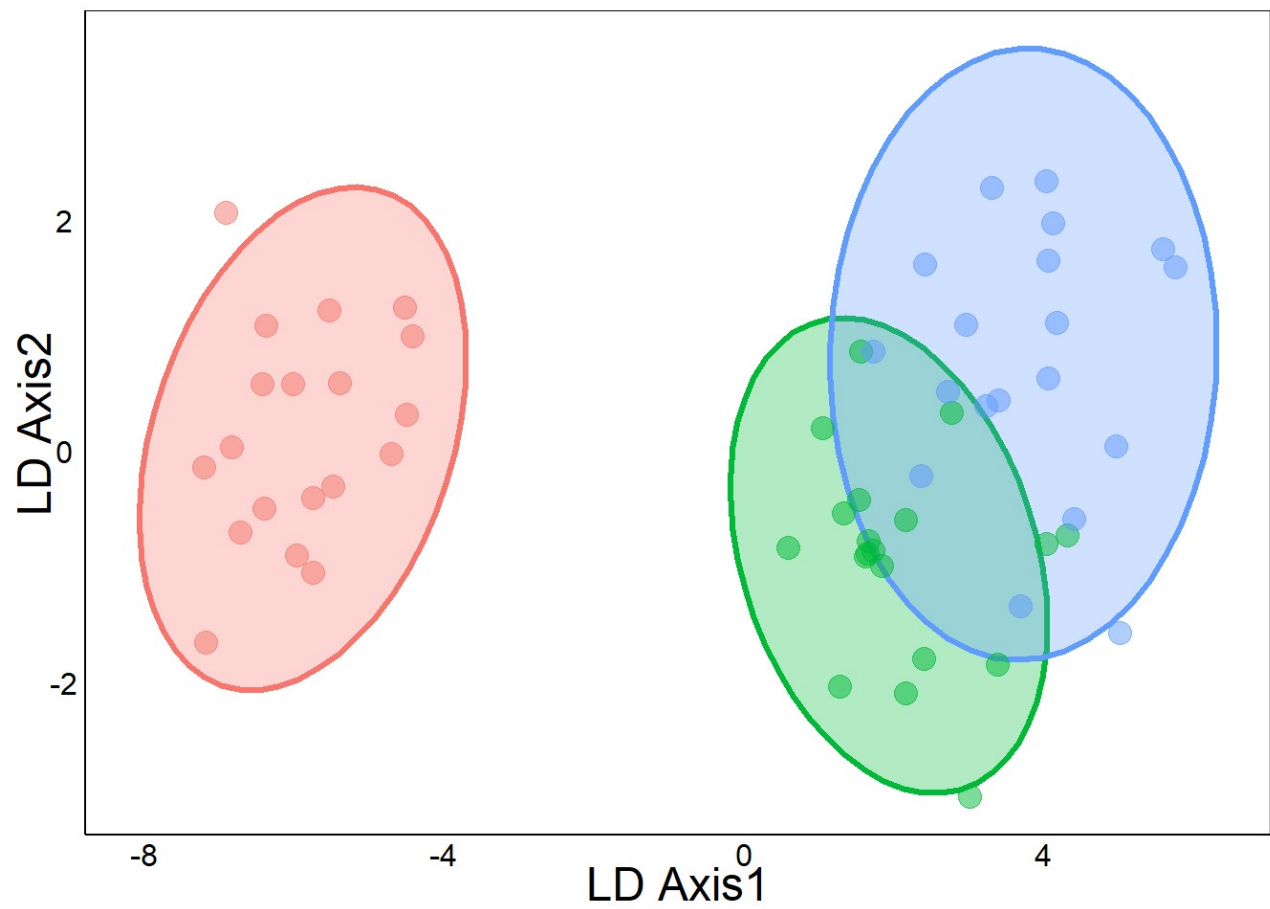
```
pdf("NMDSvsLDAplot.pdf",width=6,height=6)
  print(p2)
dev.off()
```

```
## png
##   2
```

```r
## Calculate Jaccard's distance for molecular and morphological traits
ISSRdist<-melt(as.matrix(dist(ISSR, method = "Jaccard",upper=F)))
Morphdist<-melt(as.matrix(dist(LDA[,grep("LD",names(LDA))],method = "eJaccard",upper=
F)))

### Merge distance matrices
Dist<-merge(ISSRdist,Morphdist,by=c("X1","X2"))
names(Dist)<-c("ID1","ID2","ISSR","Morph")
#### Remove duplicates
for (row in 1:nrow(Dist)){
  Dist[row,grep("ID",names(Dist))]<-sort(Dist[row,grep("ID",names(Dist))])
}
Dist<-Dist[!duplicated(Dist),]
Dist<-Dist[Dist$ID1!=Dist$ID2,]

### Categorize comparisons for colour coding
Dist$Comp<-paste0(gsub("[0-9]*","",Dist$ID1),"-",gsub("[0-9]*","",Dist$ID2))
Dist$Type<-Dist$Comp
#Dist$Type[Dist$Comp %in% c("A-A","D-D","K-K")]<-"Within"
#Dist$Type[Dist$Comp %in% c("A-D","A-K","D-K")]<-"A-D"

## Plot similarity
qplot(ISSR,Morph,colour=Comp,data=Dist)
```
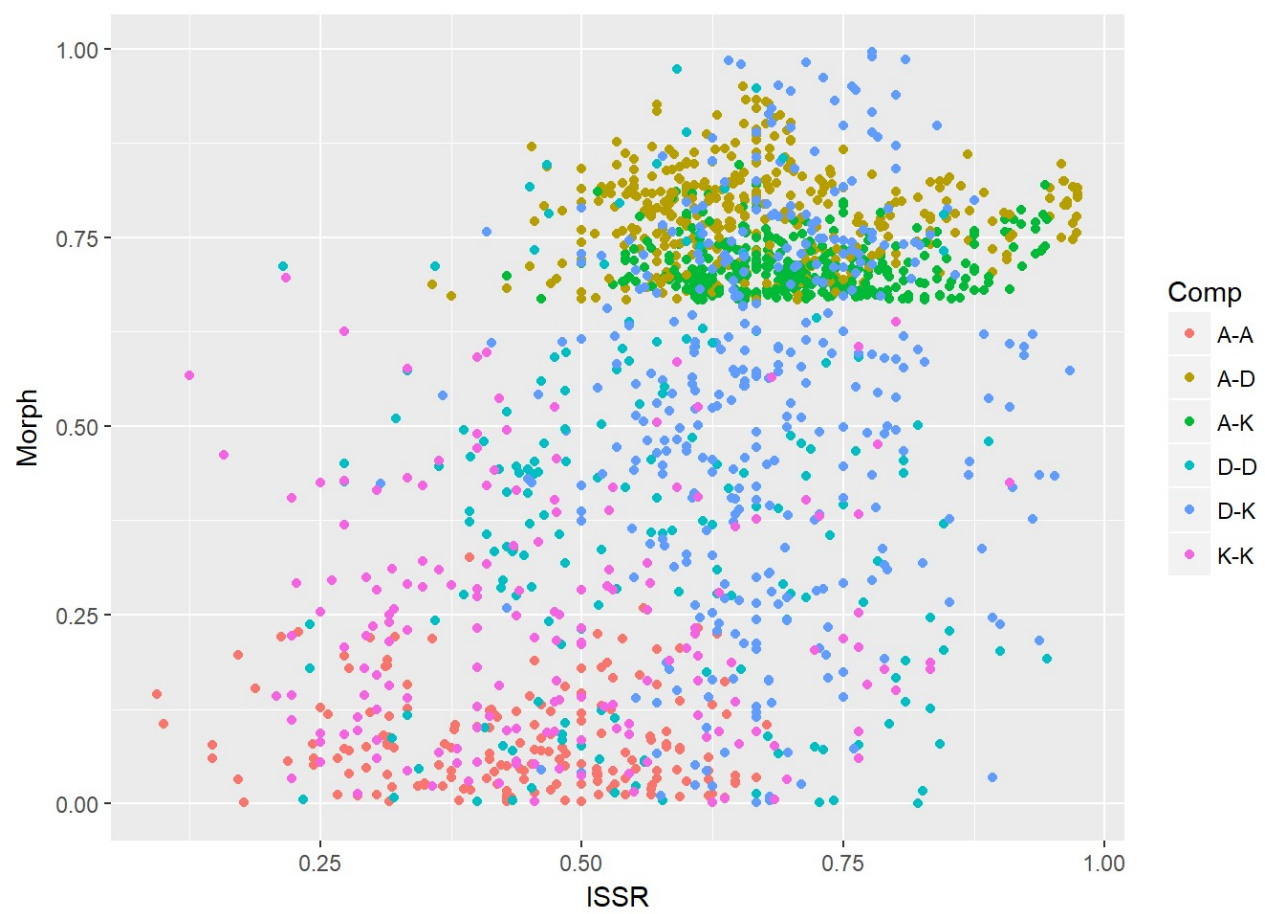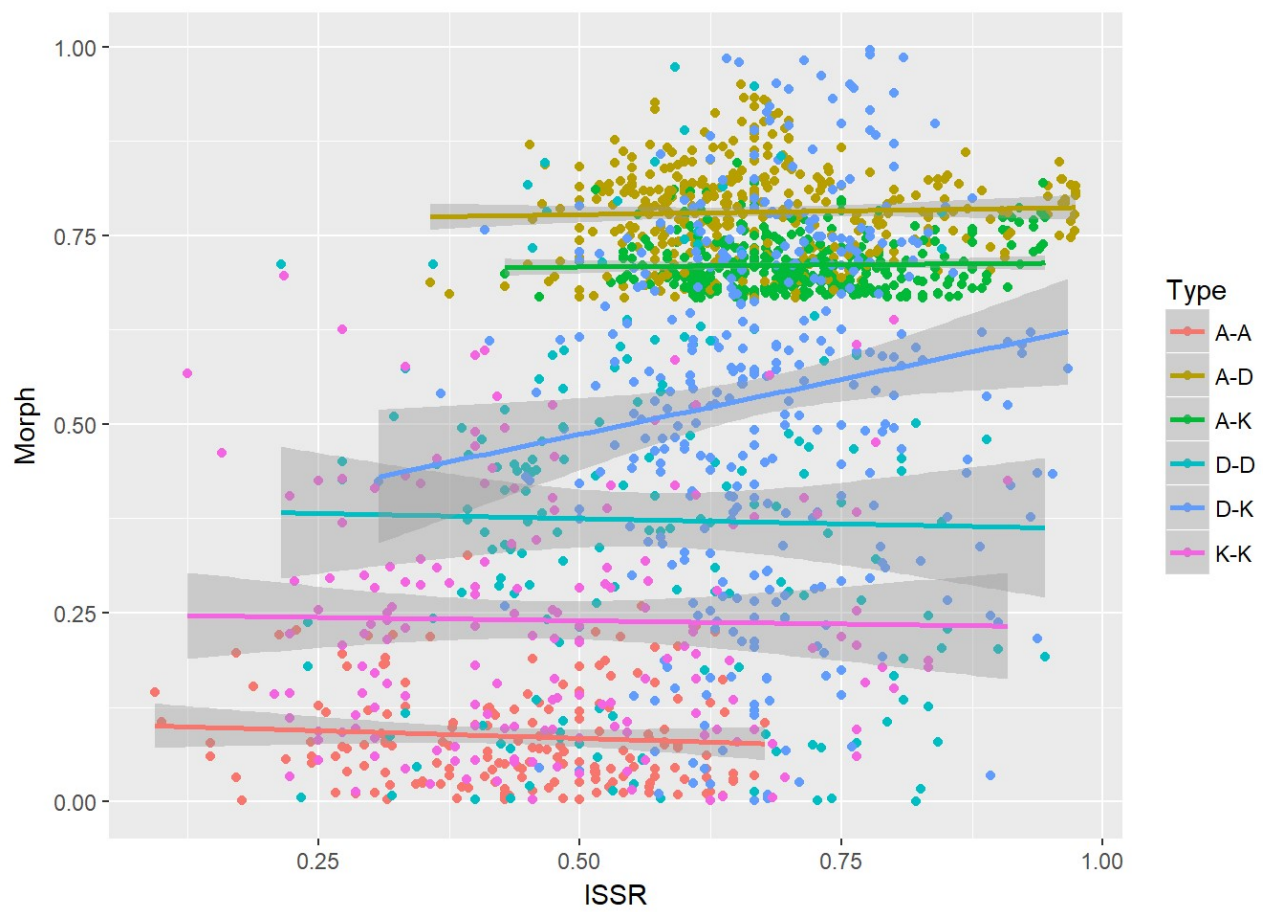
```
qplot(ISSR,Morph,colour=Type,data=Dist) + geom_smooth(method="lm")
```

```
anova(lm(Morph~ISSR+I(ISSR^2),data=Dist))
```

```
## Analysis of Variance Table
##
## Response: Morph
##             Df Sum Sq Mean Sq F value    Pr(>F)
## ISSR         1 29.917 29.9174 524.907 < 2.2e-16 ***
## I(ISSR^2)    1  2.415  2.4154  42.379 1.005e-10 ***
## Residuals 1593 90.794  0.0570
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
p<-ggplot(data=Dist,aes(x=ISSR,y=Morph)) + geom_smooth(method="lm") +
  geom_point(aes(colour=Type),alpha=0.5,size=3) +
  xlab("Genetic Distance")+ylab("Morphological Difference")+theme_simple()
p
```
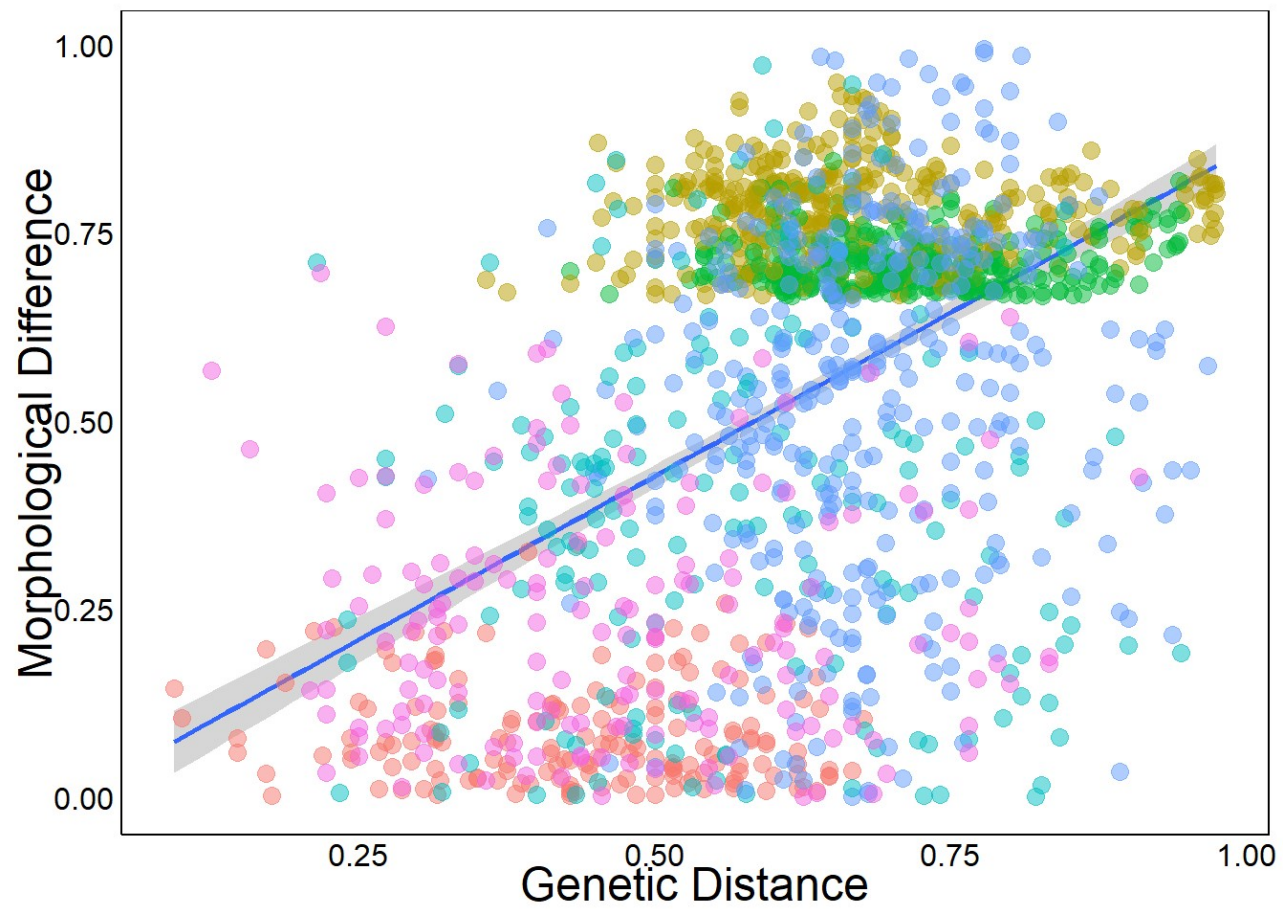
```
## Software version info
sessionInfo()
```

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 15063)
##
## locale:
## [1] LC_COLLATE=English_Canada.1252  LC_CTYPE=English_Canada.1252
## [3] LC_MONETARY=English_Canada.1252 LC_NUMERIC=C
## [5] LC_TIME=English_Canada.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] reshape_0.8.6 proxy_0.4-17  ggplot2_2.2.1
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.9      assertthat_0.1   digest_0.6.12    rprojroot_1.2
##  [5] MASS_7.3-45      plyr_1.8.4       grid_3.3.2       gtable_0.2.0
##  [9] backports_1.0.5  magrittr_1.5     evaluate_0.10    scales_0.4.1
## [13] stringi_1.1.2    lazyeval_0.2.0   rmarkdown_1.3    labeling_0.3
## [17] tools_3.3.2      stringr_1.1.0    munsell_0.4.3    colorspace_1.3-2
## [21] htmltools_0.3.5  knitr_1.15.1     tibble_1.2
```