Figure.3.4

Devon DeRaad

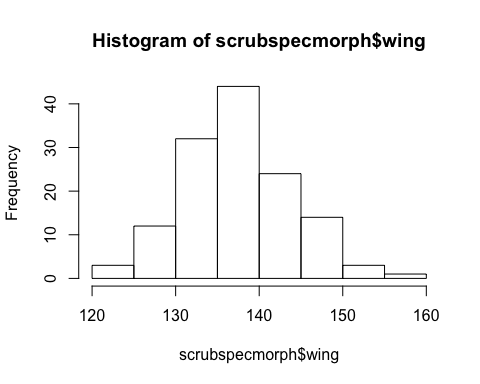
16 March 2019

library(ggplot2)  
#install.packages("devtools")  
library(devtools)  
#install\_github("vqv/ggbiplot", force = TRUE)  
library(ggbiplot)  
library(grid)  
library(gridExtra)  
#install.packages("rworldmap")  
library(rworldmap)  
library(maps)  
#install.packages("mapdata")  
library(mapdata)  
library(ggmap)  
#install.packages("ggridges")  
library(ggridges)

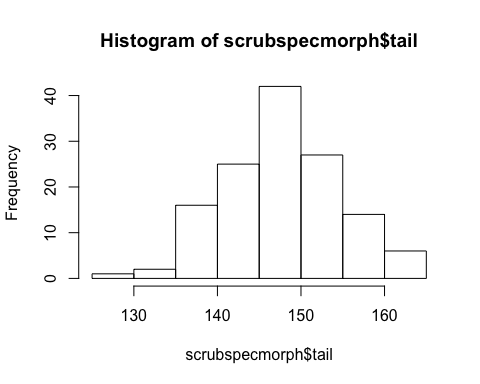
Read in dataframe containing all morphological data

scrubspecmorph <- read.csv("~/Dropbox/scrub poster/manuscript/scrubspecmorph.csv", header=T)

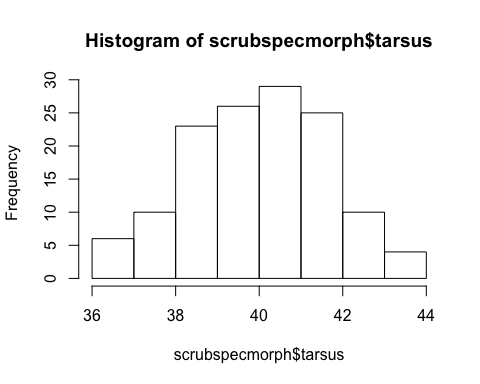
##normalitycheck  
hist(scrubspecmorph$wing)



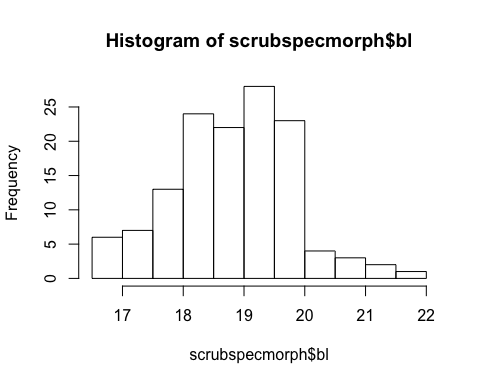
hist(scrubspecmorph$tail)



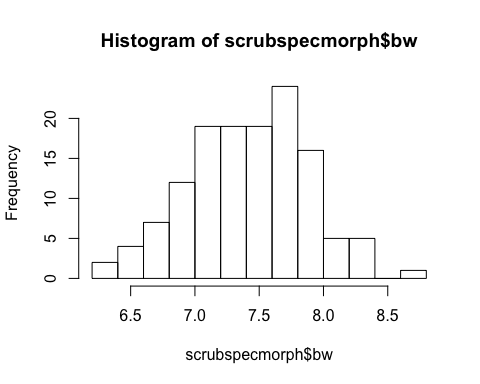
hist(scrubspecmorph$tarsus)



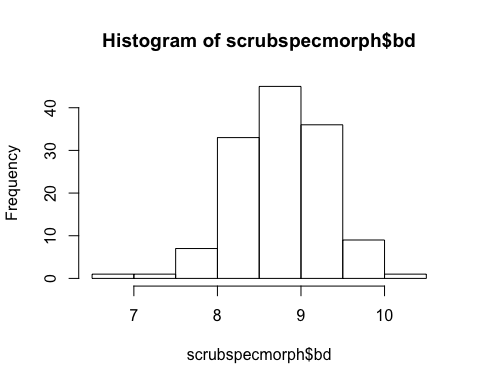
hist(scrubspecmorph$bl)



hist(scrubspecmorph$bw)



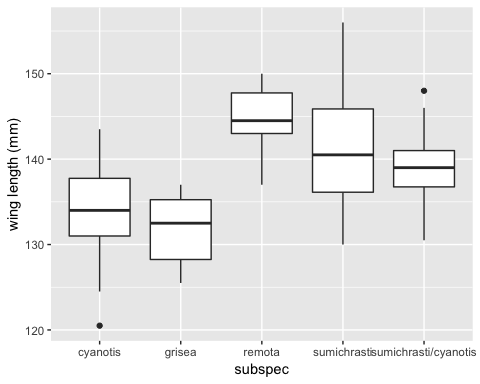
hist(scrubspecmorph$bd)



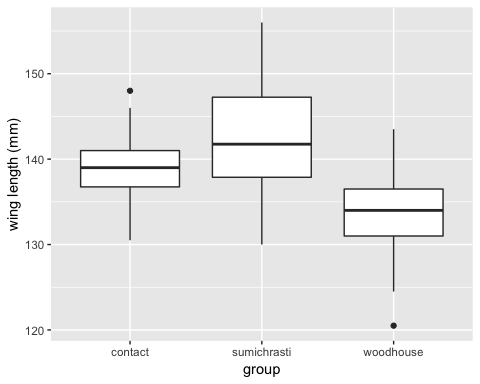
##all six characteristics look normalish

Make a few exploratory plots to get a feel for the data and the size differences between individual subspecies, and between lineages as a whole.

###compare wing length between subspecies  
subspecwing<-ggplot(scrubspecmorph, aes(subspecies, wing)) +   
 geom\_boxplot() +  
 ylab("wing length (mm)")+  
 xlab("subspec")  
#plot  
subspecwing

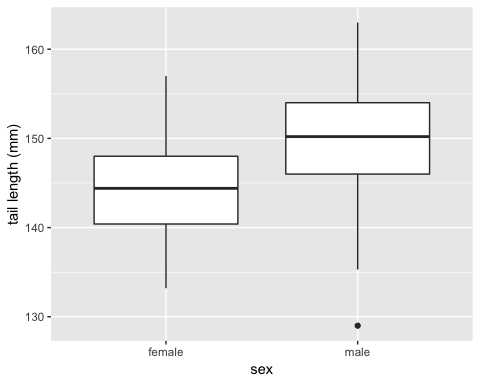


###compare lineage level differences in wing length  
specwing<-ggplot(scrubspecmorph, aes(locality, wing)) +   
 geom\_boxplot() +  
 ylab("wing length (mm)")+  
 xlab("group")  
#plot  
specwing

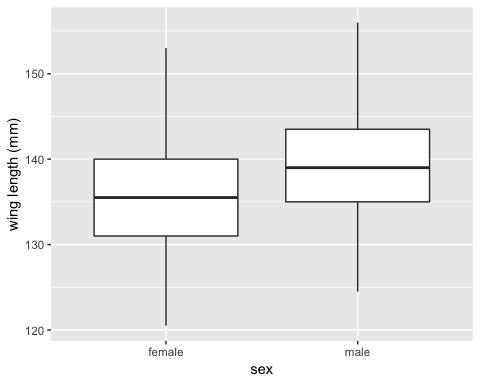


investigate sexual dimorphism identified by pitelka (1951)

#compare tail length by sex  
tailsex<-ggplot(scrubspecmorph, aes(sex, tail)) +   
 geom\_boxplot() +  
 ylab("tail length (mm)")+  
 xlab("sex")  
#plot  
tailsex #noticeable difference



#wing length by sex  
wingsex<-ggplot(scrubspecmorph, aes(sex, wing)) +   
 geom\_boxplot() +  
 ylab("wing length (mm)")+  
 xlab("sex")  
#plot  
wingsex #males slightly larger once again



Calculate p-value for effect of sex on all 6 size characteristics simultaneously using a MANOVA.

#Manova for male v female in all 6 characteristics  
fem.male <- manova(cbind(wing, tail, tarsus, bl, bw, bd) ~ sex, data = scrubspecmorph)  
summary(fem.male)

## Df Pillai approx F num Df den Df Pr(>F)   
## sex 1 0.33498 10.578 6 126 1.712e-09 \*\*\*  
## Residuals 131   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Extremely significant sex difference, likely driven by consistency (lack of variation within groups) leading to small overall size difference being considered very un-random in it’s distribution between the sexes.

Make PCA

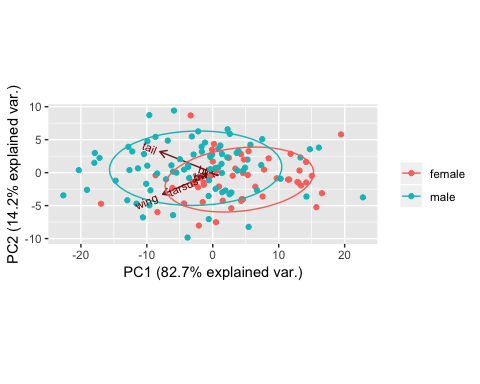
##morph PCA  
pca.morph<- prcomp(scrubspecmorph[, 8:13])  
#PrintPCA  
print(pca.morph)

## Standard deviations (1, .., p=6):  
## [1] 8.8083189 3.6485636 1.3437346 0.8979285 0.4759318 0.3447896  
##   
## Rotation (n x k) = (6 x 6):  
## PC1 PC2 PC3 PC4 PC5  
## wing -0.68212989 -0.71159382 0.16182466 -0.03084559 -3.437415e-02  
## tail -0.72026391 0.69100157 0.01371913 0.05948219 -5.930875e-05  
## tarsus -0.10985456 -0.12046045 -0.94381773 0.28596507 -2.452300e-02  
## bl -0.04803534 0.03537597 -0.27446802 -0.93051454 -2.343746e-01  
## bw -0.01995937 -0.01178350 -0.04130483 -0.13529628 6.450935e-01  
## bd -0.03383667 -0.01581283 -0.07611754 -0.17196513 7.260441e-01  
## PC6  
## wing -0.003578016  
## tail -0.003223085  
## tarsus -0.015873068  
## bl -0.017880011  
## bw -0.750537125  
## bd 0.660378062

#summarize PC's  
summary(pca.morph)

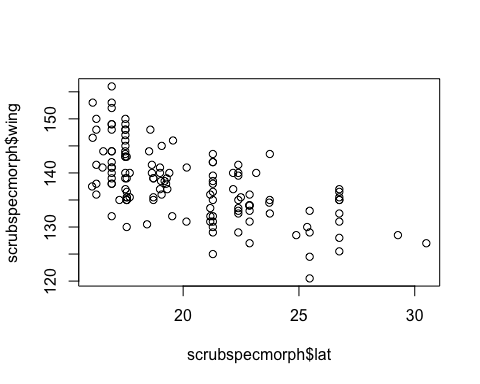
## Importance of components:  
## PC1 PC2 PC3 PC4 PC5 PC6  
## Standard deviation 8.8083 3.6486 1.34373 0.89793 0.47593 0.34479  
## Proportion of Variance 0.8267 0.1418 0.01924 0.00859 0.00241 0.00127  
## Cumulative Proportion 0.8267 0.9685 0.98773 0.99632 0.99873 1.00000

#visualize the differences between sexes  
ggbiplot(pca.morph, obs.scale = 1, var.scale = 1,  
 groups = scrubspecmorph$sex, ellipse = TRUE) +  
 scale\_color\_discrete(name = '')

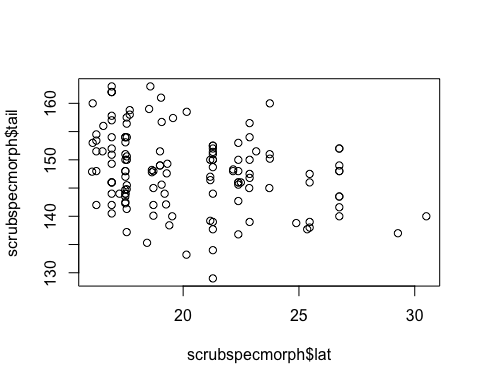


The difference between the sexes is largely explained by wing and tail length differences which load heavily on PC1. We have extremely similar sex ratios between the two lineages, and the overall difference is low, but it is worth monitoring that there is consistently detectable sexual dimorphism.

#plot winglength v latitude  
plot(scrubspecmorph$lat, scrubspecmorph$wing)



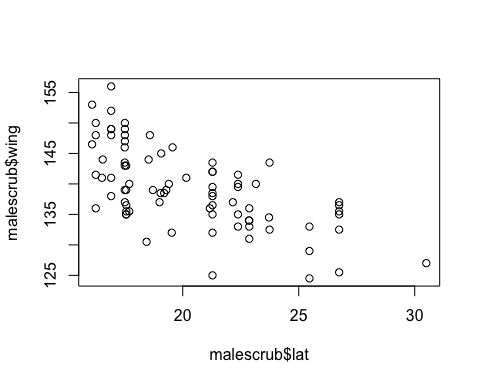
#plot taillength v latitude  
plot(scrubspecmorph$lat, scrubspecmorph$tail)



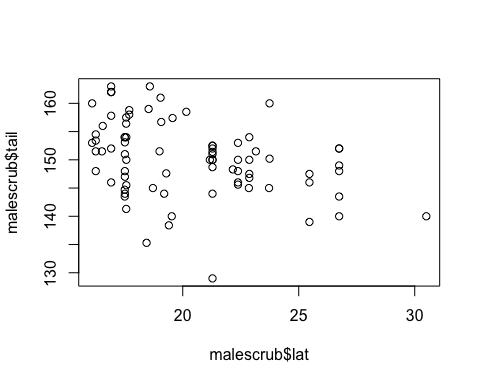
Here we see a distinct pattern of smaller birds at higher (more northern) latitudes

Lets check to see whether this pattern holds once sex is accounted for

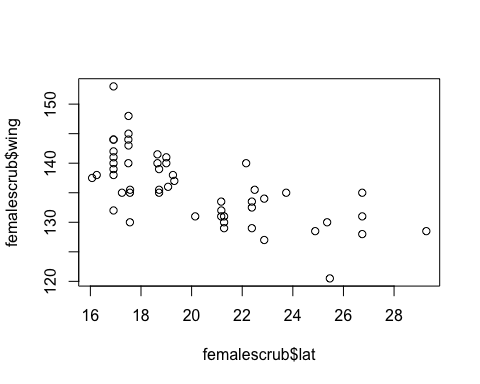
#subset dataset by sex  
malescrub<- subset(scrubspecmorph, scrubspecmorph$sex=="male")  
femalescrub<-subset(scrubspecmorph, scrubspecmorph$sex=="female")  
#plot males  
plot(malescrub$lat, malescrub$wing)



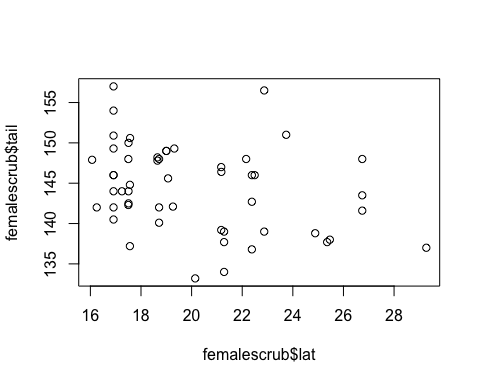
plot(malescrub$lat, malescrub$tail)



#plot females  
plot(femalescrub$lat, femalescrub$wing)

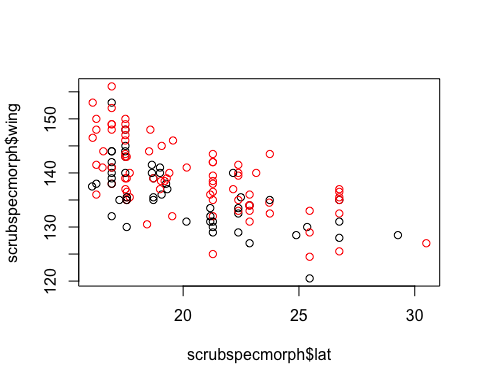


plot(femalescrub$lat, femalescrub$tail)



# Relationships between size and geography hold whether divided by sex or not divided at all.

#plot winglength v latitude color coded for sex  
plot(scrubspecmorph$lat, scrubspecmorph$wing, col = scrubspecmorph$sex)



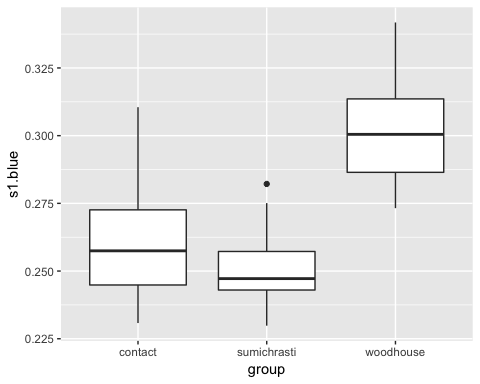
#plot taillength v latitude color coded for sex  
plot(scrubspecmorph$lat, scrubspecmorph$tail, col = scrubspecmorph$sex)



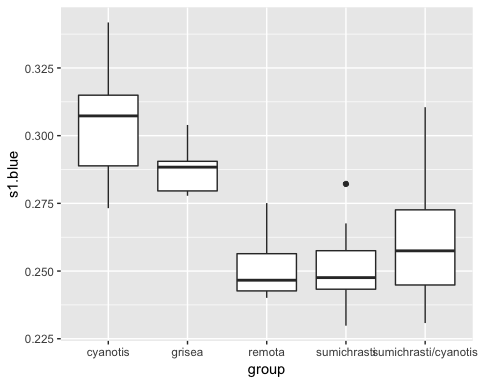
Size differences between the sexes noticeable but the pattern is identical. Males in red, females in black. The two sexes are largely overlapping and identical in signal.

Check out the difference in amount of blue on the back

#check out the spec variable based on group  
#boxplot based on the spec variable S1.blue (blue saturation of the mantle)  
S1.blueboxgroup<-ggplot(scrubspecmorph, aes(locality, S1.blue)) +   
 geom\_boxplot() +  
 ylab("s1.blue")+  
 xlab("group")  
S1.blueboxgroup

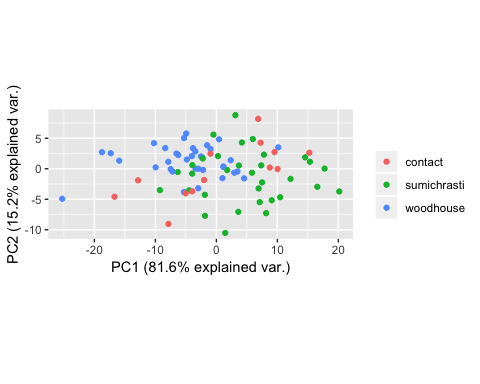


#by sub  
S1.blueboxsub<-ggplot(scrubspecmorph, aes(subspecies, S1.blue)) +   
 geom\_boxplot() +  
 ylab("s1.blue")+  
 xlab("group")  
S1.blueboxsub

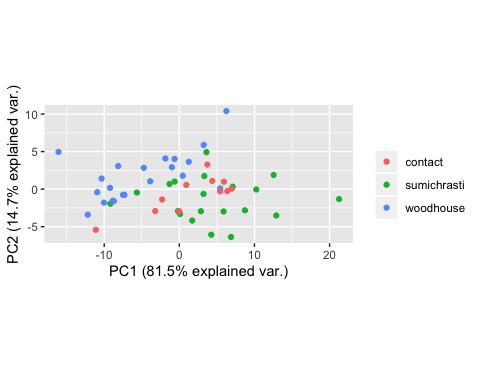


Make a PCA for each sex separately

#pca for each sex separately  
pca.malescrubmorph <- prcomp(malescrub[,8:13])  
pca.femalescrubmorph <- prcomp(femalescrub[,8:13])  
#visualize male  
ggbiplot(pca.malescrubmorph, obs.scale = 1, var.scale = 1,  
 groups = malescrub$locality, ellipse = FALSE,  
 var.axes = FALSE) +  
 scale\_color\_discrete(name = '') +  
 theme(legend.direction = 'vertical')



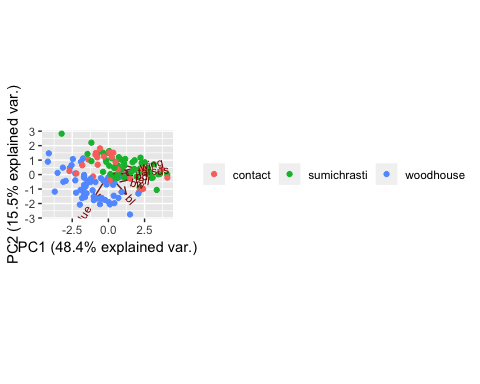
#visualize female  
ggbiplot(pca.femalescrubmorph, obs.scale = 1, var.scale = 1,  
 groups = femalescrub$locality, ellipse = FALSE,  
 var.axes = FALSE) +  
 scale\_color\_discrete(name = '') +  
 theme(legend.direction = 'vertical')



Both PCAs separate the lineages nicely

We are now going to make a PCA for all individuals

#PCA of all variables morph + blue color using correlative matrix  
pca.scruball <- princomp(scrubspecmorph[,8:14], cor=TRUE)  
  
#visualize PCA of all 6 morph variables + blue saturation of mantle by lineage  
ggbiplot(pca.scruball, obs.scale = 1, var.scale = 1,  
 groups = scrubspecmorph$locality, ellipse = FALSE) +  
 scale\_color\_discrete(name = '') +  
 theme(legend.direction = 'horizontal')

 We can see that this nicely separates the lineages based on body size differences on PC1.

We will now pull out the PC values for each individual and append it to our original dataset, so that we can use PC1 as a variable indicating general body size.

#view loadings by variable  
unclass(pca.scruball$loadings)

## Comp.1 Comp.2 Comp.3 Comp.4 Comp.5  
## wing 0.4633632 0.19945991 0.27294790 0.22938513 0.13557958  
## tail 0.4153102 -0.10267070 0.39993461 0.60705860 -0.05503053  
## tarsus 0.3915982 0.06471396 0.33071829 -0.65438235 0.48442886  
## bl 0.3164987 -0.61388721 0.03348982 -0.29837207 -0.47882936  
## bw 0.3334493 -0.17526889 -0.70571601 0.19253093 0.50068508  
## bd 0.4253020 -0.02913543 -0.34105412 -0.09159202 -0.34312376  
## S1.blue -0.2592707 -0.73284447 0.20174886 0.12713342 0.38238917  
## Comp.6 Comp.7  
## wing 0.039742048 0.7735832088  
## tail 0.009008839 -0.5342273243  
## tarsus -0.071498463 -0.2551250440  
## bl 0.438345985 0.1067653828  
## bw 0.261215277 -0.0637978419  
## bd -0.758833348 -0.0006203457  
## S1.blue -0.396255860 0.1887114214

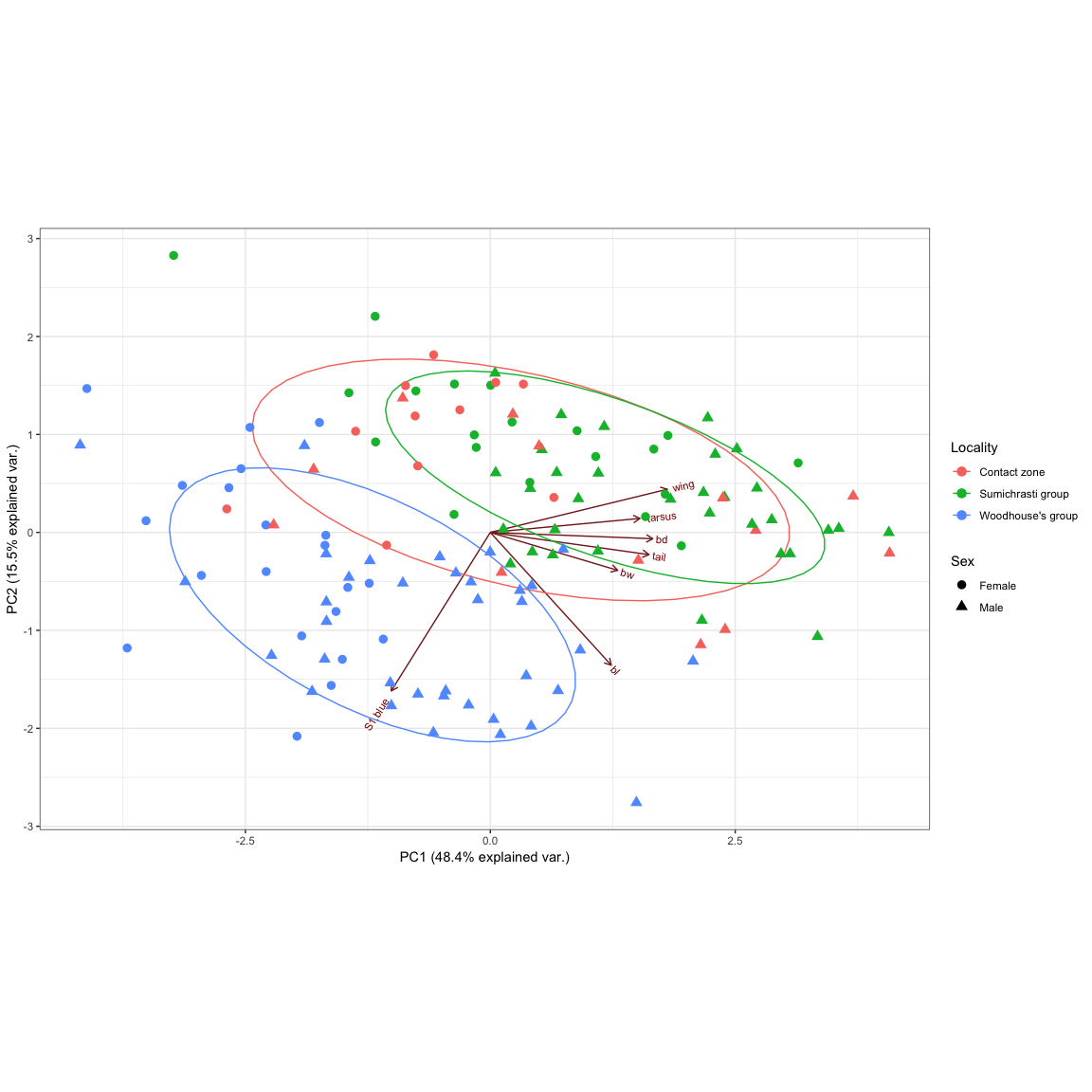
#create dataframe of PC scores  
pcascoresbyindiv<-data.frame(pca.scruball$scores)  
#insert id column from scrubspecmorph into pcascorebyindiv  
pcascoresbyindiv$id <- scrubspecmorph$id  
#merge datasets by the column ID to ensure that the correct values stay with correct individuals  
PCAscrubspecmorph<-merge(scrubspecmorph, pcascoresbyindiv, by = "id")  
head(PCAscrubspecmorph)

## id museum sex locality subspecies lat long  
## 1 18334 psm male contact sumichrasti/cyanotis 19.04498 -98.20233  
## 2 19180 mlz female woodhouse grisea 26.74261 -106.05285  
## 3 19192 MLZ male woodhouse grisea 26.74261 -106.05285  
## 4 19196 mlz female woodhouse grisea 26.74261 -106.05285  
## 5 19197 MLZ male woodhouse grisea 26.74261 -106.05285  
## 6 19201 mlz male woodhouse grisea 26.74261 -106.05285  
## wing tail tarsus bl bw bd S1.blue S1.red Comp.1  
## 1 138.5 161.0 38.9 19.4 8.7 9.7 0.2307985 0.3548928 2.7093513  
## 2 128.0 141.6 38.1 17.3 7.2 7.8 0.2864534 0.2474419 -3.1438981  
## 3 136.5 152.0 38.8 18.9 7.7 8.8 0.2898149 0.2429218 -0.1273813  
## 4 135.0 148.0 38.1 16.5 7.6 8.3 0.2778012 0.2504546 -1.7444138  
## 5 137.0 148.0 39.1 18.6 7.8 8.7 0.2883707 0.2444771 -0.3514860  
## 6 125.5 140.0 37.8 16.8 6.5 7.8 0.2900892 0.2389913 -4.1871707  
## Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7  
## 1 0.02123296 -2.1899981 1.6406742 -0.46438003 0.38948031 -1.18167296  
## 2 0.47965354 -0.1045860 0.4073618 0.53571944 0.45633253 -0.42306992  
## 3 -0.68685037 -0.3193044 0.9468791 0.03789025 0.02375313 -0.21977116  
## 4 1.12174965 -0.4521200 1.5046635 1.01159930 -0.18766426 -0.29304712  
## 5 -0.41435683 -0.5853505 0.6301926 0.46202174 0.09591630 0.05300395  
## 6 0.89200710 0.7238423 0.1632352 -0.07081998 -0.21278142 -0.47904828

# We can use the values of PC1 as a proxy for general body size in downstream analyses.

Let’s clean up that PCA for publication

#PCA color coded by lineage with sex indicated by shape  
ggbiplot(pca.scruball, obs.scale = 1, var.scale = 1,  
 groups = scrubspecmorph$locality, ellipse = TRUE) +  
 geom\_point(aes(colour=scrubspecmorph$locality, shape=scrubspecmorph$sex), size = 3)+  
 scale\_color\_discrete(name ="Locality",   
 labels = c("Contact zone", "Sumichrasti group", "Woodhouse's group")) +  
 theme(legend.direction = 'vertical') +  
 scale\_shape\_discrete(name = "Sex", labels = c("Female", "Male"))+  
 theme\_bw()

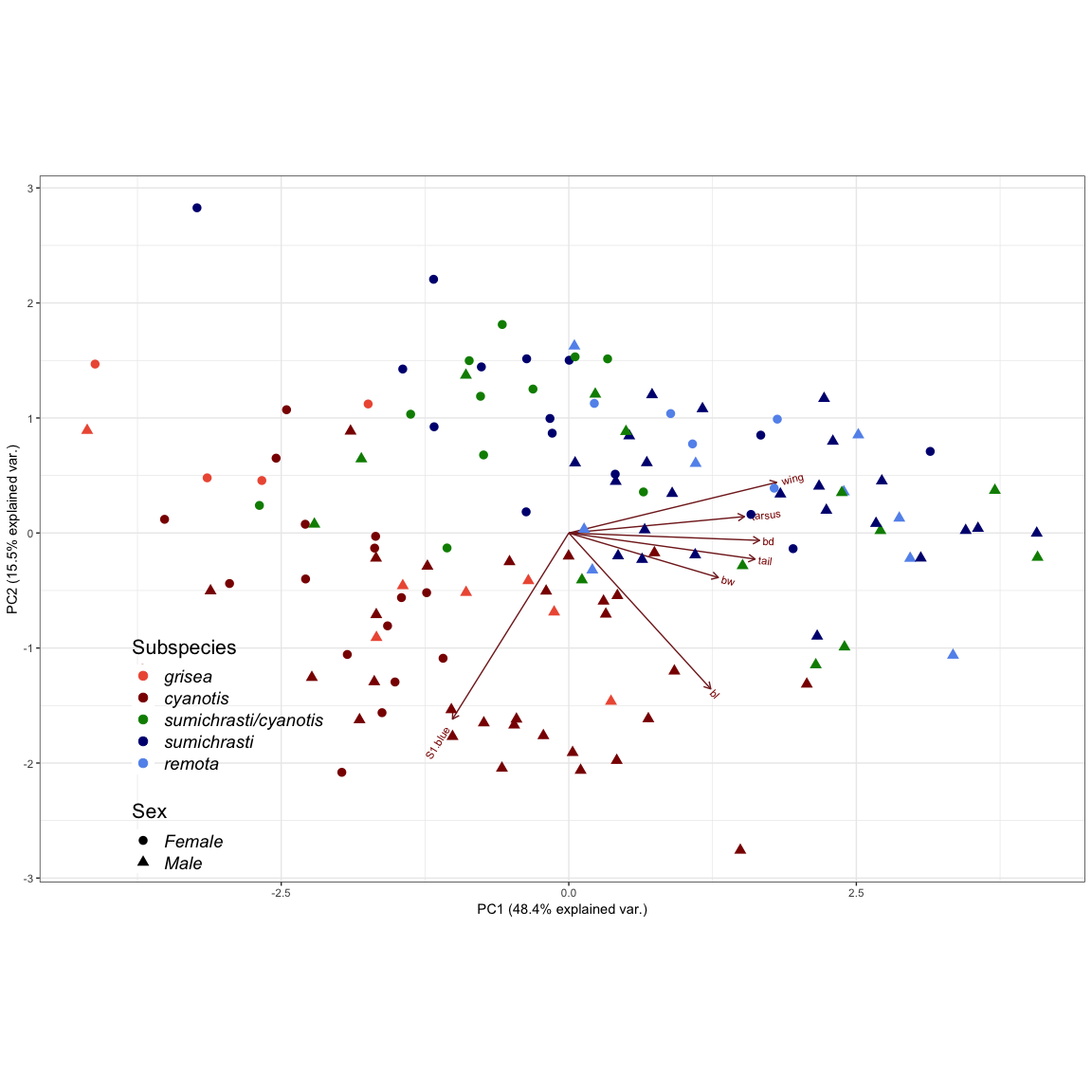


Looks good, now specify subspecies and custom color scheme to make the PCA presented in the text.

###Make manuscript Figure 3###  
##  
#  
PCAallsubspecies<-ggbiplot(pca.scruball, obs.scale = 1, var.scale = 1,  
 groups = scrubspecmorph$subspecies, ellipse = FALSE) +  
 geom\_point(aes(colour=scrubspecmorph$subspecies, shape=scrubspecmorph$sex), size = 3)+  
 scale\_color\_discrete(name ="Subspecies") +  
 scale\_color\_manual(values=c("tomato2", "darkred", "green4", "navyblue", "cornflowerblue"),  
 name="Subspecies",  
 limits=c("grisea", "cyanotis", "sumichrasti/cyanotis", "sumichrasti", "remota"))+  
 scale\_shape\_discrete(name = "Sex", labels = c("Female", "Male"))+  
 guides(color = guide\_legend(order=1),  
 shape = guide\_legend(order=2))+  
 theme\_bw()+  
 theme(legend.text=element\_text(size=14, face = "italic"),  
 legend.position =c(.18,.18),  
 legend.background = element\_rect(fill="transparent"),  
 legend.title=element\_text(size=16))

## Scale for 'colour' is already present. Adding another scale for  
## 'colour', which will replace the existing scale.

#print  
PCAallsubspecies



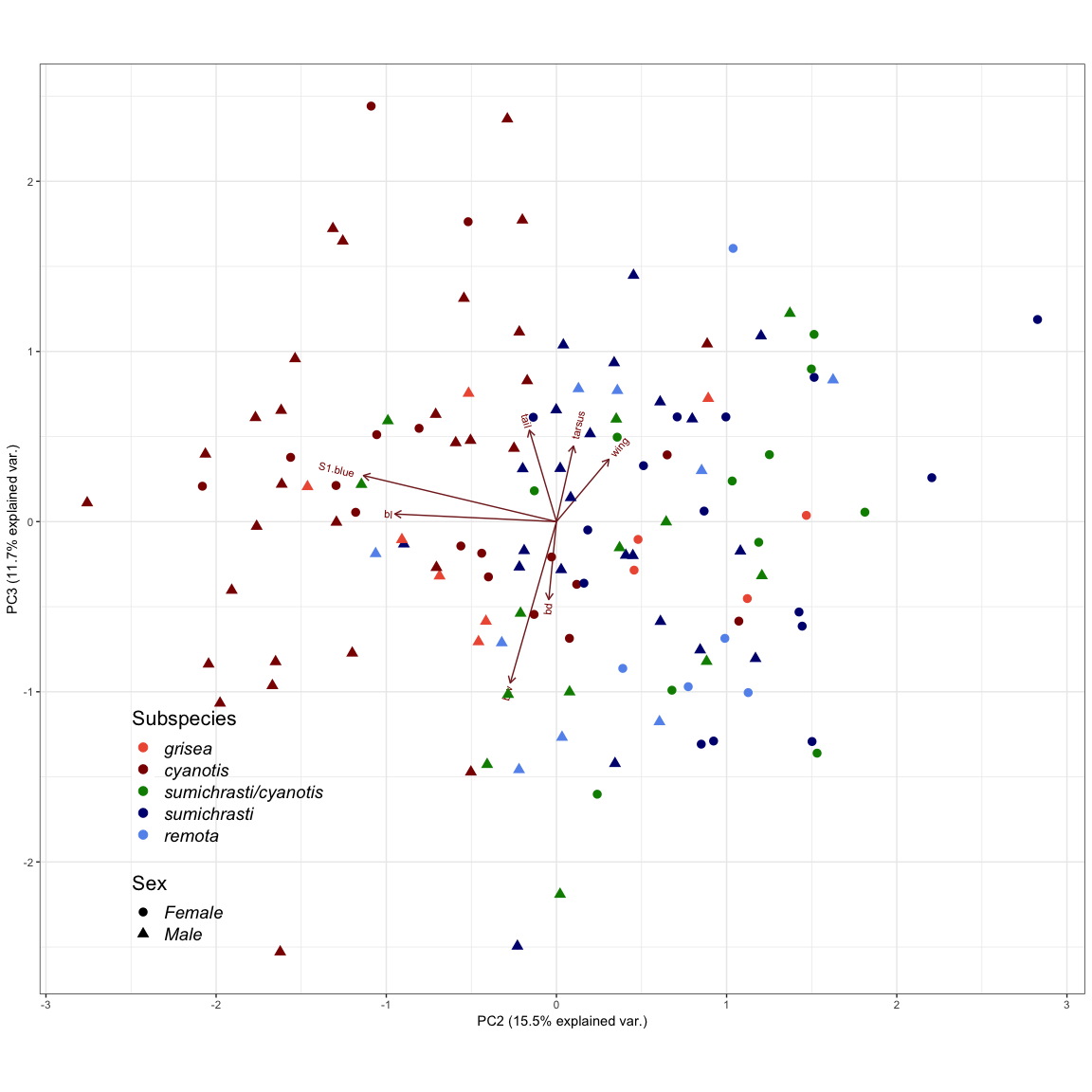
# Figure 3. PCA of 133 individuals based on 6 morphological characteristics and one variable associated with the blue saturation of the back. Individuals from the contact zone (sites E-G) are labeled as sumichrasti/cyanotis.

Plot the same PCA, showing axes 2 & 3

PCAaxestwoandthree<-ggbiplot(pca.scruball, choices = 2:3, obs.scale = 1, var.scale = 1,  
 groups = scrubspecmorph$subspecies, ellipse = FALSE) +  
 geom\_point(aes(colour=scrubspecmorph$subspecies, shape=scrubspecmorph$sex), size = 3)+  
 scale\_color\_discrete(name ="Subspecies") +  
 scale\_color\_manual(values=c("tomato2", "darkred", "green4", "navyblue", "cornflowerblue"),  
 name="Subspecies",  
 limits=c("grisea", "cyanotis", "sumichrasti/cyanotis", "sumichrasti", "remota"))+  
 scale\_shape\_discrete(name = "Sex", labels = c("Female", "Male"))+  
 guides(color = guide\_legend(order=1),  
 shape = guide\_legend(order=2))+  
 theme\_bw()+  
 theme(legend.text=element\_text(size=14, face = "italic"),  
 legend.position =c(.18,.18),  
 legend.background = element\_rect(fill="transparent"),  
 legend.title=element\_text(size=16))

## Scale for 'colour' is already present. Adding another scale for  
## 'colour', which will replace the existing scale.

PCAaxestwoandthree



Including PC axis 3 doesn’t add a ton in terms of meaningfully separating these groups, and we feel comfortable only presenting PCs 1&2 in the manuscript.

Next we visualized the differences in overall body size (PC1) by sex, and blue saturation with the sexes combined.

#separate sexes from the dataframe with PC values added as variables  
malescrub<-subset(PCAscrubspecmorph, sex == "male")  
femalescrub<-subset(PCAscrubspecmorph, sex == "female")  
  
#plot body size for males  
malebodsize<-ggplot(malescrub, aes(x = Comp.1, y = locality, fill = locality, color = locality)) +  
 geom\_density\_ridges(jittered\_points = TRUE, position = "raincloud", alpha = .4) +  
 scale\_color\_manual(values = c("green4", "navyblue", "darkred"), aesthetics = c("fill", "color")) +  
 labs(x = "Male Body Size", y = "Group") +  
 scale\_y\_discrete(limits=c("sumichrasti", "contact", "woodhouse"), labels=c("sumichrast's", "contact", "woodhouse's")) +  
 theme\_classic() +  
 theme(legend.position = "none")  
  
#plot body size for females  
femalebodsize<-ggplot(femalescrub, aes(x = Comp.1, y = locality, fill = locality, color = locality)) +  
 geom\_density\_ridges(jittered\_points = TRUE, position = "raincloud", alpha = .4) +  
 scale\_color\_manual(values = c("green4", "navyblue", "darkred"), aesthetics = c("fill", "color")) +  
 labs(x = "Female Body Size", y = "Group") +  
 scale\_y\_discrete(limits=c("sumichrasti", "contact", "woodhouse"), labels=c("sumichrast's", "contact", "woodhouse's")) +  
 theme\_classic() +  
 theme(legend.position = "none")  
  
#plot back color for all  
colorforall<-ggplot(PCAscrubspecmorph, aes(x = S1.blue, y = locality, fill = locality, color = locality)) +  
 geom\_density\_ridges(jittered\_points = TRUE, position = "raincloud", alpha = .4) +  
 scale\_color\_manual(values = c("green4", "navyblue", "darkred"), aesthetics = c("fill", "color")) +  
 labs(x = "Blue Saturation of the Mantle", y = "Group") +  
 scale\_y\_discrete(limits=c("sumichrasti", "contact", "woodhouse"), labels=c("sumichrast's", "contact", "woodhouse's")) +  
 theme\_classic() +  
 theme(legend.position = "none")

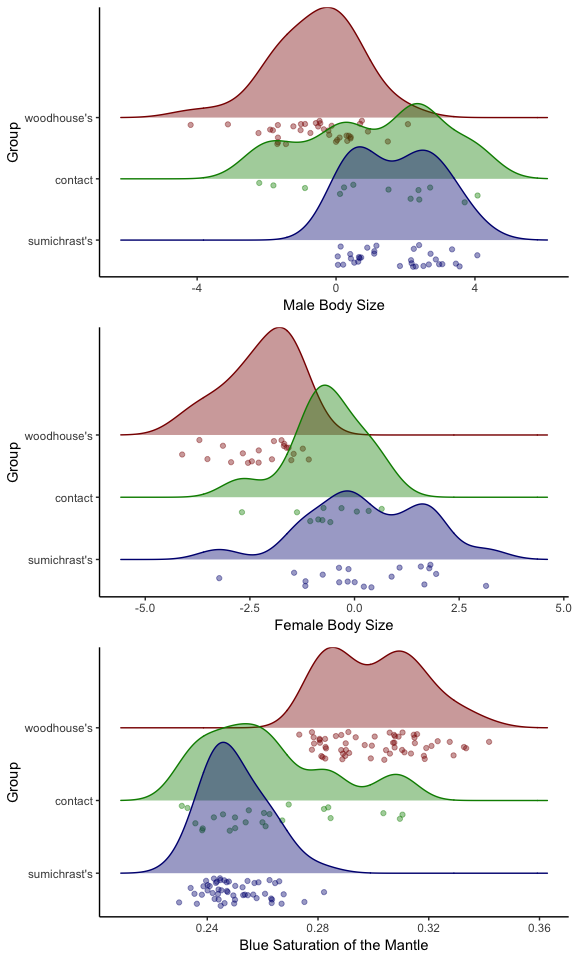
We used grid.arrange to arrange these three figures together. The arrows indicating smaller vs. larger were added in adobe photoshop.

grid.arrange(malebodsize, femalebodsize, colorforall, ncol = 1)

## Picking joint bandwidth of 0.67

## Picking joint bandwidth of 0.488

## Picking joint bandwidth of 0.00702



# Figure 4. Denisty and scatterplots based on PC1 score (general body size) for males and females, and S1B (blue saturation of back plumage) for all individuals from Sumichrast’s and Woodhouse’s groups. Individuals from sites E, F, and G (Fig. 2) are considered the contact zone.