

# Ambystoma\_GMM

David Ledesma

3/23/2021

## Load in data

```
# Generalized procrustes analysis
library(geomorph)

## Loading required package: RRPP
## Loading required package: rgl
GPA_landmarks_sub <- gpagen(GMM_data_sub$land)

##      |
# Create geomorph data frame
Amb_gdf <- geomorph.data.frame(coords = GPA_landmarks_sub$coords, size = GPA_landmarks_sub$Csize,
                                species = GMM_data_sub$species)
```

## PCA

```
Amb_PCA <- gm.prcomp(GPA_landmarks_sub$coords)
```

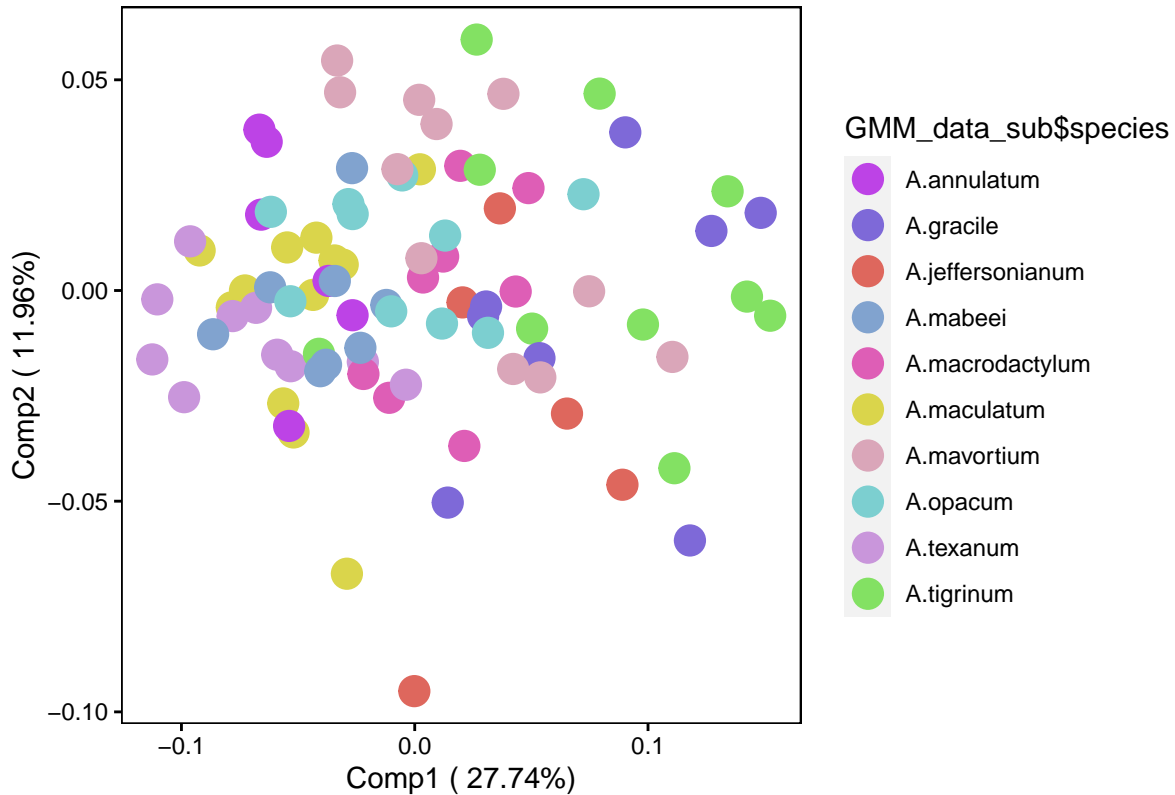
## PCA vizualization

```
PC_scores <- as.data.frame(Amb_PCA$x)
library(ggplot2)
library(grid)
library(gridExtra)
theme <- theme(panel.background = element_blank(), panel.border = element_rect(fill = NA),
               panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
               strip.background = element_blank(), axis.text.x = element_text(colour = "black"),
               axis.text.y = element_text(colour = "black"), axis.ticks = element_line(colour = "black"),
               plot.margin = unit(c(1, 1, 1, 1), "line"))
percentage <- round(Amb_PCA$sdev/sum(Amb_PCA$sdev) * 100, 2)
percentage <- paste(colnames(PC_scores), "(", paste(as.character(percentage),
            "%", ") ", sep = ""))
# library(RColorBrewer) getPalette = colorRampPalette(brewer.pal(12,
# 'Paired')) my.colors=rainbow(28) #set up color palette of rainbow colors
# with n = 14 plot(1:28, pch=19, cex=2, col=my.colors)
library(randomcolorR)
n <- 14
```

```

palette <- distinctColorPalette(n)
p <- ggplot(PC_scores, aes(x = Comp1, y = Comp2, color = GMM_data_sub$species))
p <- p + geom_point(size = 5) + theme + xlab(percentage[1]) + ylab(percentage[2]) +
  scale_color_manual(values = palette)
# p + stat_ellipse()
p

```



## ANOVA

```

# Without size
Amb_anova <- procD.lm(coords ~ species, data = Amb_gdf, iter = 999, RRPP = TRUE,
  print.progress = FALSE)
Amb_anova$aov.table

```

```

##           Df      SS      MS      Rsq      F      Z Pr(>F)
## species     9 0.28885 0.032095 0.46656 7.58 6.1822 0.001 **
## Residuals  78 0.33026 0.004234 0.53344
## Total      87 0.61911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

# With size
Amb_anova_size <- procD.lm(coords ~ species * size, data = Amb_gdf, iter = 999,
  RRPP = TRUE, print.progress = FALSE)
Amb_anova_size$aov.table

```

```

##           Df      SS      MS      Rsq      F      Z Pr(>F)
## species     9 0.28885 0.032095 0.46656 9.6972 6.7336 0.001 **

```

```
## size          1 0.05555 0.055553 0.08973 16.7849 4.6233 0.001 **
## species:size  9 0.04965 0.005517 0.08020  1.6668 2.7247 0.003 **
## Residuals     68 0.22506 0.003310 0.36352
## Total         87 0.61911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Post-hoc comparisons

```
gp <- interaction(Amb_gdf$species)
PW <- pairwise(Amb_anova, groups = gp, covariate = NULL)
summary(PW, test.type = "dist", confidence = 0.95, stat.table = TRUE)
```

```
##
## Pairwise comparisons
##
## Groups: A.annulatum A.gracile A.jeffersonianum A.mabeei A.macrodactylum A.maculatum A.mavortium A.opacum
##
## RRPP: 1000 permutations
##
## LS means:
## Vectors hidden (use show.vectors = TRUE to view)
##
## Pairwise distances between means, plus statistics
##
```

	d	UCL (95%)	Z	Pr > d
A.annulatum:A.gracile	0.13239614	0.07236325	3.12627071	0.001
A.annulatum:A.jeffersonianum	0.10926004	0.08122529	2.37627310	0.005
A.annulatum:A.mabeei	0.02404080	0.07150346	-1.45726338	0.932
A.annulatum:A.macrodactylum	0.06980446	0.07447355	1.48233490	0.077
A.annulatum:A.maculatum	0.03016095	0.06814923	-0.56429680	0.715
A.annulatum:A.mavortium	0.08066860	0.06886159	2.09737455	0.016
A.annulatum:A.opacum	0.06015993	0.07188711	1.23344009	0.111
A.annulatum:A.texanum	0.04084240	0.07001921	0.21274860	0.406
A.annulatum:A.tigrinum	0.13214751	0.07005948	3.21514964	0.001
A.gracile:A.jeffersonianum	0.04782787	0.08201598	0.39222275	0.340
A.gracile:A.mabeei	0.11979624	0.06850240	3.01992267	0.001
A.gracile:A.macrodactylum	0.06865695	0.06821479	1.67937188	0.048
A.gracile:A.maculatum	0.12605456	0.06442296	3.31277418	0.001
A.gracile:A.mavortium	0.07275562	0.06496100	1.96136696	0.028
A.gracile:A.opacum	0.08586311	0.06799658	2.29026370	0.011
A.gracile:A.texanum	0.14956254	0.06686976	3.70203029	0.001
A.gracile:A.tigrinum	0.03850558	0.06510591	0.25853742	0.386
A.jeffersonianum:A.mabeei	0.09683687	0.07838542	2.30017325	0.006
A.jeffersonianum:A.macrodactylum	0.05873597	0.07826468	0.97906457	0.173
A.jeffersonianum:A.maculatum	0.09775414	0.07287364	2.43877077	0.006
A.jeffersonianum:A.mavortium	0.06924488	0.07325083	1.55858947	0.073
A.jeffersonianum:A.opacum	0.06621109	0.07362263	1.39379349	0.085
A.jeffersonianum:A.texanum	0.12187821	0.07595853	2.90765561	0.001
A.jeffersonianum:A.tigrinum	0.06877163	0.07367072	1.47588783	0.074
A.mabeei:A.macrodactylum	0.05809304	0.06879943	1.20260983	0.129
A.mabeei:A.maculatum	0.03076645	0.06285151	-0.21393299	0.572
A.mabeei:A.mavortium	0.07799668	0.06381559	2.20384111	0.011
A.mabeei:A.opacum	0.05302748	0.06263803	1.15788616	0.129

```
## A.mabeei:A.texanum      0.04566914 0.06390798 0.79408209 0.223
## A.mabeei:A.tigrinum     0.12232671 0.06463118 3.34822467 0.001
## A.macrodactylum:A.maculatum 0.07207364 0.06373322 1.99879815 0.022
## A.macrodactylum:A.mavortium 0.03953065 0.06364719 0.42917495 0.330
## A.macrodactylum:A.opacum   0.04873358 0.06449754 0.90507767 0.192
## A.macrodactylum:A.texanum  0.09402646 0.06688709 2.55357809 0.002
## A.macrodactylum:A.tigrinum 0.07298925 0.06703932 1.93566053 0.027
## A.maculatum:A.mavortium    0.08602555 0.05817324 2.70884434 0.001
## A.maculatum:A.opacum      0.05125432 0.05902486 1.25584855 0.110
## A.maculatum:A.texanum     0.03278104 0.05993229 0.04753565 0.478
## A.maculatum:A.tigrinum     0.13091670 0.05970160 3.56966770 0.001
## A.mavortium:A.opacum      0.06020869 0.05994950 1.69268183 0.050
## A.mavortium:A.texanum     0.10658192 0.05930117 3.15414207 0.001
## A.mavortium:A.tigrinum     0.06199015 0.06033665 1.75148075 0.044
## A.opacum:A.texanum        0.07711487 0.06198869 2.24988240 0.010
## A.opacum:A.tigrinum       0.09480490 0.06247456 2.72307639 0.001
## A.texanum:A.tigrinum      0.15216326 0.06078045 4.15157469 0.001
```

```
summary(PW, test.type = "dist", confidence = 0.95, stat.table = FALSE)
```

```
##
## Pairwise comparisons
##
## Groups: A.annulatum A.gracile A.jeffersonianum A.mabeei A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum A.tigrinum
##
## RRPP: 1000 permutations
##
## LS means:
## Vectors hidden (use show.vectors = TRUE to view)
##
## Pairwise distances between means
##
##      A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.annulatum      0.00000000 0.13239614      0.10926004 0.02404080
## A.gracile        0.13239614 0.00000000      0.04782787 0.11979624
## A.jeffersonianum 0.10926004 0.04782787      0.00000000 0.09683687
## A.mabeei         0.02404080 0.11979624      0.09683687 0.00000000
## A.macrodactylum 0.06980446 0.06865695      0.05873597 0.05809304
## A.maculatum      0.03016095 0.12605456      0.09775414 0.03076645
## A.mavortium      0.08066860 0.07275562      0.06924488 0.07799668
## A.opacum         0.06015993 0.08586311      0.06621109 0.05302748
## A.texanum        0.04084240 0.14956254      0.12187821 0.04566914
## A.tigrinum       0.13214751 0.03850558      0.06877163 0.12232671
##
##      A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.annulatum      0.06980446 0.03016095 0.08066860 0.06015993 0.04084240
## A.gracile        0.06865695 0.12605456 0.07275562 0.08586311 0.14956254
## A.jeffersonianum 0.05873597 0.09775414 0.06924488 0.06621109 0.12187821
## A.mabeei         0.05809304 0.03076645 0.07799668 0.05302748 0.04566914
## A.macrodactylum 0.00000000 0.07207364 0.03953065 0.04873358 0.09402646
## A.maculatum      0.07207364 0.00000000 0.08602555 0.05125432 0.03278104
## A.mavortium      0.03953065 0.08602555 0.00000000 0.06020869 0.10658192
## A.opacum         0.04873358 0.05125432 0.06020869 0.00000000 0.07711487
## A.texanum        0.09402646 0.03278104 0.10658192 0.07711487 0.00000000
## A.tigrinum       0.07298925 0.13091670 0.06199015 0.09480490 0.15216326
##
##      A.tigrinum
## A.annulatum      0.13214751
```

```

## A.gracile      0.03850558
## A.jeffersonianum 0.06877163
## A.mabeei      0.12232671
## A.macrodactylum 0.07298925
## A.maculatum    0.13091670
## A.mavortium    0.06199015
## A.opacum       0.09480490
## A.texanum      0.15216326
## A.tigrinum     0.00000000
##
## Pairwise 95% Upper confidence limits between means
##
## A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.annulatum 0.00000000 0.07236325      0.08122529 0.07150346
## A.gracile    0.07236325 0.00000000      0.08201598 0.06850240
## A.jeffersonianum 0.08122529 0.08201598      0.00000000 0.07838542
## A.mabeei     0.07150346 0.06850240      0.07838542 0.00000000
## A.macrodactylum 0.07447355 0.06821479      0.07826468 0.06879943
## A.maculatum  0.06814923 0.06442296      0.07287364 0.06285151
## A.mavortium  0.06886159 0.06496100      0.07325083 0.06381559
## A.opacum     0.07188711 0.06799658      0.07362263 0.06263803
## A.texanum    0.07001921 0.06686976      0.07595853 0.06390798
## A.tigrinum   0.07005948 0.06510591      0.07367072 0.06463118
##
## A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.annulatum      0.07447355 0.06814923 0.06886159 0.07188711 0.07001921
## A.gracile        0.06821479 0.06442296 0.06496100 0.06799658 0.06686976
## A.jeffersonianum 0.07826468 0.07287364 0.07325083 0.07362263 0.07595853
## A.mabeei         0.06879943 0.06285151 0.06381559 0.06263803 0.06390798
## A.macrodactylum 0.00000000 0.06373322 0.06364719 0.06449754 0.06688709
## A.maculatum      0.06373322 0.00000000 0.05817324 0.05902486 0.05993229
## A.mavortium      0.06364719 0.05817324 0.00000000 0.05994950 0.05930117
## A.opacum         0.06449754 0.05902486 0.05994950 0.00000000 0.06198869
## A.texanum        0.06688709 0.05993229 0.05930117 0.06198869 0.00000000
## A.tigrinum       0.06703932 0.05970160 0.06033665 0.06247456 0.06078045
##
## A.tigrinum
## A.annulatum      0.07005948
## A.gracile        0.06510591
## A.jeffersonianum 0.07367072
## A.mabeei         0.06463118
## A.macrodactylum 0.06703932
## A.maculatum      0.05970160
## A.mavortium      0.06033665
## A.opacum         0.06247456
## A.texanum        0.06078045
## A.tigrinum       0.00000000
##
## Pairwise effect sizes (Z) between means
##
## A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.annulatum 0.0000000 3.1262707      2.3762731 -1.4572634
## A.gracile    3.1262707 0.0000000      0.3922227 3.0199227
## A.jeffersonianum 2.3762731 0.3922227      0.0000000 2.3001732
## A.mabeei     -1.4572634 3.0199227      2.3001732 0.0000000
## A.macrodactylum 1.4823349 1.6793719      0.9790646 1.2026098
## A.maculatum  -0.5642968 3.3127742      2.4387708 -0.2139330
## A.mavortium   2.0973745 1.9613670      1.5585895 2.2038411

```

```

## A.opacum          1.2334401 2.2902637          1.3937935 1.1578862
## A.texanum         0.2127486 3.7020303          2.9076556 0.7940821
## A.tigrinum        3.2151496 0.2585374          1.4758878 3.3482247
##
## A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.annulatum      1.4823349 -0.56429680      2.097375 1.2334401 0.21274860
## A.gracile         1.6793719 3.31277418      1.961367 2.2902637 3.70203029
## A.jeffersonianum 0.9790646 2.43877077      1.558589 1.3937935 2.90765561
## A.mabeei         1.2026098 -0.21393299      2.203841 1.1578862 0.79408209
## A.macrodactylum 0.0000000 1.99879815      0.429175 0.9050777 2.55357809
## A.maculatum       1.9987981 0.00000000      2.708844 1.2558485 0.04753565
## A.mavortium       0.4291750 2.70884434      0.000000 1.6926818 3.15414207
## A.opacum          0.9050777 1.25584855      1.692682 0.0000000 2.24988240
## A.texanum         2.5535781 0.04753565      3.154142 2.2498824 0.00000000
## A.tigrinum        1.9356605 3.56966770      1.751481 2.7230764 4.15157469
##
## A.tigrinum
## A.annulatum      3.2151496
## A.gracile         0.2585374
## A.jeffersonianum 1.4758878
## A.mabeei         3.3482247
## A.macrodactylum 1.9356605
## A.maculatum       3.5696677
## A.mavortium       1.7514808
## A.opacum          2.7230764
## A.texanum         4.1515747
## A.tigrinum        0.0000000
##
## Pairwise P-values between means
##
## A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.annulatum      1.000      0.001          0.005      0.932
## A.gracile         0.001      1.000          0.340      0.001
## A.jeffersonianum 0.005      0.340          1.000      0.006
## A.mabeei         0.932      0.001          0.006      1.000
## A.macrodactylum 0.077      0.048          0.173      0.129
## A.maculatum       0.715      0.001          0.006      0.572
## A.mavortium       0.016      0.028          0.073      0.011
## A.opacum          0.111      0.011          0.085      0.129
## A.texanum         0.406      0.001          0.001      0.223
## A.tigrinum        0.001      0.386          0.074      0.001
##
## A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.annulatum      0.077      0.715          0.016      0.111      0.406
## A.gracile         0.048      0.001          0.028      0.011      0.001
## A.jeffersonianum 0.173      0.006          0.073      0.085      0.001
## A.mabeei         0.129      0.572          0.011      0.129      0.223
## A.macrodactylum 1.000      0.022          0.330      0.192      0.002
## A.maculatum       0.022      1.000          0.001      0.110      0.478
## A.mavortium       0.330      0.001          1.000      0.050      0.001
## A.opacum          0.192      0.110          0.050      1.000      0.010
## A.texanum         0.002      0.478          0.001      0.010      1.000
## A.tigrinum        0.027      0.001          0.044      0.001      0.001
##
## A.tigrinum
## A.annulatum      0.001
## A.gracile         0.386
## A.jeffersonianum 0.074
## A.mabeei         0.001

```

```
## A.macrodactylum      0.027
## A.maculatum            0.001
## A.mavortium            0.044
## A.opacum               0.001
## A.texanum              0.001
## A.tigrinum             1.000
```

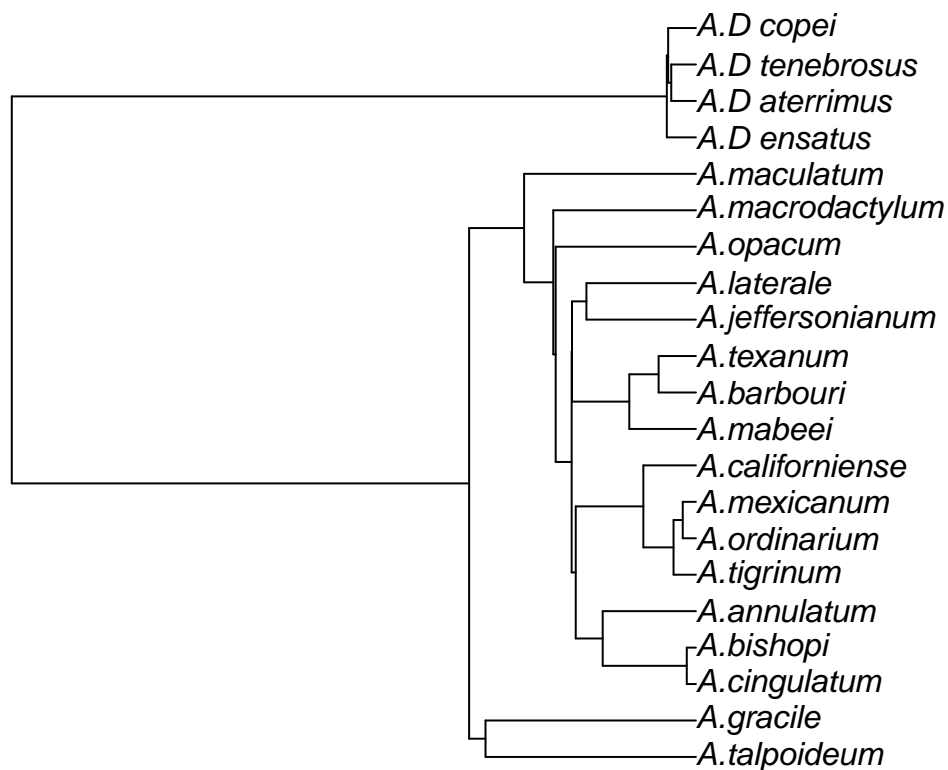
## Phylogenetic signal

### Load in data

```
require(phytools)

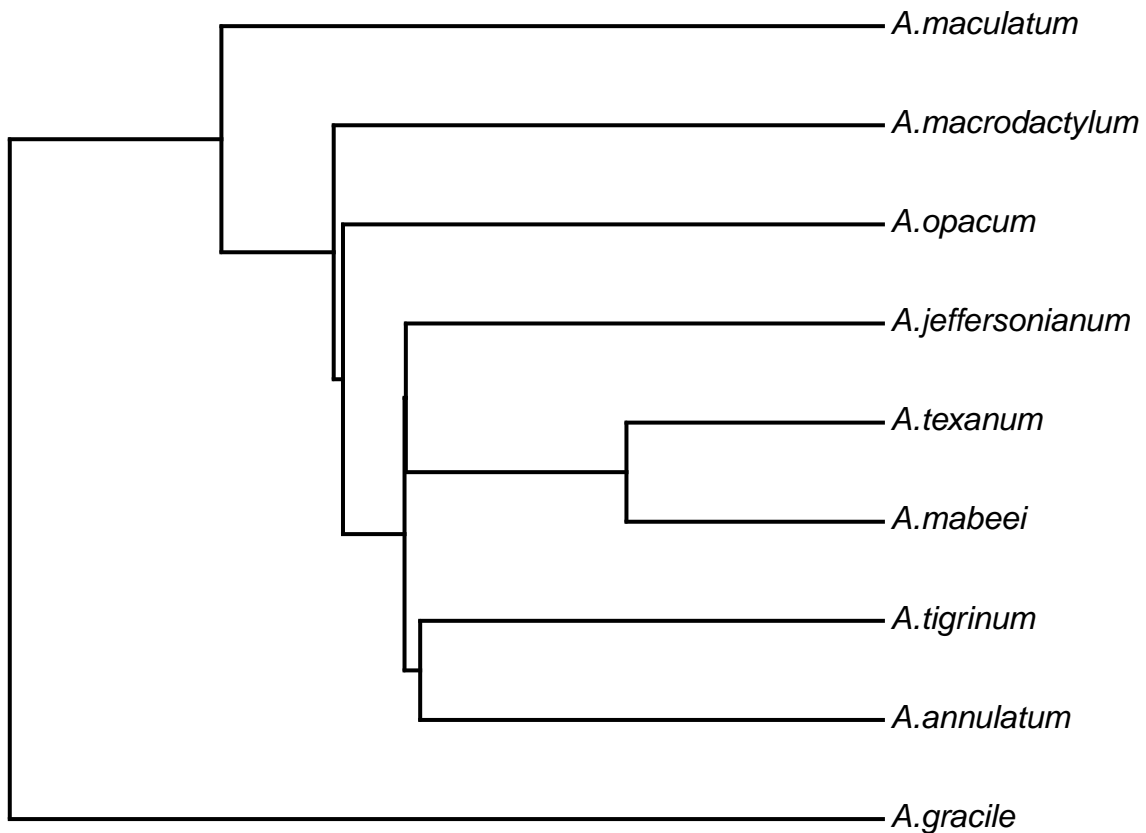
## Loading required package: phytools
## Loading required package: ape
## Loading required package: maps

download.file("https://github.com/TIMAVID/Ambystoma/blob/master/GMM/Data/Amb_species?raw=true",
  "Amb_species.txt")
# Read in tree
tree <- read.newick("Amb_species.txt") #tree from Williams et al. 2013
par(mar = c(1, 1, 1, 1))
tree$tip.label <- gsub("^", "A.", tree$tip.label)
plot(tree)
```



```
# Subset tree to include only GMM species
Amb_species <- unique(GMM_data_sub$species)
tips <- tree$tip.label
```

```
ii <- sapply(Amb_species, function(x, y) grep(x, y)[1], y = tips)
tree <- drop.tip(tree, setdiff(tree$tip.label, tips[ii]))
plotTree(tree, ftype = "i")
```



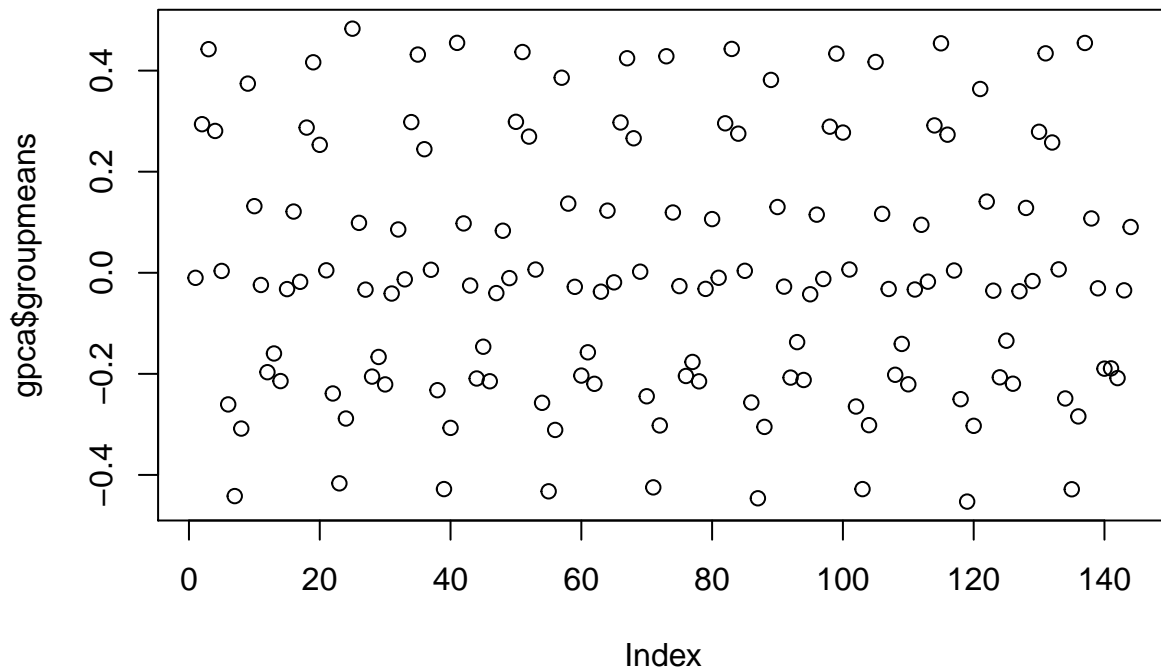
```
# Tree did not include A.mavortium so I lumped that species with A.tigrinum
Amb_gdf$species <- gsub("A.mavortium", "A.tigrinum", Amb_gdf$species, fixed = TRUE)
Amb_gdf$species <- as.factor(Amb_gdf$species)
```

## Preformed a group PCA

```
library(Morpho)

## Registered S3 method overwritten by 'Morpho':
##   method      from
##   print.classify RRPP
##
## Attaching package: 'Morpho'
## The following object is masked from 'package:RRPP':
##
##   classify
gpca <- groupPCA(Amb_gdf$coords, Amb_gdf$species, rounds = 0)
plot(gpca$groupmeans)
```





Performed a Phylogenetic PCA based on group means

```
phylo.PCA <- gm.prcomp(gpca$groupmeans, phy = tree)
summary(phylo.PCA)
```

```
##
## Ordination type: Principal Component Analysis
## Centering by OLS mean
## Orthogonal projection of OLS residuals
## Number of observations: 9
## Number of vectors 9
##
## Importance of Components:
##
```

	Comp1	Comp2	Comp3	Comp4
## Eigenvalues	0.002786587	0.0003175153	0.0001206304	0.0001026932
## Proportion of Variance	0.815069827	0.0928724554	0.0352841101	0.0300375020
## Cumulative Proportion	0.815069827	0.9079422824	0.9432263926	0.9732638946

```
##
```

	Comp5	Comp6	Comp7	Comp8
## Eigenvalues	5.699343e-05	1.597011e-05	1.473938e-05	3.703327e-06
## Proportion of Variance	1.667044e-02	4.671219e-03	4.311234e-03	1.083214e-03
## Cumulative Proportion	9.899343e-01	9.946056e-01	9.989168e-01	1.000000e+00

```
##
##
```

	Comp9
## Eigenvalues	5.409390e-34
## Proportion of Variance	1.582234e-31
## Cumulative Proportion	1.000000e+00

```
##
##
## Dispersion (variance) of points, after projection:
##
```

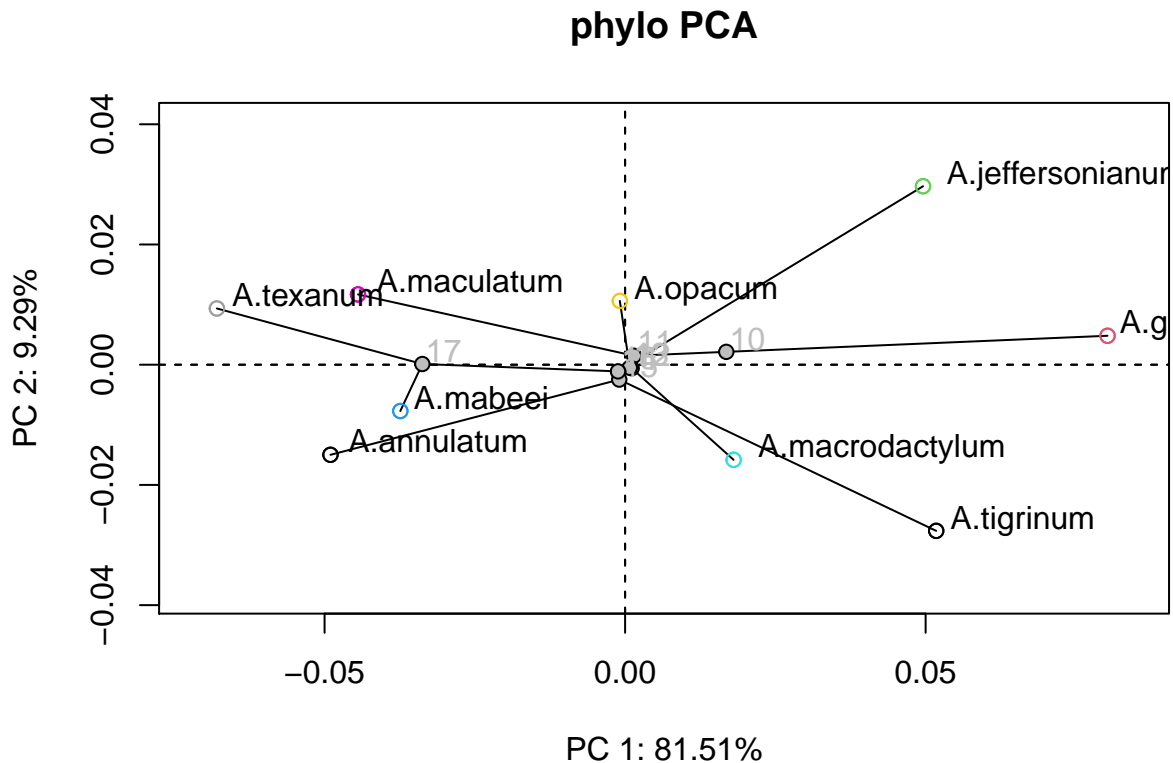
	Comp1	Comp2	Comp3
## Tips Dispersion	0.0027865866	3.175153e-04	0.0001206304
## Proportion Tips Dispersion	0.8150698271	9.287246e-02	0.0352841101

```
## Cumulative Tips Dispersion      0.8150698271 9.079423e-01 0.9432263926
## Ancestors Dispersion           0.0001990754 2.231723e-06 0.0000028151
## Proportion Ancestors Dispersion 0.9407409454 1.054612e-02 0.0133028988
## Cumulative Ancestors Dispersion 0.9407409454 9.512871e-01 0.9645899644
##                               Comp4      Comp5      Comp6
## Tips Dispersion                1.026932e-04 5.699343e-05 1.597011e-05
## Proportion Tips Dispersion      3.003750e-02 1.667044e-02 4.671219e-03
## Cumulative Tips Dispersion      9.732639e-01 9.899343e-01 9.946056e-01
## Ancestors Dispersion            1.429229e-06 3.973386e-06 1.456089e-06
## Proportion Ancestors Dispersion 6.753896e-03 1.877644e-02 6.880821e-03
## Cumulative Ancestors Dispersion 9.713439e-01 9.901203e-01 9.970011e-01
##                               Comp7      Comp8      Comp9
## Tips Dispersion                1.473938e-05 3.703327e-06 2.459722e-36
## Proportion Tips Dispersion      4.311234e-03 1.083214e-03 7.194627e-34
## Cumulative Tips Dispersion      9.989168e-01 1.000000e+00 1.000000e+00
## Ancestors Dispersion            5.936065e-07 4.100346e-08 7.525691e-38
## Proportion Ancestors Dispersion 2.805118e-03 1.937639e-04 3.556303e-34
## Cumulative Ancestors Dispersion 9.998062e-01 1.000000e+00 1.000000e+00
```

```
A_species <- attributes(gpca$groupmeans) #access attributes names
A_species <- (A_species$dimnames[[3]])
A_species <- as.factor(A_species)
```

## Plot phylogenetic PCA

```
plot(phylo.PCA, phylo = TRUE, main = "phylo PCA", col = A_species)
```



### 3D plot of phylogenetic PCA

```
plotdat <- phylo.PCA$x[, 1:3]
colnames(plotdat) <- c("", "", "") #prevent axis labels
obj <- phytools::phylomorphospace3d(tree, plotdat, method = "dynamic", control = list(ftype = "off",
  spin = FALSE, box = FALSE), cex.symbol = 0.5)
spheres3d(phylo.PCA$x[, 1:3], color = palette()[A_species], r = 0.01)
bbox3d(color = c("white"), shininess = 15, alpha = 0.3, xat = c(10), xlab = "x",
  yat = c(10), ylab = "y", zat = c(10), zlab = "z")
text3d((phylo.PCA$x[, 1:3] + 0.005), texts = substr(A_species, 1, 6))
```

Test for phylogenetic signal, uses Blomberg's K to test for strength and significance of phylogenetic signal.

```
physignal(gpca$groupmeans, tree, print.progress = F)

##
## Call:
## physignal(A = gpca$groupmeans, phy = tree, print.progress = F)
##
##
## Observed Phylogenetic Signal (K): 0.9576
##
## P-value: 0.124
##
## Effect Size: 1.1131
##
## Based on 1000 random permutations
```

### Phylogenetic generalized least squares

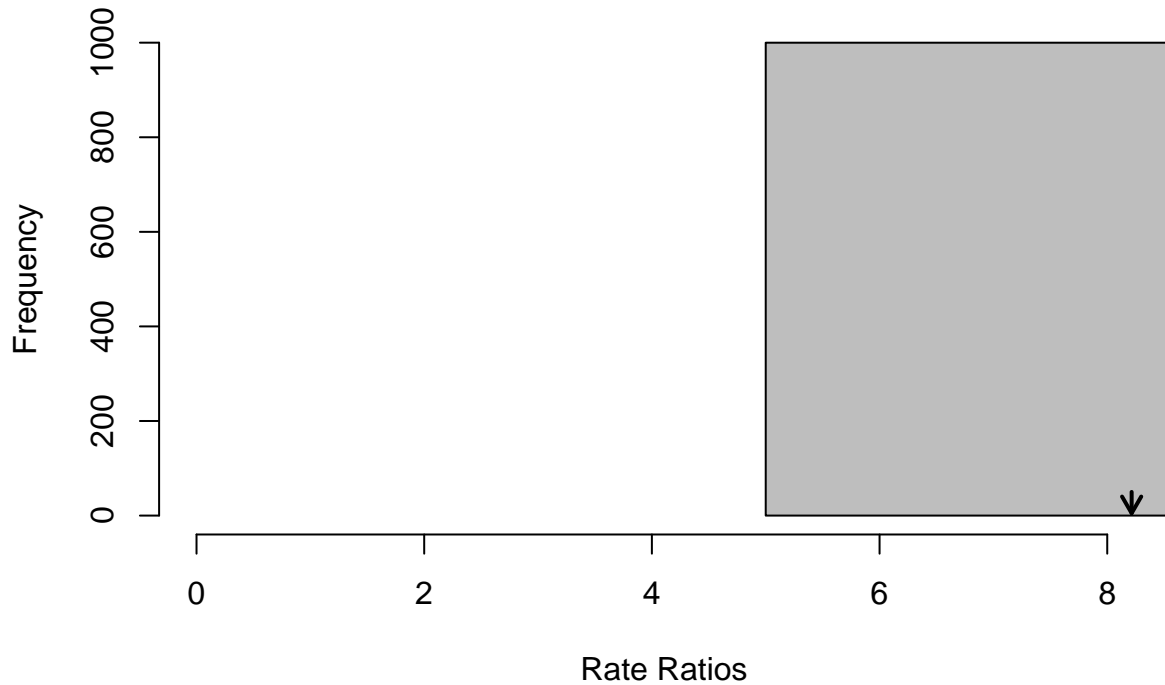
```
avg_gdf <- geomorph.data.frame(coords = gpca$groupmeans, species = A_species) #make new geomorph dataf
pgls <- procD.pgls(coords ~ species, phy = tree, data = avg_gdf, print.progress = F) #Phylogenetic gen
pgls$aov.table
```

```
##           Df      SS  MS Rsq  F Z Pr(>F)
## species    8 1.0748   0   1  0  0.7295
## Residuals   0 0.0000 Inf   0
## Total      8 1.0748
```

Compare evolutionary rates in different portions of the tree based on brownian motion

```
names(avg_gdf$species) <- avg_gdf$species
rate.comp <- compare.evol.rates(avg_gdf$coords, tree, gp = avg_gdf$species,
  print.progress = F)
plot(rate.comp)
```

Observed Rate Ratio = 8.2142 ; P-value = 1



```
rate.comp$sigma.d.gp
```

```
##      A.annulatum      A.gracile A.jeffersonianum      A.mabeei
##      0.009432017      0.006943721      0.010206699      0.003679637
## A.macrodactylum      A.maculatum      A.opacum      A.texanum
##      0.001940983      0.006185997      0.002044237      0.015943548
##      A.tigrinum
##      0.010795768
```

```
rate.comp$pairwise.pvalue
```

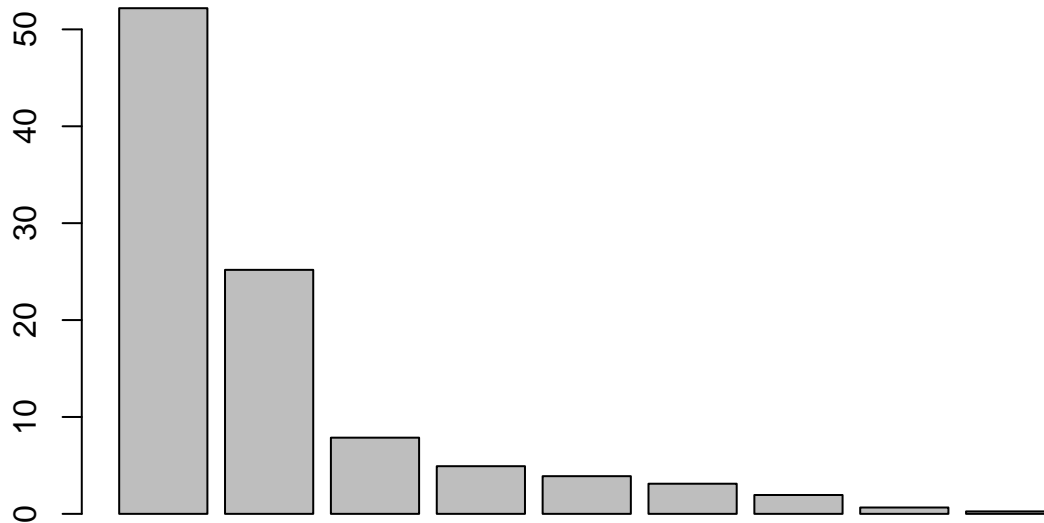
```
##      A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.gracile      0.826
## A.jeffersonianum 0.954      0.837
## A.mabeei      0.427      0.536      0.410
## A.macrodactylum 0.189      0.277      0.144      0.515
## A.maculatum      0.801      0.918      0.703      0.654
## A.opacum      0.222      0.294      0.162      0.590
## A.texanum      0.652      0.517      0.674      0.246
## A.tigrinum      0.864      0.719      0.939      0.343
##      A.macrodactylum A.maculatum A.opacum A.texanum
## A.gracile
## A.jeffersonianum
## A.mabeei
## A.macrodactylum
## A.maculatum      0.349
## A.opacum      0.973      0.374
## A.texanum      0.030      0.437      0.042
## A.tigrinum      0.058      0.587      0.082      0.814
```

## Discriminant Function Analysis

```
library(Morpho)
DFA <- CVA(GPA_landmarks_sub$coords, GMM_data_sub$species, cv = TRUE)
```

```
## singular Covariance matrix: General inverse is used. Threshold for zero eigenvalue is 1e-10
```

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



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

```
##
##           A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.annulatum           1         0                 0         3
## A.gracile             0         2                 2         0
## A.jeffersonianum      0         2                 0         0
## A.mabeei              3         0                 0         0
## A.macrodactylum      0         1                 0         1
## A.maculatum           1         0                 1         3
## A.mavortium           1         1                 0         0
## A.opacum              0         2                 0         0
## A.texanum             0         0                 0         1
## A.tigrinum            1         1                 0         0
##
##           A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.annulatum                0           0           0         0         2
## A.gracile                  2           0           1         1         0
## A.jeffersonianum           0           1           0         2         0
## A.mabeei                   2           1           0         1         1
## A.macrodactylum           5           0           0         1         0
## A.maculatum                 0           6           0         0         1
## A.mavortium                 1           0           4         0         0
## A.opacum                    0           2           0         6         0
## A.texanum                   1           1           0         0         7
## A.tigrinum                  2           0           2         0         0
##
```

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

```
diag(prop.table(accJack, 1)) #accuracy per species as %
```

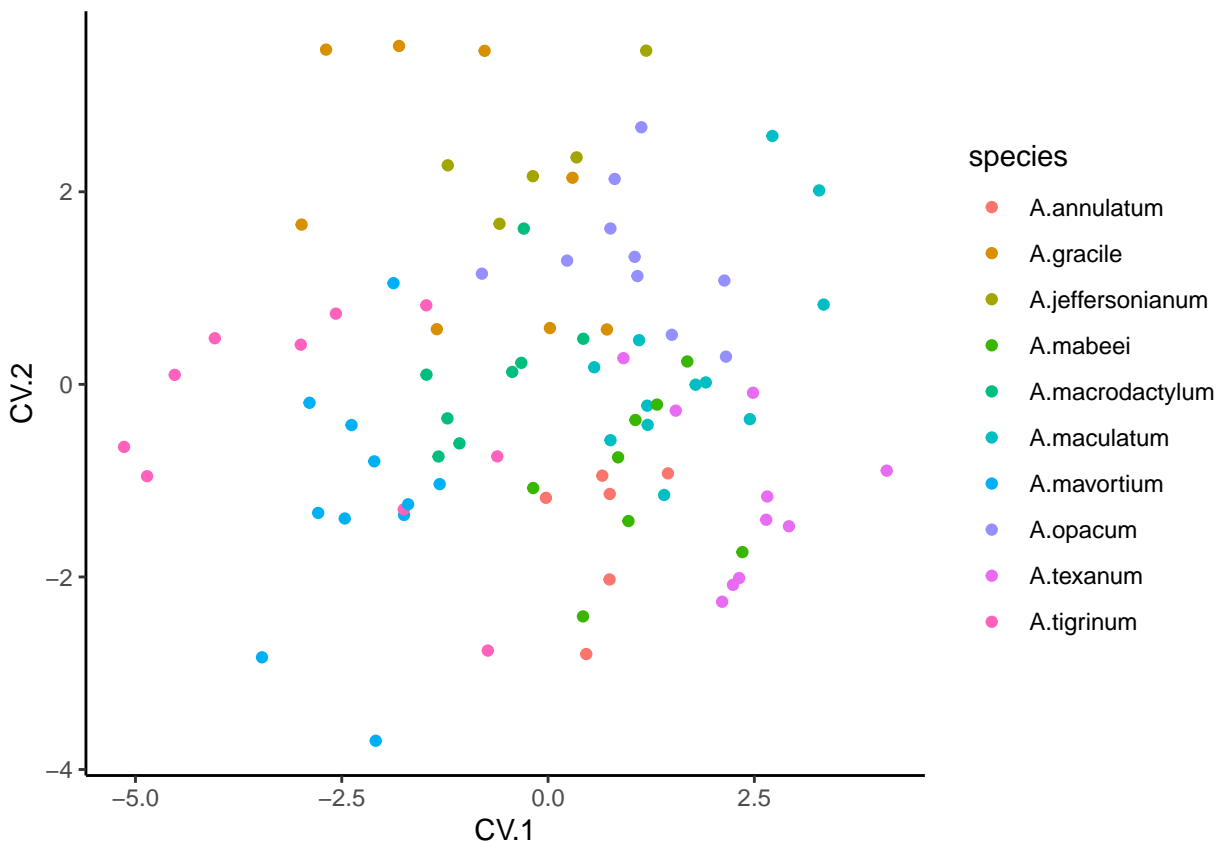
```
##      A.annulatum      A.gracile A.jeffersonianum      A.mabeei
##      0.1666667      0.2500000      0.0000000      0.0000000
## A.macrodactylum A.maculatum      A.mavortium      A.opacum
##      0.6250000      0.5000000      0.3636364      0.6000000
##      A.texanum      A.tigrinum
##      0.7000000      0.4000000
```

```
sum(accJack[row(accJack) == col(accJack)])/sum(accJack) #overall accuracy
```

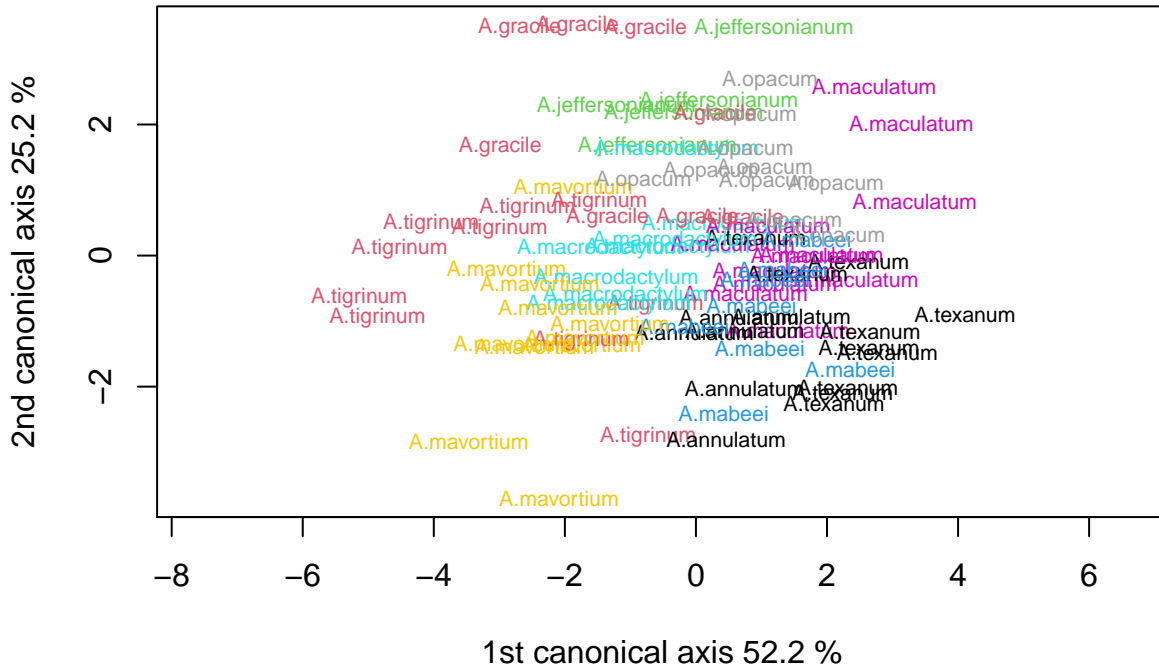
```
## [1] 0.3977273
```

```
# Plot first two DF axes #
```

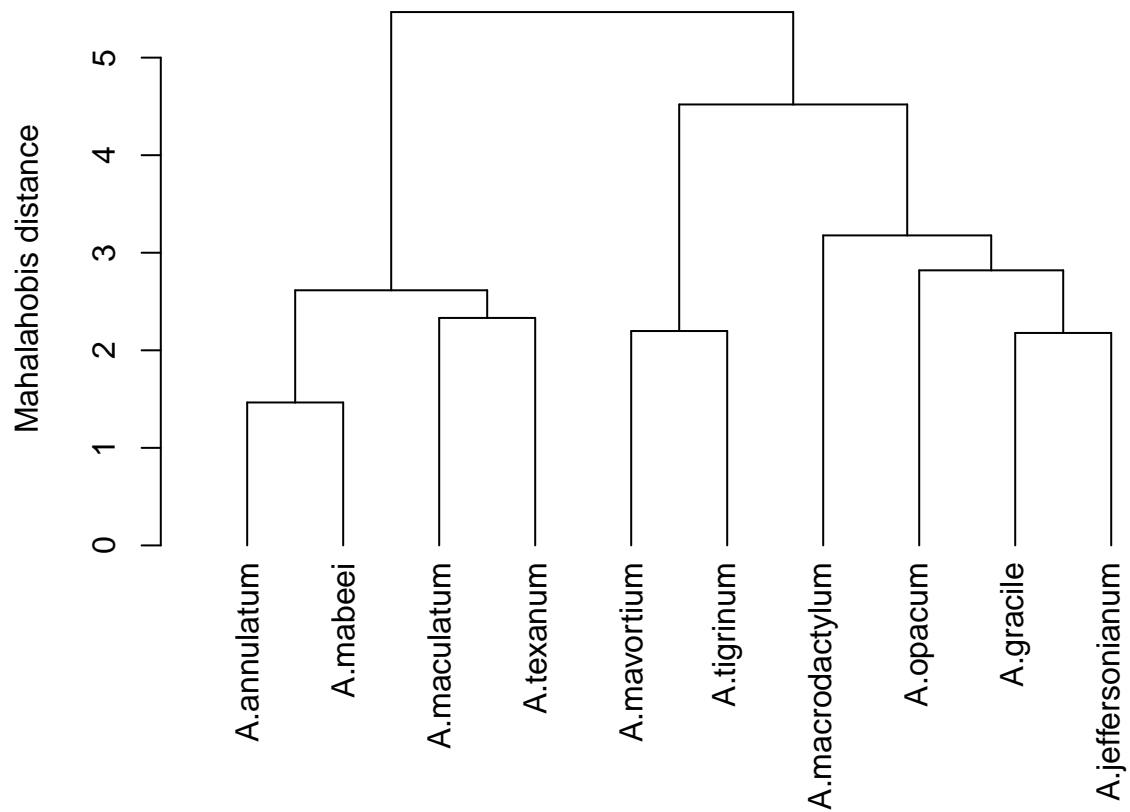
```
DFA_cva <- data.frame(DFA$CVscores, species = DFA$groups)
ggplot(DFA_cva, aes(CV.1, CV.2)) + geom_point(aes(color = species)) + theme_classic()
```



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



```
# Plot Mahalahobis distances as dendrogram #
dendroS = hclust(DFA$Dist$GroupdistMaha)
dendroS$labels = levels(GMM_data_sub$species)
par(mar = c(6.5, 4.5, 1, 1))
dendroS = as.dendrogram(dendroS)
plot(dendroS, main = "", sub = "", xlab = "", ylab = "Mahalahobis distance")
```



## K Nearest neighbor ###:Non-parametric

```
library(caret)
```

```
## Loading required package: lattice
```

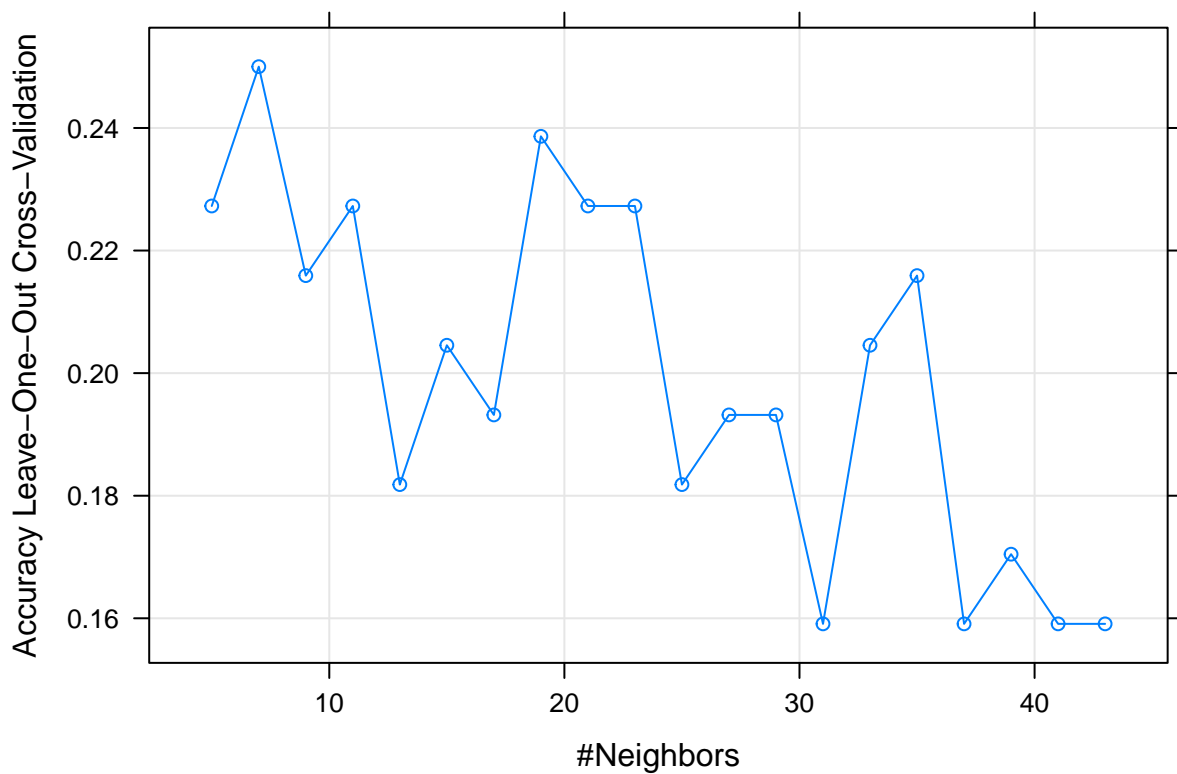
```
Atlas_PC_scores <- data.frame(Amb_PCA$x,species=GMM_data_sub$species)
```

```
set.seed(123)
```

```
KNNmodel <- train(
  species ~., data = Atlas_PC_scores, method = "knn",
  trControl = trainControl("LOOCV", number = 1),
  preProcess = c("center","scale"), #scale the data
  tuneLength = 20)
```

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





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

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

```
predicted.classes <- KNNmodel %>% predict(Atlas_PC_scores[,1:17]) # predict class based on KNN model
head(predicted.classes)
```

```
## [1] A.maculatum      A.jeffersonianum A.gracile         A.jeffersonianum
## [5] A.maculatum      A.maculatum
## 10 Levels: A.annulatum A.gracile A.jeffersonianum A.mabeei ... A.tigrinum
```

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

```
## [1] 0.5340909
```

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

```
##               predicted.classes
##               A.annulatum A.gracile A.jeffersonianum A.mabeei
## A.annulatum             4         0         0         0
## A.gracile                0         2         0         1
## A.jeffersonianum         0         1         2         0
## A.mabeei                 0         1         0         1
## A.macrodactylum         0         1         0         2
## A.maculatum              0         0         0         1
## A.mavortium              1         1         0         0
## A.opacum                 0         0         0         0
## A.texanum                0         0         0         2
## A.tigrinum               0         0         0         1
##               predicted.classes
```

```
##           A.macrodactylum A.maculatum A.mavortium A.opacum A.texanum
## A.annulatum                0            1            1            0            0
## A.gracile                  1            1            1            1            0
## A.jeffersonianum           0            2            0            0            0
## A.mabeei                   1            2            0            1            2
## A.macrodactylum           2            2            0            1            0
## A.maculatum                 0            9            0            2            0
## A.mavortium                 1            0            8            0            0
## A.opacum                    0            1            0            8            1
## A.texanum                   0            2            0            1            5
## A.tigrinum                  1            0            1            1            0
```

```
##           predicted.classes
```

```
##           A.tigrinum
```

```
## A.annulatum                0
## A.gracile                  1
## A.jeffersonianum           0
## A.mabeei                   0
## A.macrodactylum           0
## A.maculatum                 0
## A.mavortium                 0
## A.opacum                    0
## A.texanum                   0
## A.tigrinum                  6
```

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

```
##           A.annulatum           A.gracile A.jeffersonianum           A.mabeei
##           0.6666667           0.2500000           0.4000000           0.1250000
## A.macrodactylum           A.maculatum           A.mavortium           A.opacum
##           0.2500000           0.7500000           0.7272727           0.8000000
##           A.texanum           A.tigrinum
##           0.5000000           0.6000000
```