

Atlas

David Ledesma

3/22/2021

Load in data

```
library(curl)
library(readxl)
Amb_linear_data <- curl("https://raw.githubusercontent.com/TIMAVID/Ambystoma/master/Linear_data/Data/Amb_linear_data.csv")
Amb_linear_data <- read.csv(Amb_linear_data)
```

Tidy data

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

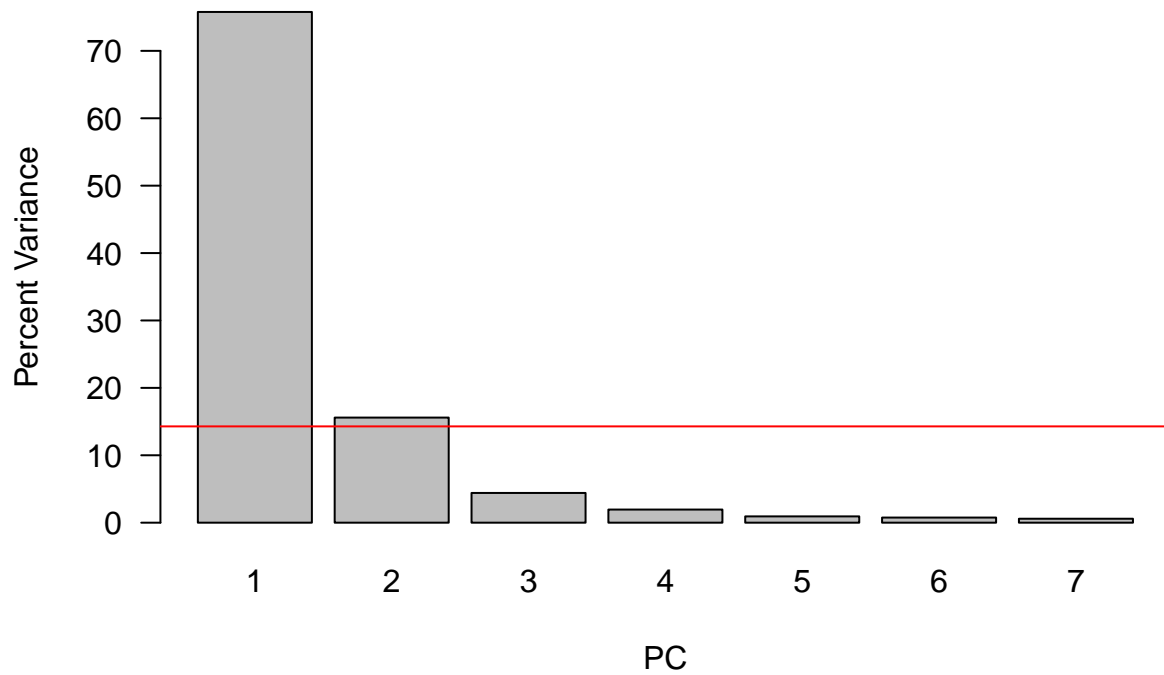
PCA's

```
Atlas.pca <- prcomp(Atlas_wofossil_noNA[c(1:7)], center = TRUE, scale = TRUE) # PCA
# Summary stats #
summary(Atlas.pca)
```

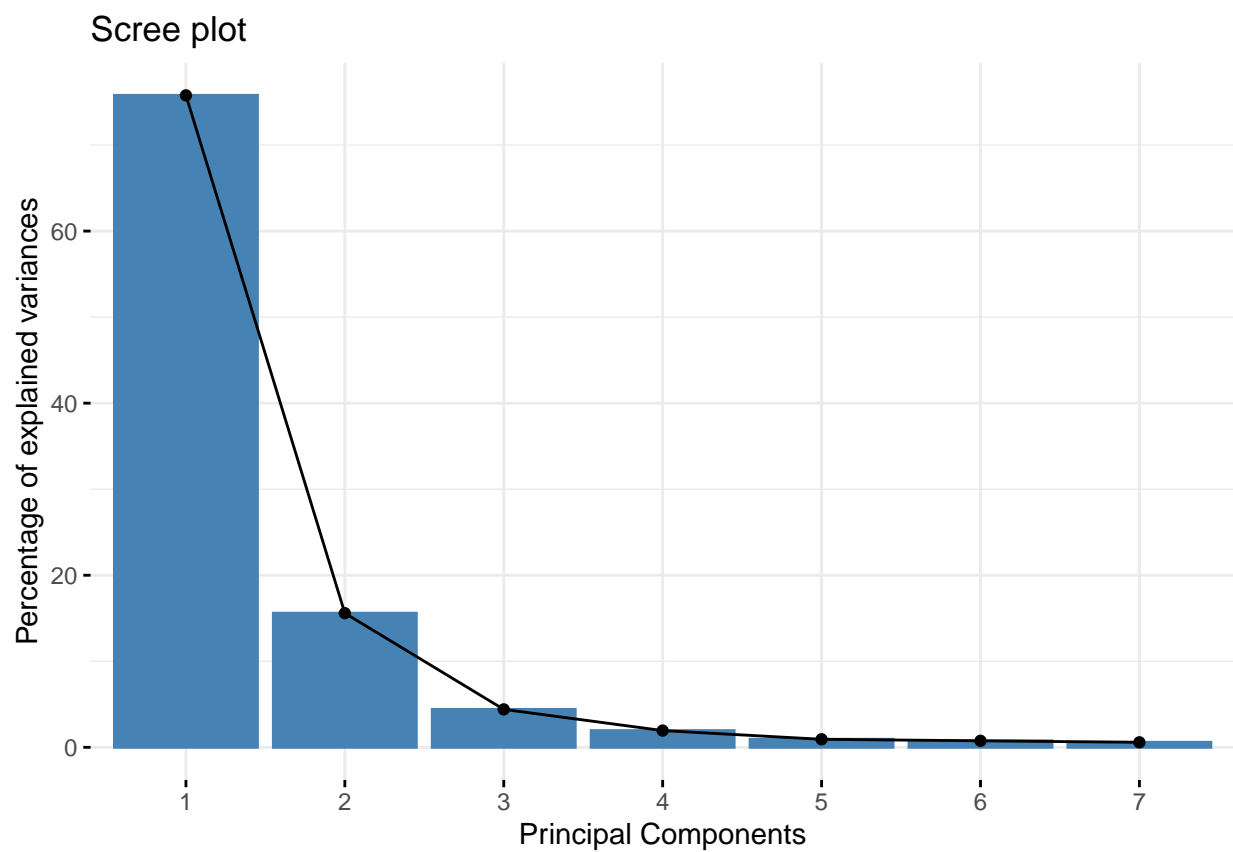
```
## Importance of components:
##              PC1      PC2      PC3      PC4      PC5      PC6      PC7
## Standard deviation    2.3031 1.0447 0.55567 0.36920 0.25577 0.23039 0.20233
## Proportion of Variance 0.7577 0.1559 0.04411 0.01947 0.00935 0.00758 0.00585
## Cumulative Proportion 0.7577 0.9136 0.95775 0.97722 0.98657 0.99415 1.00000
```

```
sd <- Atlas.pca$sdev
loadings <- Atlas.pca$rotation
rownames(loadings) <- colnames(Atlas[c(1:7)])
scores <- Atlas.pca$x
# Show variance explained by PC's #
var <- sd^2
varPercent <- var/sum(var) * 100
barplot(varPercent, xlab = "PC", ylab = "Percent Variance", names.arg = 1:length(varPercent),
```

```
las = 1, ylim = c(0, max(varPercent)), col = "gray")
abline(h = 1/ncol(Atlas[c(1:7)]) * 100, col = "red")
```



```
fviz_eig(Atlas.pca, xlab = "Principal Components")
```



```
# Show loadings #
loadings
```

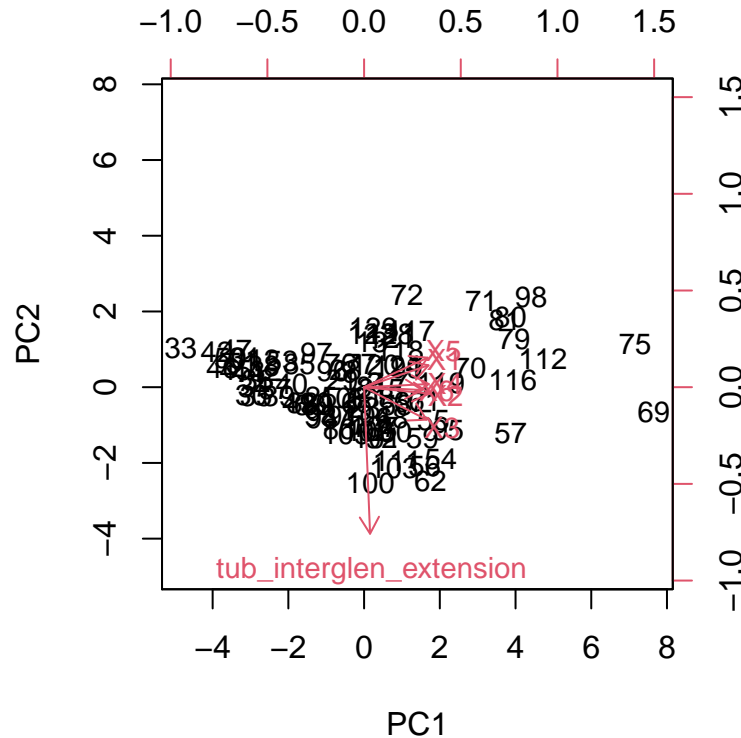
```
##          PC1          PC2          PC3          PC4
## tub_interglen_extension 0.03870518 -0.94714699  0.1063967 -0.23874181
## X1                    0.41652039  0.14147859  0.2646479 -0.04799568
## X2                    0.42197173 -0.04128891  0.1441653  0.29389133
## X3                    0.40400635 -0.21003360 -0.0244320  0.71347417
## X4                    0.41821678  0.02044259  0.1837650 -0.49111174
## X5                    0.41053379  0.19060650  0.2312434 -0.26560742
## X6                    0.37455058 -0.01813215 -0.9000098 -0.18314807
##          PC5          PC6          PC7
## tub_interglen_extension 0.01387657 -0.16293082 -0.07973435
## X1                    0.48873145 -0.09764476 -0.69696451
## X2                    0.47031170 -0.27244438  0.64626642
## X3                   -0.39377589  0.30964846 -0.17911119
## X4                   -0.03666985  0.70273932  0.23351511
## X5                   -0.61565777 -0.53937106  0.03398159
## X6                    0.06570612 -0.09511705 -0.04957562
```

```
sqrt(1/ncol(Atlas[c(1:7)])) # cutoff for 'important' loadings
```

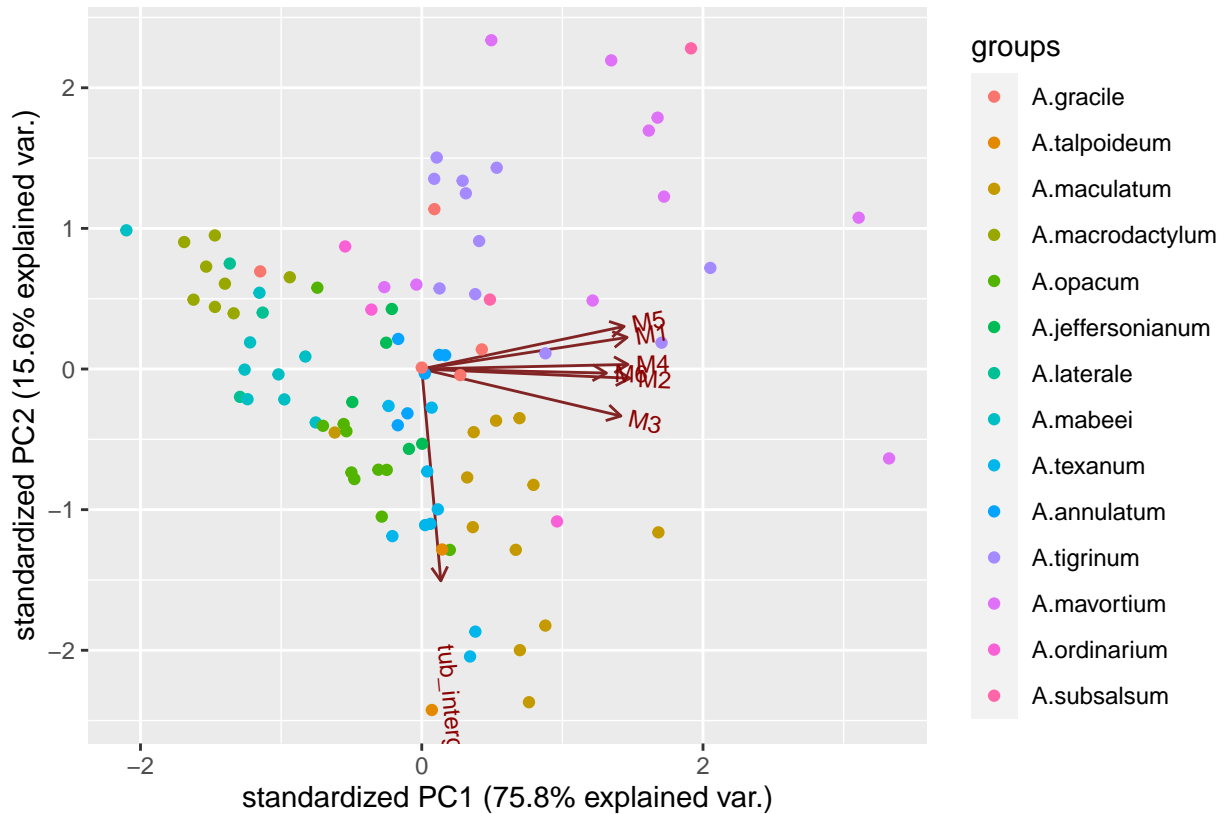
```
## [1] 0.3779645
```

```
# dev.new(height=7, width=7)
```

```
biplot(scores[, 1:2], loadings[, 1:2], cex = 0.9)
```

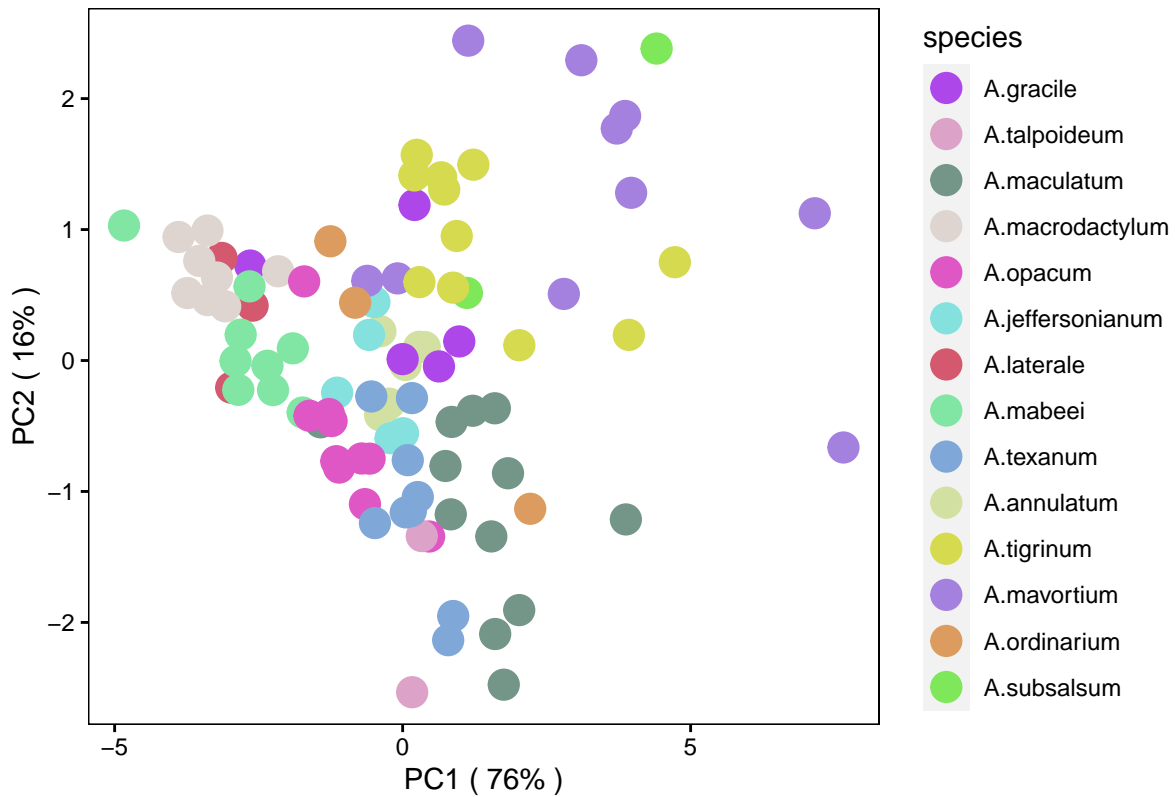


```
ggbiplot(Atlas.pca, ellipse = FALSE, groups = Atlas_wofossil_noNA$species)
```



Plot

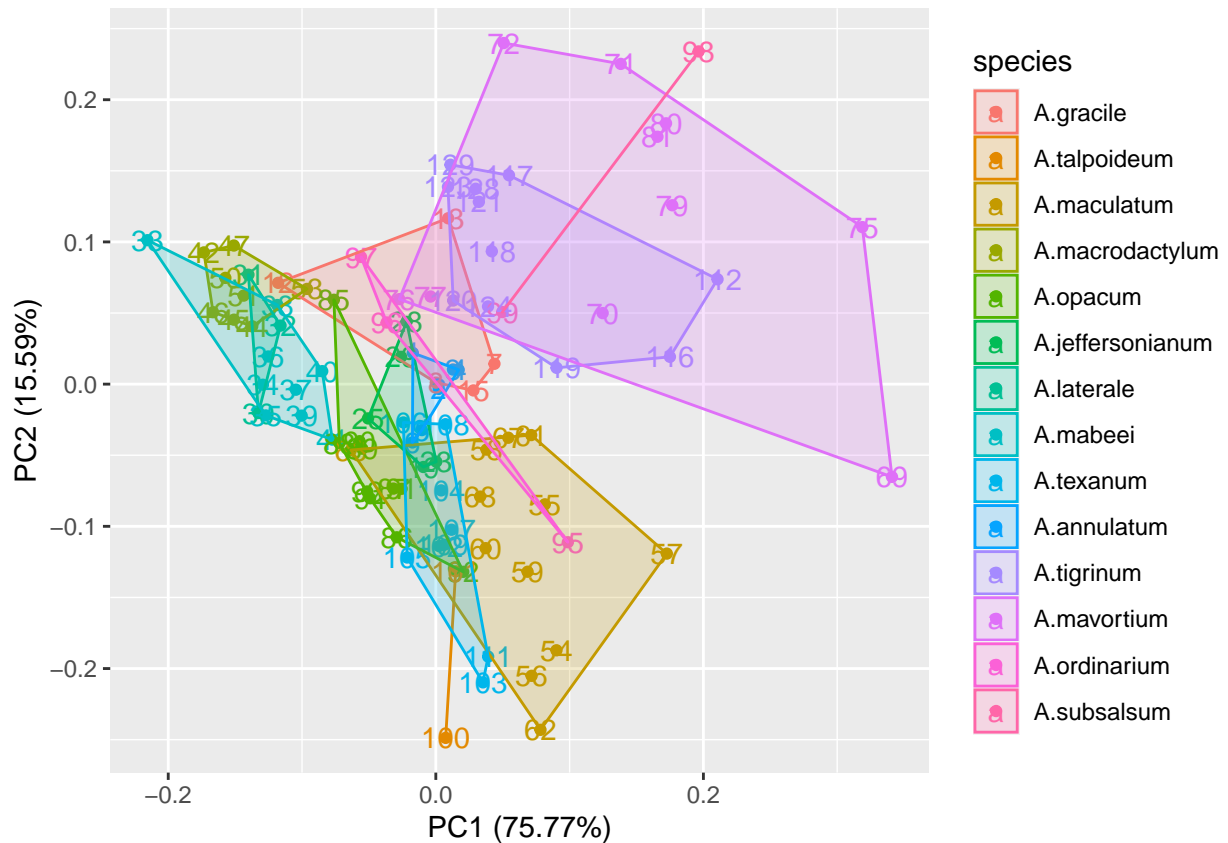
```
scores <- as.data.frame(scores)
scores$species <- Atlas_wofossil_noNA$species # reattach species
theme <- theme(panel.background = element_blank(), panel.border = element_rect(fill = NA),
  panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  strip.background = element_blank(), axis.text.x = element_text(colour = "black"),
  axis.text.y = element_text(colour = "black"), axis.ticks = element_line(colour = "black"),
  plot.margin = unit(c(1, 1, 1, 1), "line"))
percentage <- paste(colnames(scores), "(", paste(as.character(round(varPercent)),
  "%", " )", sep = ""))
n <- 14
palette <- distinctColorPalette(n)
p <- ggplot(scores, aes(x = PC1, y = PC2, color = species))
p <- p + geom_point(size = 5) + theme + xlab(paste(colnames(scores), "(", paste(as.character(round(varPercent)),
  "%", " )", sep = "")) + ylab(paste(colnames(scores), "(", paste(as.character(round(varPercent)),
  "%", " )", sep = "")) +
  scale_color_manual(values = palette)
# p + stat_ellipse()
p
```



```
# alternative plot fviz_pca_ind(Atlas.pca) fviz_pca_ind(Atlas.pca,
# label='none', habillage=Atlas_wofossil_noNA$species, addEllipses=TRUE,
# ellipse.level=0.95, palette = palette) alternative plot 2
autoplot(Atlas.pca, data = Atlas_wofossil_noNA, colour = "species", frame = TRUE,
label = TRUE)
```

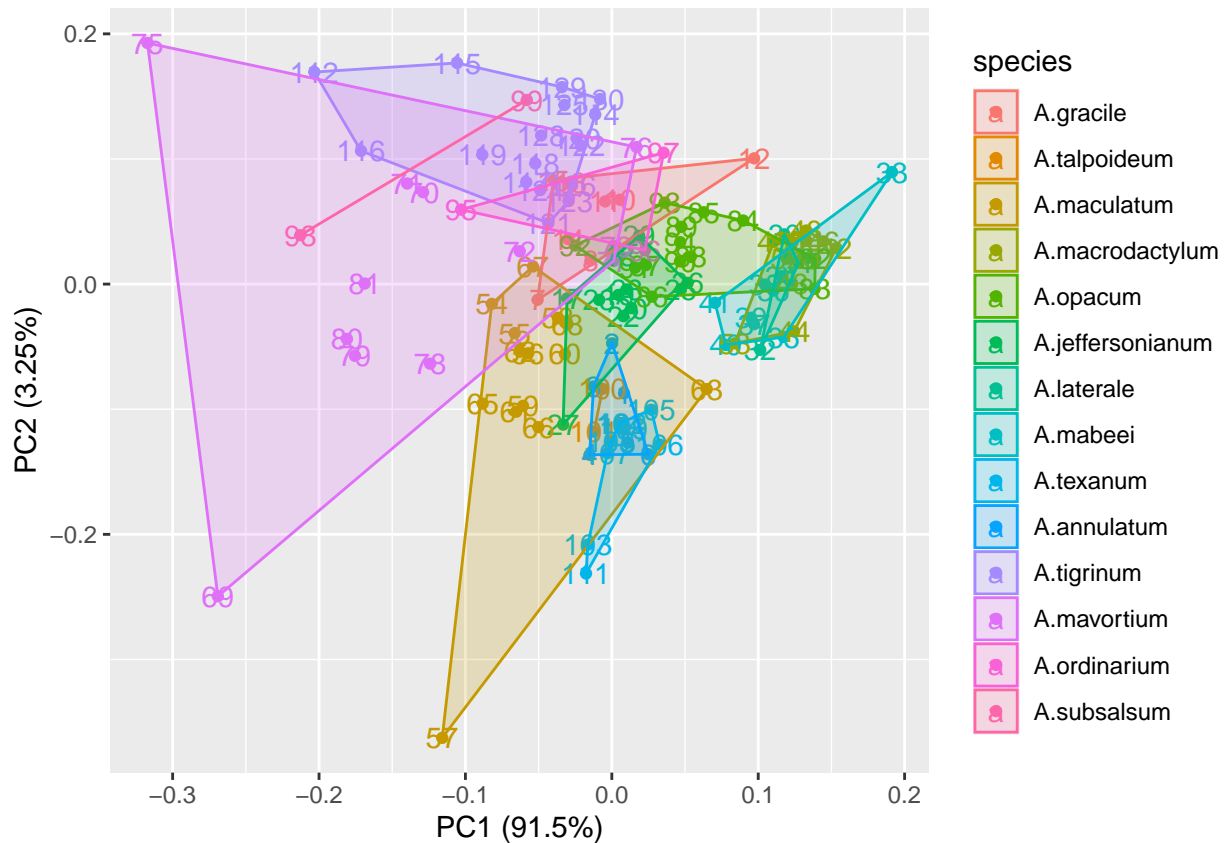
```
## Warning: `select_()` was deprecated in dplyr 0.7.0.
## Please use `select()` instead.

## Warning: `group_by_()` was deprecated in dplyr 0.7.0.
## Please use `group_by()` instead.
## See vignette('programming') for more help
```



No tuberculum interglenoideum ventral extent measurement

```
Atlas_wofossil_noTub <- Atlas_wofossil[, -1]
Atlas_wofossil_noTub <- na.omit(Atlas_wofossil_noTub) # remove rows with N/A's
Atlas_wofossil_noTub$species <- as.factor(Atlas_wofossil_noTub$species)
Atlas_wofossil_noTub$species <- factor(Atlas_wofossil_noTub$species, levels = c("A.gracile",
  "A.talpoideum", "A.maculatum", "A.macrodactylum", "A.opacum", "A.jeffersonianum",
  "A.laterale", "A.mabeei", "A.texanum", "A.annulatum", "A.tigrinum", "A.mavortium",
  "A.ordinarium", "A.subsalsum")) # Reorder species
Atlas.pca_2 <- prcomp(Atlas_wofossil_noTub[c(1:6)], center = TRUE, scale = FALSE) # PCA
autoplot(Atlas.pca_2, data = Atlas_wofossil_noTub, colour = "species", frame = TRUE,
  label = TRUE)
```



Statistical Tests

```
# Removing 'A.laterale|A.talpoideum|A.subsalsum|A.ordinarium' due to low
# sample sizes#
Atlas_wofossil_noTub_sub <- dplyr::filter(Atlas_wofossil_noTub, !grepl("A.laterale|A.talpoideum|A.subsalsum|A.ordinarium",
species))
Atlas_wofossil_noTub_sub$species <- factor(Atlas_wofossil_noTub_sub$species,
levels = c("A.gracile", "A.maculatum", "A.macrodactylum", "A.opacum", "A.jeffersonianum",
"A.mabeei", "A.texanum", "A.annulatum", "A.tigrinum", "A.mavortium")) # Reorder species
```

Various checks for MANOVA

```
# Atlas_wofossil_noTub_sub <- Atlas_wofossil_noTub_sub %>% add_column(id =
# rownames(Atlas_wofossil_noTub_sub), .after = 8)
# Check sample sizes:PASS
Atlas_wofossil_noTub_sub %>% dplyr::group_by(species) %>% dplyr::summarise(N = n())

## # A tibble: 10 x 2
##   species      N
##   <fct>      <int>
## 1 A.gracile      8
## 2 A.maculatum    14
## 3 A.macrodactylum 12
```

```
## 4 A.opacum 13
## 5 A.jeffersonianum 11
## 6 A.mabeei 9
## 7 A.texanum 9
## 8 A.annulatum 6
## 9 A.tigrinum 17
## 10 A.mavortium 12
```

```
# Identify univariate outliers for each variable:FAIL
```

```
Atlas_wofossil_noTub_sub %>% dplyr::group_by(species) %>% identify_outliers(5) #input variable column
```

```
## # A tibble: 15 x 10
```

##	species	M1	M2	M3	M4	M5	M6	specimen_num	is.outlier
##	<fct>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<lgl>
## 1	A.gracile	2.27	2.34	0.86	1.67	2.68	0.98	A.gracile_UTEP_16~	TRUE
## 2	A.maculatum	2.58	2.97	1.24	1.95	2.47	0.97	A.maculatum_UTEP_~	TRUE
## 3	A.macrodac~	2.62	2.68	0.98	1.86	2.28	0.96	A.macrodactylum_U~	TRUE
## 4	A.opacum	1.93	2.18	0.85	1.47	2.13	0.83	A.opacum_UTEP_1276	TRUE
## 5	A.opacum	2.05	2.34	0.78	1.25	2.09	0.78	A.opacum_UTEP_1280	TRUE
## 6	A.opacum	2.54	2.53	0.87	1.74	2.49	0.84	A.opacum_UTEP_1285	TRUE
## 7	A.opacum	3.46	3.39	1.27	2.6	3.35	1.17	A.opacum_UTEP_1310	TRUE
## 8	A.jefferso~	2.89	2.79	1.14	2.11	2.63	0.99	A.jeffersonianum_~	TRUE
## 9	A.mabeei	1.75	1.44	0.6	1.19	1.72	0.78	A.mabeei_UTEP_1576	TRUE
## 10	A.mabeei	2.51	2.71	1.01	1.94	2.6	1.15	A.mabeei_UTEP_1584	TRUE
## 11	A.texanum	2.8	3.19	1.28	2.31	2.62	1.26	A.texanum_UTEP_17~	TRUE
## 12	A.tigrinum	4.52	4.54	1.87	3.63	5.38	1.73	A.tigrinum_UTEP_1~	TRUE
## 13	A.tigrinum	3.69	3.67	1.47	3.35	4.49	1.51	A.tigrinum_UTEP_1~	TRUE
## 14	A.tigrinum	4.46	4.35	1.66	3.36	4.91	1.79	A.tigrinum_UTEP_1~	TRUE
## 15	A.mavortium	5.65	5.13	1.74	4.97	6.03	2.18	A.mavortium_UTEP_~	TRUE

```
## # ... with 1 more variable: is.extreme <lgl>
```

```
Atlas_wofossil_noTub_sub %>% dplyr::group_by(species) %>% identify_outliers(6) #input variable column
```

```
## # A tibble: 7 x 10
```

##	species	M1	M2	M3	M4	M5	M6	specimen_num	is.outlier	is.extreme
##	<fct>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<lgl>	<lgl>
## 1	A.macu~	3.68	4.18	1.58	3.13	3.71	3.29	A.maculatum~	TRUE	TRUE
## 2	A.mabe~	2.51	2.71	1.01	1.94	2.6	1.15	A.mabeei_UT~	TRUE	FALSE
## 3	A.annu~	3.38	3.38	1.17	2.32	2.73	1.08	A.annulatum~	TRUE	FALSE
## 4	A.tigr~	4.52	4.54	1.87	3.63	5.38	1.73	A.tigrinum_~	TRUE	FALSE
## 5	A.tigr~	3.45	2.99	1.51	2.18	3.56	1.11	A.tigrinum_~	TRUE	FALSE
## 6	A.tigr~	4.46	4.35	1.66	3.36	4.91	1.79	A.tigrinum_~	TRUE	FALSE
## 7	A.mavo~	4.92	5.12	2.22	3.87	5.32	3.65	A.mavortium~	TRUE	TRUE

```
# ...
```

```
# Detect multivariate outliers:MOSTLY PASS
```

```
head(mahalanobis_distance(Atlas_wofossil_noTub_sub[, 1:6]))
```

##	M1	M2	M3	M4	M5	M6	mahal.dist	is.outlier
## 1	3.38	3.38	1.17	2.32	2.73	1.08	6.035	FALSE
## 2	3.36	3.21	1.31	2.49	2.91	1.22	6.431	FALSE
## 3	3.50	3.42	1.29	2.46	2.96	1.30	4.895	FALSE
## 4	3.42	3.61	1.30	2.57	2.87	1.28	4.941	FALSE
## 5	3.12	3.44	1.24	2.29	2.77	1.24	2.674	FALSE
## 6	3.05	3.18	1.24	2.27	2.61	1.38	4.128	FALSE


```
# Check univariate normality assumption:FAIL
Atlas_wofossil_noTub_sub %>% group_by(species) %>% shapiro_test(M1, M2, M3,
  M4, M5, M6) %>% arrange(variable)
```

```
## # A tibble: 60 x 4
##   species      variable statistic      p
##   <fct>      <chr>      <dbl>    <dbl>
## 1 A.gracile      M1          0.782 0.0183
## 2 A.maculatum    M1          0.787 0.00350
## 3 A.macrodactylum M1          0.968 0.894
## 4 A.opacum       M1          0.957 0.703
## 5 A.jeffersonianum M1          0.945 0.576
## 6 A.mabeei       M1          0.909 0.308
## 7 A.texanum      M1          0.972 0.910
## 8 A.annulatum    M1          0.885 0.293
## 9 A.tigrinum     M1          0.787 0.00136
## 10 A.mavortium   M1          0.961 0.791
## # ... with 50 more rows
```

```
# Check Multivariate normality:FAIL
Atlas_wofossil_noTub_sub %>% dplyr::select(, 1:6) %>% mshapiro_test()
```

```
## # A tibble: 1 x 2
##   statistic p.value
##   <dbl>    <dbl>
## 1      0.482 4.95e-18
```

```
# Identify multicollinearity:FAIL
Atlas_wofossil_noTub_sub %>% cor_test(, 1:6)
```

```
## # A tibble: 36 x 8
##   var1 var2 cor statistic      p conf.low conf.high method
##   <chr> <chr> <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <chr>
## 1 M1    M1    1      Inf    0.      1      1      Pearson
## 2 M1    M2    0.94   30.2 1.02e-54 0.921    0.962 Pearson
## 3 M1    M3    0.85   17.1 1.29e-32 0.793    0.897 Pearson
## 4 M1    M4    0.93   25.9 2.59e-48 0.896    0.950 Pearson
## 5 M1    M5    0.93   27.0 4.50e-50 0.903    0.953 Pearson
## 6 M1    M6    0.76   12.1 7.33e-22 0.664    0.827 Pearson
## 7 M2    M1    0.94   30.2 1.02e-54 0.921    0.962 Pearson
## 8 M2    M2    1      Inf    0.      1      1      Pearson
## 9 M2    M3    0.92   23.8 5.21e-45 0.880    0.942 Pearson
## 10 M2   M4    0.92   24.8 1.22e-46 0.888    0.946 Pearson
## # ... with 26 more rows
```

```
# PROBLEM!! Absence of multicollinearity. The dependent (outcome) variables
# cannot be too correlated to each other. No correlation should be above r =
# 0.90 [Tabachnick and Fidell (2012)].
cor(Atlas_wofossil_noTub_sub[, 1:6])
```

```
##           M1           M2           M3           M4           M5           M6
## M1 1.0000000 0.9449857 0.8533793 0.9272517 0.9326445 0.7568932
## M2 0.9449857 1.0000000 0.9158753 0.9217077 0.8823482 0.7989239
## M3 0.8533793 0.9158753 1.0000000 0.8476626 0.8159639 0.7816608
## M4 0.9272517 0.9217077 0.8476626 1.0000000 0.9208258 0.7833759
## M5 0.9326445 0.8823482 0.8159639 0.9208258 1.0000000 0.7618434
```

```
## M6 0.7568932 0.7989239 0.7816608 0.7833759 0.7618434 1.0000000
```

MANOVA# : *Failed multiple checks

```
Atlas.man <- manova(cbind(M1, M2, M3, M4, M5, M6) ~ species, data = Atlas_wofossil_noTub_sub)
summary(Atlas.man)
```

```
##              Df Pillai approx F num Df den Df      Pr(>F)
## species        9 2.1061      6.07    54    606 < 2.2e-16 ***
## Residuals    101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary.aov(Atlas.man)
```

```
## Response M1 :
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        9 40.104  4.4560  27.517 < 2.2e-16 ***
## Residuals    101 16.355  0.1619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Response M2 :
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        9 37.119  4.1244  26.124 < 2.2e-16 ***
## Residuals    101 15.945  0.1579
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Response M3 :
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        9 5.5115 0.61239  21.615 < 2.2e-16 ***
## Residuals    101 2.8615 0.02833
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Response M4 :
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        9 21.584 2.39828  16.607 < 2.2e-16 ***
## Residuals    101 14.586 0.14441
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Response M5 :
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        9 52.724  5.8582  35.407 < 2.2e-16 ***
## Residuals    101 16.711  0.1655
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Response M6 :
##              Df Sum Sq Mean Sq F value    Pr(>F)
## species        9  8.1785 0.90872   9.1615 4.933e-10 ***
## Residuals    101 10.0181 0.09919
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Permutation MANOVA ## *Overcome failed checks

```
# Permutational Multivariate Analysis of Variance Using Distance Matrices
adonis(Atlas_wofossil_noTub_sub[, 1:6] ~ species, data = Atlas_wofossil_noTub_sub,
       method = "euclidean")
```

```
##
## Call:
## adonis(formula = Atlas_wofossil_noTub_sub[, 1:6] ~ species, data = Atlas_wofossil_noTub_sub,
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## species        9   165.222  18.3580  24.245 0.68359 0.001 ***
## Residuals    101    76.477   0.7572      0.31641
## Total       110   241.698      1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# pairwise comparisons between group levels with corrections for multiple
# testing
pairwise.perm.manova(Atlas_wofossil_noTub_sub[, 1:6], Atlas_wofossil_noTub_sub$species,
                    nperm = 50, progress = FALSE) #needs more permutation but takes a long time
```

```
##
## Pairwise comparisons using permutation MANOVAs (test: Pillai)
##
## data:  Atlas_wofossil_noTub_sub[, 1:6] by Atlas_wofossil_noTub_sub$species
## 50 permutations
##
##              A.gracile A.maculatum A.macrodactylum A.opacum
## A.maculatum      0.023      -            -            -
## A.macrodactylum 0.023      0.023          -            -
## A.opacum          0.118      0.023          0.023          -
## A.jeffersonianum 0.042      0.023          0.023          0.023
## A.mabeei          0.023      0.023          0.062          0.023
## A.texanum         0.023      0.042          0.023          0.023
## A.annulatum       0.023      0.023          0.023          0.023
## A.tigrinum        0.042      0.023          0.023          0.023
## A.mavortium       0.100      0.023          0.023          0.023
##
##              A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.maculatum      -            -            -            -
## A.macrodactylum -            -            -            -
## A.opacum          -            -            -            -
## A.jeffersonianum -            -            -            -
## A.mabeei          0.023      -            -            -
## A.texanum         0.023      0.023      -            -
## A.annulatum       0.023      0.023      0.042      -            -
```

```
## A.tigrinum      0.023      0.023      0.023      0.023      -
## A.mavortium     0.023      0.023      0.023      0.023      0.023
##
## P value adjustment method: fdr
# or using euclidean distances
AtlasPPM <- pairwise.perm.manova(dist(Atlas_wofossil_noTub_sub[, 1:6], "euclidean"),
  Atlas_wofossil_noTub_sub$species, nperm = 999)
AtlasPPM

##
## Pairwise comparisons using permutation MANOVAs on a distance matrix
##
## data: dist(Atlas_wofossil_noTub_sub[, 1:6], "euclidean") by Atlas_wofossil_noTub_sub$species
## 999 permutations
##
##      A.gracile A.maculatum A.macrodactylum A.opacum
## A.maculatum    0.0084      -                -                -
## A.macrodactylum 0.0019    0.0019          -                -
## A.opacum        0.0093    0.0019          0.0019          -
## A.jeffersonianum 0.2550    0.0019          0.0019          0.0084
## A.mabeei        0.0019    0.0019          0.1141          0.0084
## A.texanum       0.0084    0.0019          0.0019          0.0019
## A.annulatum     0.0771    0.0050          0.0019          0.0084
## A.tigrinum      0.0254    0.0315          0.0019          0.0019
## A.mavortium     0.0163    0.0170          0.0019          0.0019
##
##      A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.maculatum          -          -          -          -
## A.macrodactylum    -          -          -          -
## A.opacum             -          -          -          -
## A.jeffersonianum    -          -          -          -
## A.mabeei            0.0019      -          -          -
## A.texanum           0.0093      0.0019      -          -
## A.annulatum         0.1268      0.0019    0.0249      -
## A.tigrinum          0.0019      0.0019    0.0019    0.0035
## A.mavortium         0.0019      0.0019    0.0035    0.0103    0.0340
##
## P value adjustment method: fdr
```

tuberculum interglenoideum measurement only

```
Atlas_wofossil_Tub_only <- as.data.frame(Atlas_wofossil[c(1, 8:9)])
Atlas_wofossil_Tub_only <- na.omit(Atlas_wofossil_Tub_only) # remove rows with N/A's
Atlas_wofossil_Tub_only$species <- as.factor(Atlas_wofossil_Tub_only$species)
Atlas_wofossil_Tub_only$species <- factor(Atlas_wofossil_Tub_only$species, levels = c("A.gracile",
  "A.maculatum", "A.macrodactylum", "A.opacum", "A.jeffersonianum", "A.mabeei",
  "A.texanum", "A.annulatum", "A.tigrinum", "A.mavortium")) # Reorder species
Atlas_wofossil_Tub_only <- dplyr::filter(Atlas_wofossil_Tub_only, !grepl("A.laterale|A.talpoideum|A.sub",
  species))
# Permutational Anova
perm.anova(Atlas_wofossil_Tub_only$tub_interglen_extension ~ Atlas_wofossil_Tub_only$species,
  nperm = 1000)
```

```
## |
```

```
|
```

```
## Permutation Analysis of Variance Table
##
## Response: Atlas_wofossil_Tub_only$tub_interglen_extension
## 1000 permutations
##
##              Sum Sq Df  Mean Sq F value    Pr(>F)
## Atlas_wofossil_Tub_only$species 0.79265   9 0.088073  16.161 0.000999 ***
## Residuals                0.43597  80 0.005450
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# pairwise comparisons between group levels with corrections for multiple
# testing
library(rcompanion)

##
## Attaching package: 'rcompanion'

## The following object is masked from 'package:psych':
##
##      phi

PT <- pairwisePermutationTest(tub_interglen_extension ~ species, data = Atlas_wofossil_Tub_only,
                              method = "fdr")
PT

##              Comparison      Stat    p.value p.adjust
## 1      A.gracile - A.maculatum = 0   -2.822  0.004772 0.014320
## 2      A.gracile - A.macrodactylum = 0    2.412  0.01585 0.031010
## 3      A.gracile - A.opacum = 0    -2.317  0.02051 0.032960
## 4      A.gracile - A.jeffersonianum = 0  -0.9899  0.3222 0.362500
## 5      A.gracile - A.mabeei = 0     0.539  0.5899 0.617300
## 6      A.gracile - A.texanum = 0    -2.611  0.009034 0.021400
## 7      A.gracile - A.annulatum = 0   -0.3469  0.7287 0.745300
## 8      A.gracile - A.tigrinum = 0     2.177  0.02947 0.044200
## 9      A.gracile - A.mavortium = 0     2.255  0.02415 0.037470
## 10     A.maculatum - A.macrodactylum = 0   3.711 0.0002068 0.002126
## 11     A.maculatum - A.opacum = 0     1.505  0.1323 0.170100
## 12     A.maculatum - A.jeffersonianum = 0   2.361  0.01823 0.032800
## 13     A.maculatum - A.mabeei = 0     3.31  0.000934 0.004203
## 14     A.maculatum - A.texanum = 0    0.7083  0.4788 0.513000
## 15     A.maculatum - A.annulatum = 0     2.744  0.00607 0.016070
## 16     A.maculatum - A.tigrinum = 0     3.9 9.615e-05 0.002126
## 17     A.maculatum - A.mavortium = 0     3.798 0.000146 0.002126
## 18     A.macrodactylum - A.opacum = 0   -3.564 0.0003649 0.002126
## 19     A.macrodactylum - A.jeffersonianum = 0  -3.043 0.002342 0.008107
## 20     A.macrodactylum - A.mabeei = 0   -2.511  0.01205 0.024650
## 21     A.macrodactylum - A.texanum = 0    -3.6 0.0003188 0.002126
## 22     A.macrodactylum - A.annulatum = 0   -3.144 0.001667 0.006450
## 23     A.macrodactylum - A.tigrinum = 0  -0.04369  0.9652 0.965200
## 24     A.macrodactylum - A.mavortium = 0    1.122  0.262 0.318200
## 25     A.opacum - A.jeffersonianum = 0     1.699  0.08939 0.120000
## 26     A.opacum - A.mabeei = 0     2.913  0.003578 0.011500
## 27     A.opacum - A.texanum = 0   -0.8956  0.3705 0.406600
## 28     A.opacum - A.annulatum = 0     2.347  0.01895 0.032800
## 29     A.opacum - A.tigrinum = 0     3.62 0.0002945 0.002126
## 30     A.opacum - A.mavortium = 0     3.433 0.0005967 0.002984
```

```
## 31      A.jeffersonianum - A.mabeei = 0      1.692      0.09068 0.120000
## 32      A.jeffersonianum - A.texanum = 0     -2.133      0.03296 0.046350
## 33      A.jeffersonianum - A.annulatum = 0      1.017      0.3089 0.356400
## 34      A.jeffersonianum - A.tigrinum = 0      2.771      0.00559 0.015720
## 35      A.jeffersonianum - A.mavortium = 0      2.548      0.01084 0.023230
## 36      A.mabeei - A.texanum = 0     -3.135      0.00172 0.006450
## 37      A.mabeei - A.annulatum = 0     -1.106      0.2687 0.318200
## 38      A.mabeei - A.tigrinum = 0      2.139      0.03245 0.046350
## 39      A.mabeei - A.mavortium = 0      2.32      0.02033 0.032960
## 40      A.texanum - A.annulatum = 0      2.612      0.009011 0.021400
## 41      A.texanum - A.tigrinum = 0      3.725      0.0001954 0.002126
## 42      A.texanum - A.mavortium = 0      3.555      0.000378 0.002126
## 43      A.annulatum - A.tigrinum = 0      2.562      0.01041 0.023230
## 44      A.annulatum - A.mavortium = 0      2.355      0.01852 0.032800
## 45      A.tigrinum - A.mavortium = 0      1.206      0.2279 0.284900
```

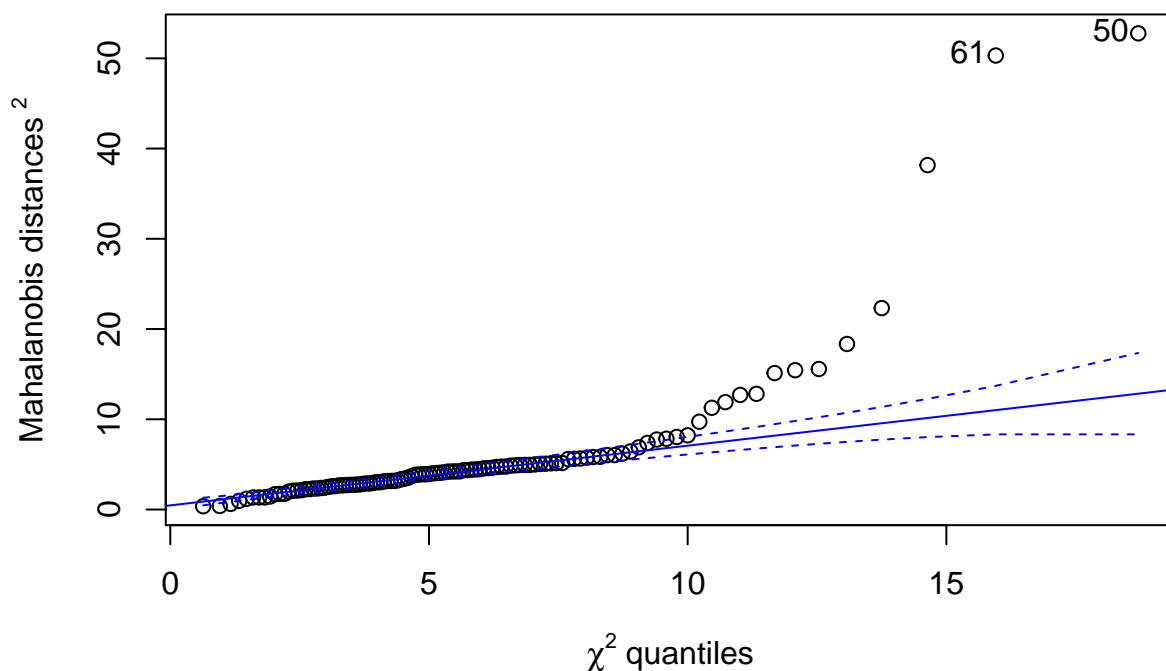
```
pairwise.perm.t.test(Atlas_wofossil_Tub_only$tub_interglac_extension, Atlas_wofossil_Tub_only$species,
  nperm = 999, progress = FALSE)
```

```
##
## Pairwise comparisons using permutation t tests
##
## data: Atlas_wofossil_Tub_only$tub_interglac_extension and Atlas_wofossil_Tub_only$species
## 999 permutations
##
##      A.gracile A.maculatum A.macrodactylum A.opacum
## A.maculatum    0.006      -              -              -
## A.macrodactylum 0.028    0.006      -              -
## A.opacum        0.028    0.165    0.006      -
## A.jeffersonianum 0.387    0.016    0.006    0.130
## A.mabeei        0.659    0.006    0.016    0.016
## A.texanum       0.016    0.486    0.006    0.421
## A.annulatum     0.745    0.006    0.006    0.041
## A.tigrinum      0.039    0.006    0.976    0.006
## A.mavortium     0.031    0.006    0.346    0.006
##
##      A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.maculatum          -          -          -          -
## A.macrodactylum     -          -          -          -
## A.opacum              -          -          -          -
## A.jeffersonianum     -          -          -          -
## A.mabeei             0.098      -          -          -
## A.texanum            0.029      0.006      -          -
## A.annulatum          0.346      0.346    0.016      -
## A.tigrinum           0.016      0.051    0.006    0.020
## A.mavortium          0.016      0.029    0.006    0.026    0.346
##
## P value adjustment method: fdr
```

DFA

```
# PROBLEM: Check Multivariate normality: FAIL
mqnorm(Atlas_wofossil_noTub_sub[, 1:6], main = "Multi-normal Q-Q Plot")
```

Multi-normal Q-Q Plot



```
## [1] 50 61
```

```
# DFA# With MASS
```

```
library(MASS)
```

```
##
```

```
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:rstatix':
```

```
##
```

```
## select
```

```
## The following object is masked from 'package:dplyr':
```

```
##
```

```
## select
```

```
AtlasLDA <- lda(species ~ M1 + M2 + M3 + M4 + M5 + M6, data = Atlas_wofossil_noTub_sub,
  CV = FALSE) #DFA no jackknife
```

```
AtlasLDA
```

```
## Call:
```

```
## lda(species ~ M1 + M2 + M3 + M4 + M5 + M6, data = Atlas_wofossil_noTub_sub,
```

```
## CV = FALSE)
```

```
##
```

```
## Prior probabilities of groups:
```

## A.gracile	A.maculatum	A.macrodactylum	A.opacum
## 0.07207207	0.12612613	0.10810811	0.11711712
## A.jeffersonianum	A.mabeei	A.texanum	A.annulatum
## 0.09909910	0.08108108	0.08108108	0.05405405
## A.tigrinum	A.mavortium		
## 0.15315315	0.10810811		

```
##
```

```

## Group means:
##           M1           M2           M3           M4           M5           M6
## A.gracile      3.301250  3.227500  1.1475000  2.425000  3.305000  1.1750000
## A.maculatum    3.542143  3.787143  1.4935714  2.729286  3.442857  1.4521429
## A.macrodactylum 2.200833  2.234167  0.8316667  1.655000  2.088333  0.8066667
## A.opacum       2.709231  2.836154  1.0546154  1.982308  2.799231  1.0130769
## A.jeffersonianum 3.144545  3.211818  1.1481818  2.434545  2.997273  1.1618182
## A.mabeei       2.218889  2.440000  0.9411111  1.713333  2.268889  0.9200000
## A.texanum      3.011111  3.490000  1.3822222  2.382222  2.864444  1.3155556
## A.annulatum    3.305000  3.373333  1.2583333  2.400000  2.808333  1.2500000
## A.tigrinum     3.534118  3.487059  1.3176471  2.744706  3.968235  1.4082353
## A.mavortium    4.222500  4.206667  1.5216667  3.038333  4.310833  1.7458333
##
## Coefficients of linear discriminants:
##           LD1           LD2           LD3           LD4           LD5           LD6
## M1  0.55663683 -0.8606379 -5.5795956 -0.3483209  2.989576  1.097501598
## M2  0.86359598 -2.5415371  1.9312656 -3.1297756 -4.765508  0.001897505
## M3  4.33837399 -3.6814217  4.6634719  2.8925828  6.711399  2.898321950
## M4  3.68648363  0.3634334 -0.8034773  5.2104307 -1.277601 -1.351860806
## M5 -7.27345664  2.1171683  2.6638922 -0.7621234 -0.491493  0.461170573
## M6 -0.01259303  0.2267898 -0.7343062 -1.3772889  1.482124 -3.660023730
##
## Proportion of trace:
##           LD1           LD2           LD3           LD4           LD5           LD6
## 0.6619 0.2516 0.0620 0.0176 0.0047 0.0023

AtlasLDA_jack <- lda(species ~ M1 + M2 + M3 + M4 + M5 + M6, data = Atlas_wofossil_noTub_sub,
  CV = TRUE) #DFA with jackknife
# Assess the accuracy of jackknife #
accAtlasLDA <- table(Atlas_wofossil_noTub_sub$species, AtlasLDA_jack$class)
accAtlasLDA

##
##           A.gracile A.maculatum A.macrodactylum A.opacum
## A.gracile           3           0           0           2
## A.maculatum          0          12           0           0
## A.macrodactylum     0           0           9           0
## A.opacum              0           1           0           9
## A.jeffersonianum      1           2           0           2
## A.mabeei              0           0           3           1
## A.texanum             0           2           0           0
## A.annulatum           0           0           0           0
## A.tigrinum            2           0           0           0
## A.mavortium           1           2           0           1
##
##           A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.gracile                 2           0           0           0           0
## A.maculatum                0           0           2           0           0
## A.macrodactylum           0           3           0           0           0
## A.opacum                   2           1           0           0           0
## A.jeffersonianum           5           1           0           0           0
## A.mabeei                   0           5           0           0           0
## A.texanum                  0           0           7           0           0
## A.annulatum                0           0           1           5           0
## A.tigrinum                 0           0           0           0          14

```



```
## A.mavortium          0          0          0          0          4
##
##           A.mavortium
## A.gracile          1
## A.maculatum        0
## A.macrodactylum   0
## A.opacum            0
## A.jeffersonianum    0
## A.mabeei            0
## A.texanum           0
## A.annulatum         0
## A.tigrinum          1
## A.mavortium         4
```

```
diag(prop.table(accAtlasLDA, 1))
```

```
##           A.gracile      A.maculatum  A.macrodactylum      A.opacum
##           0.3750000      0.8571429      0.7500000      0.6923077
## A.jeffersonianum      A.mabeei      A.texanum      A.annulatum
##           0.4545455      0.5555556      0.7777778      0.8333333
##           A.tigrinum      A.mavortium
##           0.8235294      0.3333333
```

```
sum(accAtlasLDA[row(accAtlasLDA) == col(accAtlasLDA)])/sum(accAtlasLDA)
```

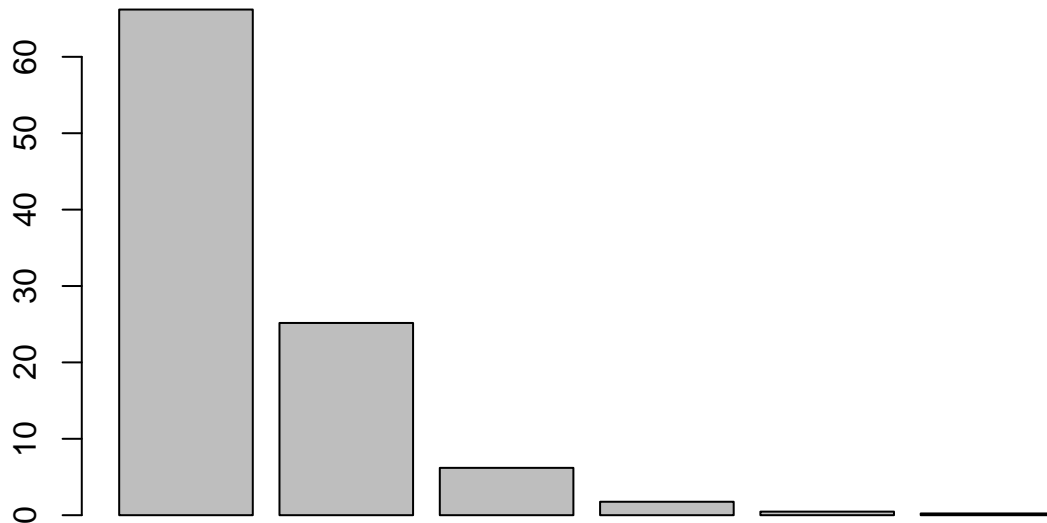
```
## [1] 0.6576577
```

```
# DFA# With MORPHO
```

```
library(Morpho)
```

```
Atlascva = CVA(Atlas_wofossil_noTub_sub[, 1:6], groups = Atlas_wofossil_noTub_sub$species,
  rounds = 10000, cv = TRUE)
```

```
barplot(Atlascva$Var[, 2]) # Variance explained by the canonical roots
```



```
# get the typicality probabilities and resulting classifications all
# specimens with a probability of < 0.01 as outliers (assigned to no class)
typprobs <- typprobClass(Atlascva$CVscores, groups = Atlas_wofossil_noTub_sub$species,
  outlier = 0.01, cv = TRUE)
print(typprobs)
```

```
## cross-validated classification results in frequencies
```

```

##
##      A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.gracile      0      3      2      0
## A.maculatum    0      0      0      0
## A.macrodactylum 0      0      0      3
## A.opacum       0      0      2      1
## A.jeffersonianum 0      2      6      1
## A.mabeei       0      0      0      7
## A.texanum      1      0      0      0
## A.annulatum    5      0      0      0
## A.tigrinum     0      2      0      0
## A.mavortium    0      1      0      0
##
##      A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.gracile      0      0      1      2      0
## A.maculatum    0      11     0      0      2
## A.macrodactylum 9      0      0      0      0
## A.opacum       0      1      0      9      0
## A.jeffersonianum 0      1      0      0      1
## A.mabeei       2      0      0      0      0
## A.texanum      0      1      0      0      7
## A.annulatum    0      0      0      0      1
## A.tigrinum     0      0      1      0      0
## A.mavortium    0      1      4      1      0
##
##      A.tigrinum none
## A.gracile      0      0
## A.maculatum    0      1
## A.macrodactylum 0      0
## A.opacum       0      0
## A.jeffersonianum 0      0
## A.mabeei       0      0
## A.texanum      0      0
## A.annulatum    0      0
## A.tigrinum     12     2
## A.mavortium    3      2
##
## cross-validated classification result in %
##
##      A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.gracile      0.0000 37.5000      25.0000  0.0000
## A.maculatum    0.0000  0.0000      0.0000  0.0000
## A.macrodactylum 0.0000  0.0000      0.0000 25.0000
## A.opacum       0.0000  0.0000     15.3846  7.6923
## A.jeffersonianum 0.0000 18.1818     54.5455  9.0909
## A.mabeei       0.0000  0.0000      0.0000 77.7778
## A.texanum      11.1111  0.0000      0.0000  0.0000
## A.annulatum    83.3333  0.0000      0.0000  0.0000
## A.tigrinum     0.0000 11.7647      0.0000  0.0000
## A.mavortium    0.0000  8.3333      0.0000  0.0000
##
##      A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.gracile      0.0000  0.0000     12.5000 25.0000  0.0000

```

```
## A.maculatum 0.0000 78.5714 0.0000 0.0000 14.2857
## A.macrodactylum 75.0000 0.0000 0.0000 0.0000 0.0000
## A.opacum 0.0000 7.6923 0.0000 69.2308 0.0000
## A.jeffersonianum 0.0000 9.0909 0.0000 0.0000 9.0909
## A.mabeei 22.2222 0.0000 0.0000 0.0000 0.0000
## A.texanum 0.0000 11.1111 0.0000 0.0000 77.7778
## A.annulatum 0.0000 0.0000 0.0000 0.0000 16.6667
## A.tigrinum 0.0000 0.0000 5.8824 0.0000 0.0000
## A.mavortium 0.0000 8.3333 33.3333 8.3333 0.0000
##
## A.tigrinum none
## A.gracile 0.0000 0.0000
## A.maculatum 0.0000 7.1429
## A.macrodactylum 0.0000 0.0000
## A.opacum 0.0000 0.0000
## A.jeffersonianum 0.0000 0.0000
## A.mabeei 0.0000 0.0000
## A.texanum 0.0000 0.0000
## A.annulatum 0.0000 0.0000
## A.tigrinum 70.5882 11.7647
## A.mavortium 25.0000 16.6667
##
##
## overall classification accuracy: 65.76577 %
```

```
# Assess the accuracy of jackknife #
accJack <- table(Atlascva$groups, Atlascva$class)
accJack
```

```
##
## A.gracile A.maculatum A.macrodactylum A.opacum
## A.gracile 3 0 0 2
## A.maculatum 0 12 0 0
## A.macrodactylum 0 0 9 0
## A.opacum 0 1 0 9
## A.jeffersonianum 1 2 0 2
## A.mabeei 0 0 3 1
## A.texanum 0 2 0 0
## A.annulatum 0 0 0 0
## A.tigrinum 2 0 0 0
## A.mavortium 1 2 0 1
##
## A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.gracile 2 0 0 0 0
## A.maculatum 0 0 2 0 0
## A.macrodactylum 0 3 0 0 0
## A.opacum 2 1 0 0 0
## A.jeffersonianum 5 1 0 0 0
## A.mabeei 0 5 0 0 0
## A.texanum 0 0 7 0 0
## A.annulatum 0 0 1 5 0
## A.tigrinum 0 0 0 0 14
## A.mavortium 0 0 0 0 4
##
## A.mavortium
```

```
## A.gracile 1
## A.maculatum 0
## A.macrodactylum 0
## A.opacum 0
## A.jeffersonianum 0
## A.mabeei 0
## A.texanum 0
## A.annulatum 0
## A.tigrinum 1
## A.mavortium 4
```

```
diag(prop.table(accJack, 1))
```

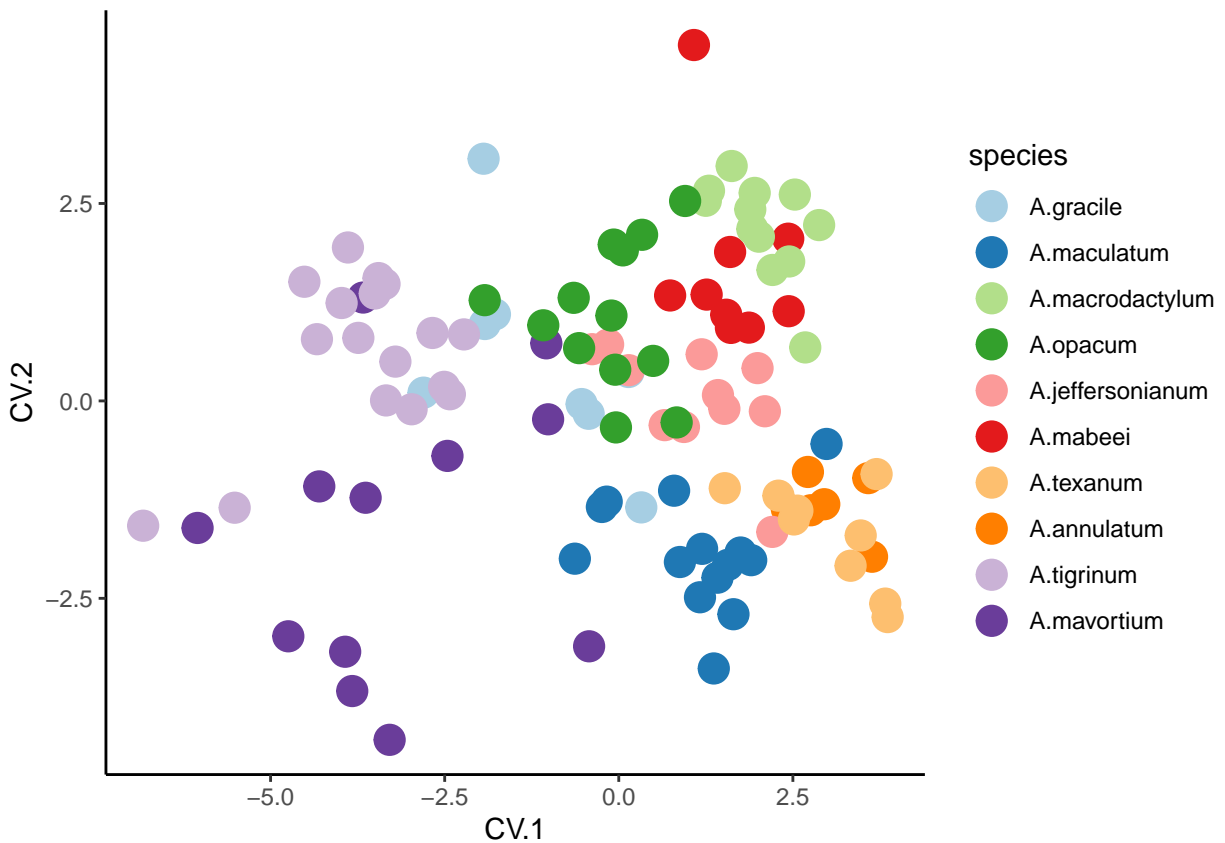
```
## A.gracile A.maculatum A.macrodactylum A.opacum
## 0.3750000 0.8571429 0.7500000 0.6923077
## A.jeffersonianum A.mabeei A.texanum A.annulatum
## 0.4545455 0.5555556 0.7777778 0.8333333
## A.tigrinum A.mavortium
## 0.8235294 0.3333333
```

```
sum(accJack[row(accJack) == col(accJack)])/sum(accJack)
```

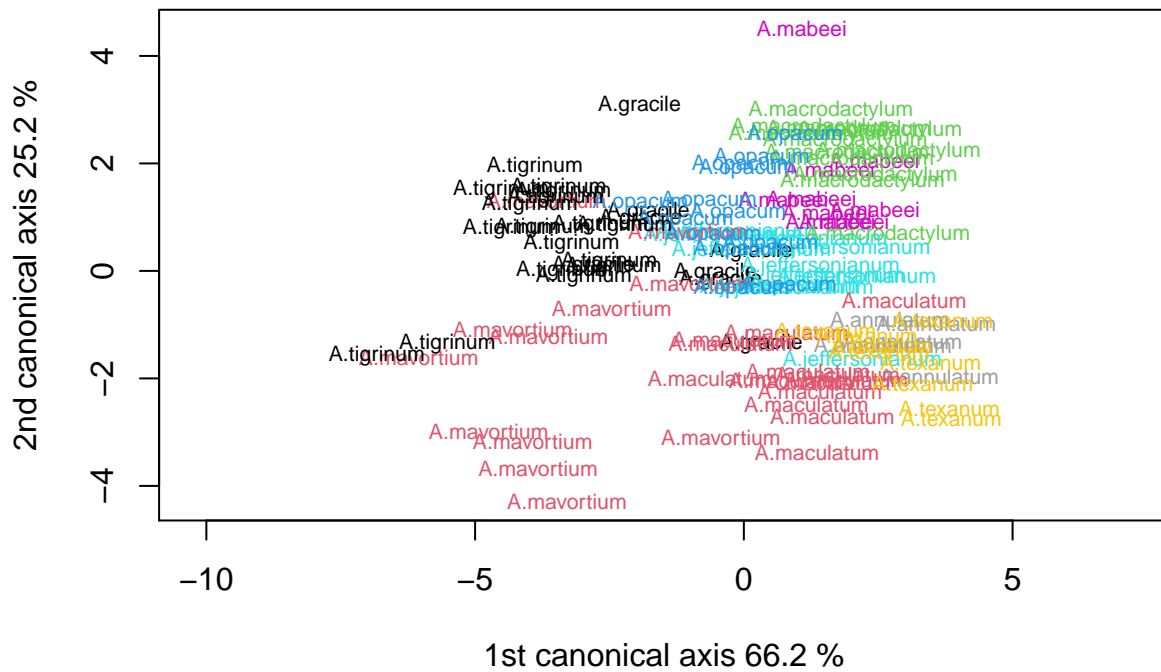
```
## [1] 0.6576577
```

Plot first two DF axes

```
AT_cva <- data.frame(AtlasCva$CVscores, species = AtlasCva$groups)
ggplot(AT_cva, aes(CV.1, CV.2)) + geom_point(size = 5, aes(color = species)) +
  theme_classic() + scale_color_brewer(palette = "Paired")
```

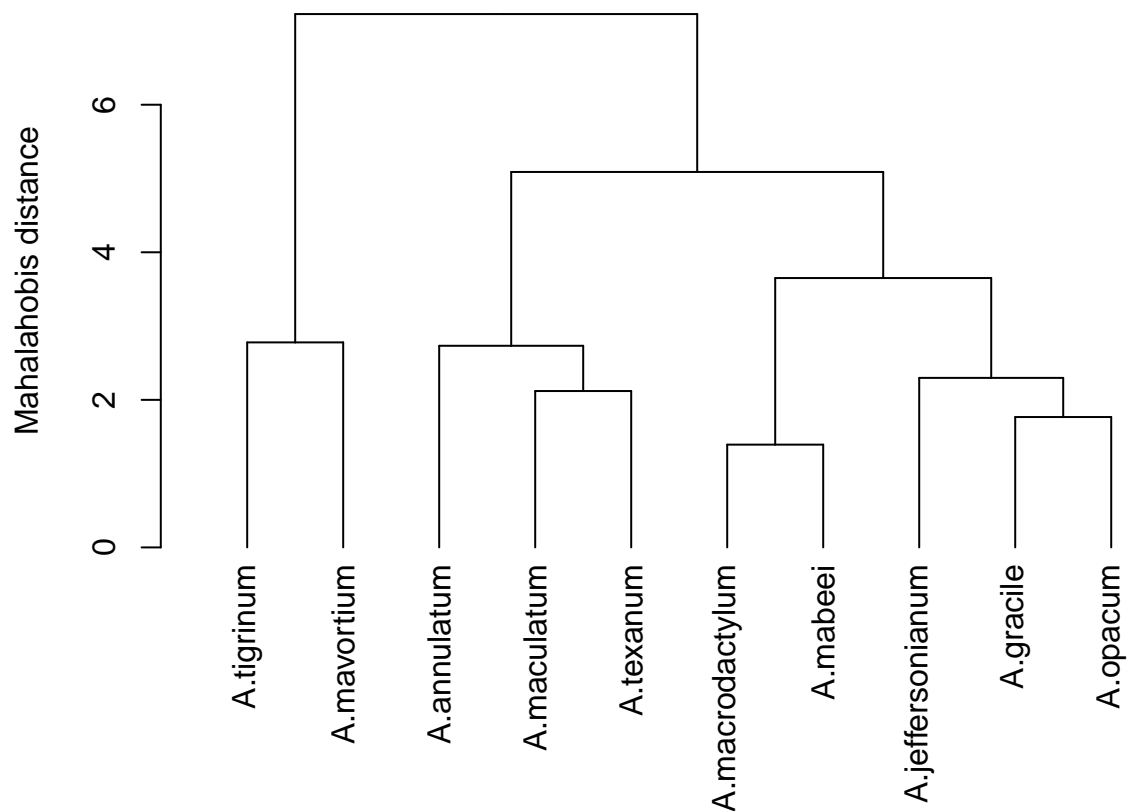


```
# alternative plot
plot(Atlascva$CVscores, col = Atlas_wofossil_noTub_sub$species, pch = as.numeric(Atlas_wofossil_noTub_sub$species),
     typ = "n", asp = 1, xlab = paste("1st canonical axis", paste(round(Atlascva$Var[1,
2], 1), "%")), ylab = paste("2nd canonical axis", paste(round(Atlascva$Var[2,
2], 1), "%")))
text(Atlascva$CVscores, as.character(Atlas_wofossil_noTub_sub$species), col = as.numeric(Atlas_wofossil_noTub_sub$species),
     cex = 0.7)
```



Plot Mahalanobis distances as dendrogram

```
dendroS = hclust(Atlascva$Dist$GroupdistMaha)
dendroS$labels = levels(Atlas_wofossil_noTub_sub$species)
par(mar = c(6.5, 4.5, 1, 1))
dendroS = as.dendrogram(dendroS)
plot(dendroS, main = "", sub = "", xlab = "", ylab = "Mahalanobis distance")
```



Random Forest ###:Non-parametric

```
library(randomForest)

## randomForest 4.6-14
## Type rfNews() to see new features/changes/bug fixes.
##
## Attaching package: 'randomForest'
## The following object is masked from 'package:gridExtra':
##
##   combine
## The following object is masked from 'package:psych':
##
##   outlier
## The following object is masked from 'package:ggplot2':
##
##   margin
## The following object is masked from 'package:dplyr':
##
##   combine
Atlas.rf <- randomForest(species ~ M1 + M2 + M3 + M4 + M5 + M6, data = Atlas_wofossil_noTub_sub,
  importance = TRUE, proximity = TRUE)
```

```
print(Atlas.rf)
```

```
##
## Call:
## randomForest(formula = species ~ M1 + M2 + M3 + M4 + M5 + M6,      data = Atlas_wofossil_noTub_sub,
##               Type of random forest: classification
##               Number of trees: 500
## No. of variables tried at each split: 2
##
## OOB estimate of error rate: 34.23%
## Confusion matrix:
##               A.gracile A.maculatum A.macrodactylum A.opacum
## A.gracile           0           1           0           0
## A.maculatum          0          12           0           1
## A.macrodactylum     0           0          10           0
## A.opacum              1           0           2           7
## A.jeffersonianum      1           0           0           2
## A.mabeei              0           0           3           1
## A.texanum             0           0           0           0
## A.annulatum           0           0           0           0
## A.tigrinum            0           1           0           0
## A.mavortium           0           0           0           1
##               A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.gracile                     2           1           0           0           3
## A.maculatum                    0           0           0           0           0
## A.macrodactylum              0           2           0           0           0
## A.opacum                       1           2           0           0           0
## A.jeffersonianum               6           0           0           2           0
## A.mabeei                       0           5           0           0           0
## A.texanum                      0           0           8           1           0
## A.annulatum                    1           0           1           4           0
## A.tigrinum                     0           0           0           0          14
## A.mavortium                    1           0           1           0           2
##               A.mavortium class.error
## A.gracile              1  1.0000000
## A.maculatum             1  0.1428571
## A.macrodactylum        0  0.1666667
## A.opacum                 0  0.4615385
## A.jeffersonianum         0  0.4545455
## A.mabeei                 0  0.4444444
## A.texanum                0  0.1111111
## A.annulatum              0  0.3333333
## A.tigrinum               2  0.1764706
## A.mavortium              7  0.4166667
```

```
rf_acc <- Atlas.rf$confusion
rf_acc <- 1 - rf_acc[, 11] # percent correct classification
rf_acc
```

```
##           A.gracile      A.maculatum A.macrodactylum      A.opacum
##           0.0000000      0.8571429      0.8333333      0.5384615
## A.jeffersonianum      A.mabeei      A.texanum      A.annulatum
##           0.5454545      0.5555556      0.8888889      0.6666667
##           A.tigrinum      A.mavortium
```



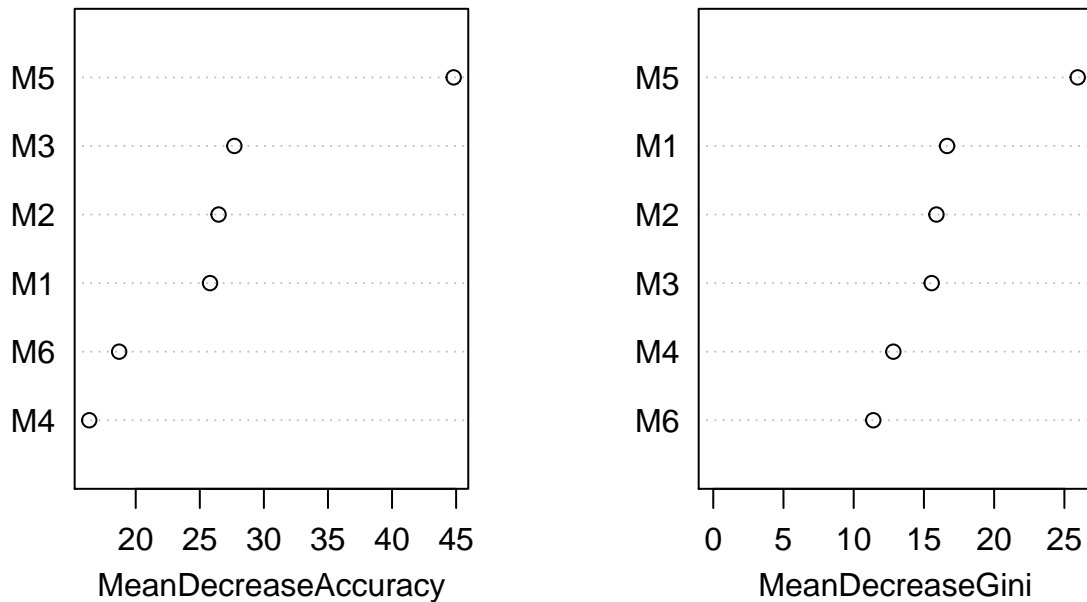
```
##          0.8235294          0.5833333
```

```
# Look at variable importance
round(importance(Atlas.rf), 2)
```

```
##      A.gracile A.maculatum A.macrodactylum A.opacum A.jeffersonianum A.mabeei
## M1      1.97      11.13          5.12      8.44          4.47      16.07
## M2     -0.87      18.45          18.73      7.38          0.02      6.89
## M3     -2.16      23.27          11.84     10.25          11.06     -0.09
## M4     -0.73      9.98           7.36      7.72          4.82      6.69
## M5      5.73      15.79          22.77     13.50          12.47     10.31
## M6     -2.74      6.69           11.86     10.87          2.02      2.89
##      A.texanum A.annulatum A.tigrinum A.mavortium MeanDecreaseAccuracy
## M1      10.64      3.55          6.12      18.74          25.80
## M2      8.84      3.57          1.05      15.18          26.46
## M3     16.00      9.56         -0.89      3.30          27.70
## M4      5.06      5.19          8.20      1.26          16.37
## M5     16.91     10.89         31.19     16.91          44.81
## M6      3.63      0.55          8.24     10.03          18.71
##      MeanDecreaseGini
## M1             16.65
## M2             15.89
## M3             15.55
## M4             12.82
## M5             25.95
## M6             11.39
```

```
varImpPlot(Atlas.rf)
```

Atlas.rf



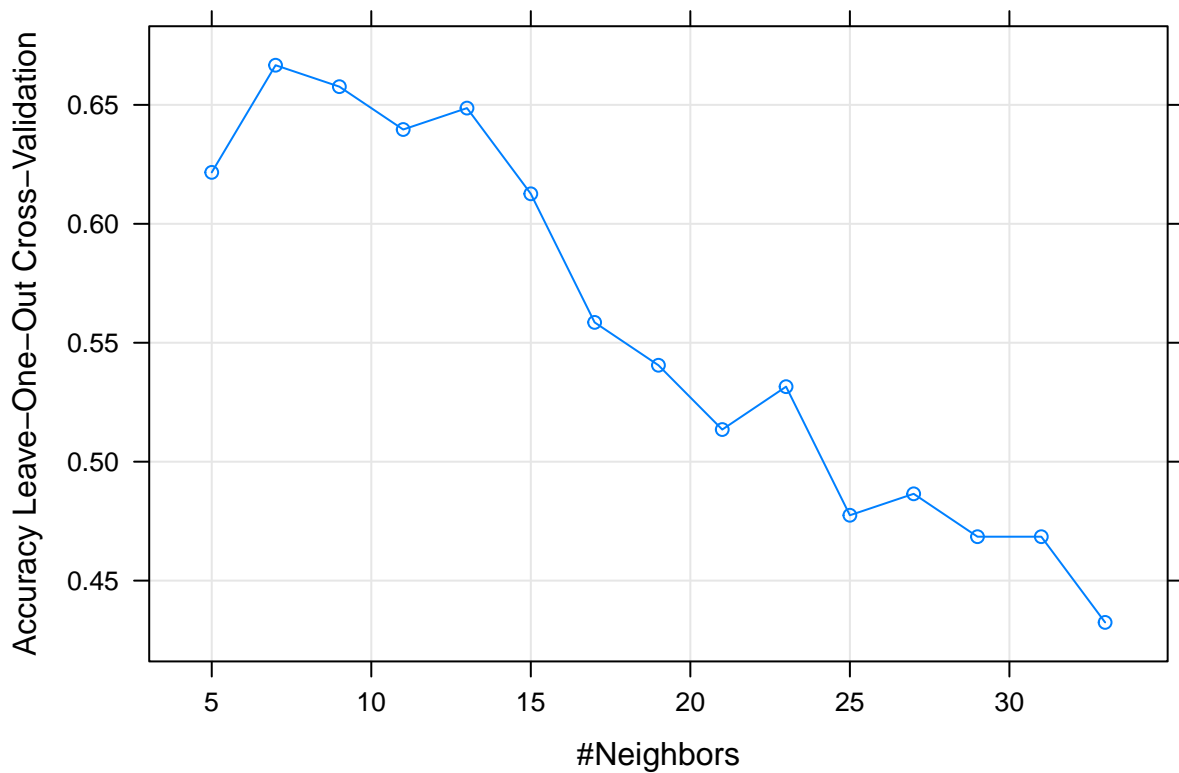
K Nearest neighbor ###Non-parametric

```
library(tidyverse)
library(caret)

##
## Attaching package: 'caret'
## The following object is masked from 'package:vegan':
##
##   tolerance
## The following object is masked from 'package:purrr':
##
##   lift
Atlas_wofossil_noTub_sub <- column_to_rownames(Atlas_wofossil_noTub_sub, var = "specimen_num")

#make KNN model using LOOCV to find optimal k
KNNmodel <- train(
  species ~., data = Atlas_wofossil_noTub_sub, method = "knn",
  trControl = trainControl("LOOCV", number = 1),
  preProcess = c("center"), #center the data
  tuneLength = 15)

plot(KNNmodel) # plot accuracy vs k
```



```
KNNmodel$bestTune # optimal k
```

```
## k
## 2 7
```

```
predicted.classes <- KNNmodel %>% predict(Atlas_wofossil_noTub_sub[,1:6]) # predict class based on KNN
head(predicted.classes)
```

```
## [1] A.annulatum      A.jeffersonianum A.annulatum      A.annulatum
## [5] A.texanum          A.texanum
## 10 Levels: A.gracile A.maculatum A.macrodactylum A.opacum ... A.mavortium
```

```
mean(predicted.classes == Atlas_wofossil_noTub_sub$species) #overall accuracy
```

```
## [1] 0.7477477
```

```
# assess accuracy per species
```

```
accKNN <- table(Atlas_wofossil_noTub_sub$species,predicted.classes)
accKNN
```

```
##                predicted.classes
##                A.gracile A.maculatum A.macrodactylum A.opacum
## A.gracile                1            1                0            0
## A.maculatum              0            13               0            0
## A.macrodactylum        0            0               11            0
## A.opacum                 0            1                2            8
## A.jeffersonianum        0            1                0            3
## A.mabeei                0            0                1            0
## A.texanum               0            0                0            0
## A.annulatum             0            0                0            0
## A.tigrinum              0            0                0            0
## A.mavortium             0            1                0            1
##                predicted.classes
##                A.jeffersonianum A.mabeei A.texanum A.annulatum A.tigrinum
## A.gracile                    2            1            0            0            3
## A.maculatum                  1            0            0            0            0
## A.macrodactylum            0            1            0            0            0
## A.opacum                     1            1            0            0            0
## A.jeffersonianum            7            0            0            0            0
## A.mabeei                    0            8            0            0            0
## A.texanum                   0            0            9            0            0
## A.annulatum                 1            0            2            3            0
## A.tigrinum                  0            0            0            0           15
## A.mavortium                 1            0            0            0            1
##                predicted.classes
##                A.mavortium
## A.gracile                0
## A.maculatum              0
## A.macrodactylum        0
## A.opacum                 0
## A.jeffersonianum        0
## A.mabeei                0
## A.texanum               0
## A.annulatum             0
## A.tigrinum              2
## A.mavortium             8
```

```
diag(prop.table(accKNN, 1))
```

```
##          A.gracile      A.maculatum A.macrodactylum      A.opacum
##          0.1250000      0.9285714      0.9166667      0.6153846
```

## A.jeffersonianum	A.mabeei	A.texanum	A.annulatum
## 0.6363636	0.8888889	1.0000000	0.5000000
## A.tigrinum	A.mavortium		
## 0.8823529	0.6666667		

Model selection (multinomial regression)

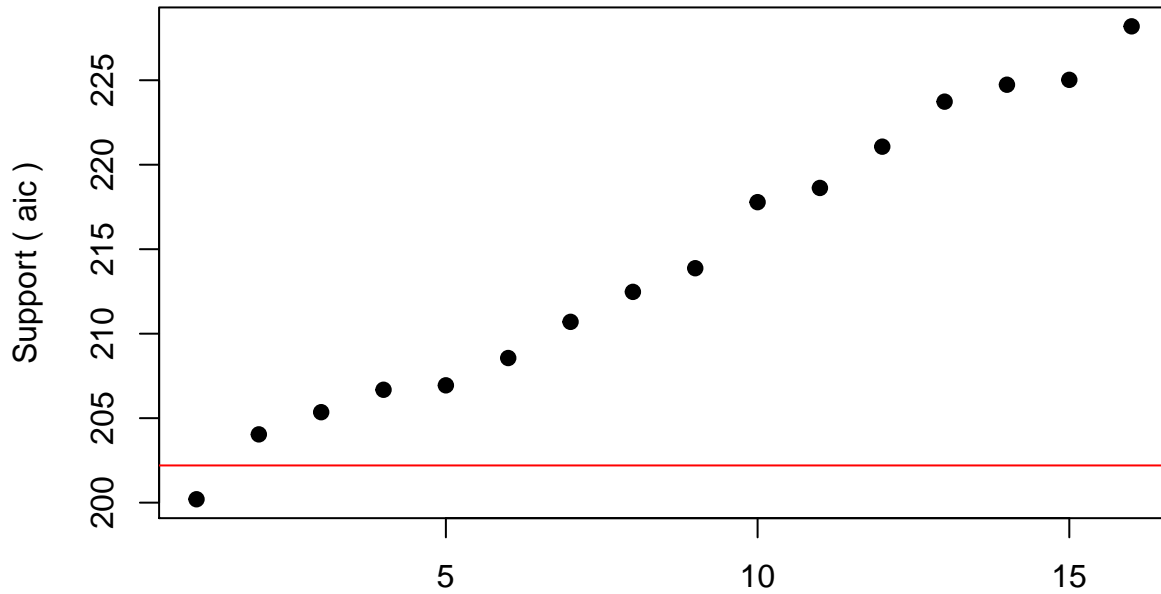
```
library(glmulti)
library(nnet)
multinom.glmulti <- function(formula, data, ...) multinom(formula, data, ...)
res <- glmulti(species ~ ., level = 1, data = Atlas_wofossil_noTub_sub, report = FALSE,
  plotty = FALSE, fitfunction = multinom.glmulti, method = "h", crit = "aic",
  confsetsize = 16)

print(res)

## glmulti.analysis
## Method: h / Fitting: multinom.glmulti / IC used: aic
## Level: 1 / Marginality: FALSE
## From 16 models:
## Best IC: 200.201066755855
## Best model:
## [1] "species ~ 1 + M1 + M3 + M4 + M5 + M6"
## Evidence weight: 0.757413727546765
## Worst IC: 228.189504107244
## 1 models within 2 IC units.
## 3 models to reach 95% of evidence weight.

plot(res)
```

IC profile



Best models

```
top <- weightable(res)
top
```

##	model	aic	weights
## 1	species ~ 1 + M1 + M3 + M4 + M5 + M6	200.2011	7.574137e-01
## 2	species ~ 1 + M1 + M3 + M4 + M5	204.0410	1.110440e-01
## 3	species ~ 1 + M1 + M2 + M3 + M4 + M5 + M6	205.3535	5.760868e-02
## 4	species ~ 1 + M1 + M2 + M3 + M4 + M5	206.6808	2.966786e-02
## 5	species ~ 1 + M2 + M3 + M4 + M5 + M6	206.9433	2.601858e-02
## 6	species ~ 1 + M3 + M4 + M5 + M6	208.5596	1.159589e-02
## 7	species ~ 1 + M2 + M3 + M4 + M5	210.7009	3.974899e-03
## 8	species ~ 1 + M1 + M2 + M3 + M5	212.4752	1.636927e-03
## 9	species ~ 1 + M1 + M2 + M3 + M5 + M6	213.8743	8.132637e-04
## 10	species ~ 1 + M3 + M4 + M5	217.7840	1.151447e-04
## 11	species ~ 1 + M2 + M3 + M5 + M6	218.6262	7.557261e-05
## 12	species ~ 1 + M2 + M3 + M5	221.0684	2.228732e-05
## 13	species ~ 1 + M1 + M3 + M5	223.7328	5.881411e-06
## 14	species ~ 1 + M1 + M2 + M4 + M5 + M6	224.7346	3.564098e-06
## 15	species ~ 1 + M1 + M3 + M5 + M6	225.0239	3.084009e-06
## 16	species ~ 1 + M1 + M2 + M4 + M5	228.1895	6.334630e-07