Plover Preen Profiles

Meinolf Ottensmann

2016-08-12

## Intro

This is R Markdown document provides all code necessary code to run the analysis on chemical profiles of preen gland samples of three Madagascan plover species (Kittlitz´s-, Madagascar- and White-fronted). Samples have been processed by gas-chromatography and peaks were called using **Xcalibur 2.0.5**. All further steps are reproducible with the code delivered within this document. The Document is under development and will be updated while running the analysis!

## Bullet points

* Possibility to compare scent profiles of the preen gland with the body swabs: Mantel Test to test for correlations in the similarity/ dissimilarity among individuals based on both scent profiles.
* Comparison of odour diverstiy between both data sets



## Getting started: Load the data and packages to the Global Environment

### Prerequisites

For reproducing the presented code completely, you will need to install **GCalignR** first. The package can be downloaded in the most recent development version from GitHub:

install.packages("devtools") devtools::install\_github("mastoffel/GCalignR",  
build\_vignettes = TRUE)

library(GCalignR) # does the alignment of gc-data, already run before

### Load the raw data

The raw data was already transformed into the working format of GCalignR. In the data set there are six blanks (i.e. controls only filled with dichlormethane, the solvent used) and an additional sample ("w41") containing a septum, that was detached from one of the samples ("w65") was included to control for potential influences of dislodged septums on the chemical profiles. Since a high number of peaks (n=129) was called, the suitability to act as a control seems very limited, hence we will exclude this sample now.

load("data/charadrius\_peaks.RData")  
charadrius\_peaks[["w41"]] <- NULL # remove w41

Lets do the alignment of the gas-chromatography peaks. We use the chromatogram with the highest peak count to apply linear transformations of peak retention times in order to control for systematic temporal shifts in the gas-chromatography run.

charadrius\_peaks\_aligned <- align\_chromatograms(data =charadrius\_peaks ,  
 rt\_col\_name = "rt", # retention time  
 conc\_col\_name = "area", # peak abundance  
 reference = "w62", # sample with the most peaks   
 write\_output = c("rt","area"),   
 blanks = c("w17","w37","w47","w57","w67","w77"),  
 delete\_single\_peak = T, # peaks present in one sample are not informative  
 min\_diff\_peak2peak = 0.03,  
 max\_diff\_peak2mean = 0.03,  
 rt\_cutoff\_low = 8 # peaks before the solvent are treated as uncertain  
 )  
save(charadrius\_peaks\_aligned,file = "data/charadrius\_peaks\_aligned.RData")

## Extract the scent data and load covariates

The lines of code above have been already executed, so we can load the data now. Some tweaking is done to format the data for a easy usage in the following analytical steps.

load("data/charadrius\_peaks\_aligned.RData") # GCalignR output  
# normalise the peak abundancies  
scent <- GCalignR::norm\_peaks(charadrius\_peaks\_aligned, conc\_col\_name = "area",   
 rt\_col\_name = "rt", out = "data.frame")  
scent <- log(scent + 1) # log+1 transformation to reduce mean-variance trends  
factors <- read.csv("data/factors.csv", sep = ";", skip = 1) # Comprises all available covariates  
factors$pair <- paste0(factors$Species, factors$Nest\_Brood) # unique 'pairs'  
factors$pair[is.na(factors$Nest\_Brood)] <- NA # Drop unknown Individuals from 'pairs'  
row.names(factors) <- as.character(tolower(factors$GC\_Sample)) # for cross-reference with gc-data  
scent <- scent[match(row.names(factors), row.names(scent)), ] # same order of rows is crucial   
# The data includes duplicates for three plover individuals  
indices <- which(factors$Ring %in% factors$Ring[duplicated(factors$Ring)]) # duplicated Individuals

pander::pandoc.table(factors[indices, c(2:7, 9, 12, 17, 18, 19)])

Table continues below

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Species | Ring | Nest\_Brood | Observer | Sample\_Date | Brood\_Status |
| **w15** | KIP | FH68752 | 417 | LEP | 404 | Nest |
| **w33** | KIP | FH68899 | 405 | LEP | 322 | Nest |
| **w43** | KIP | FH68752 | 417 | LEP | 409 | Brood |
| **w45** | MP | FH47040 | 101 | OV | 428 | Brood |
| **w53** | KIP | FH68899 | 405 | LEP | 411 | Brood |
| **w64** | MP | FH47040 | 101 | OV | 415 | Brood |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | GC\_Sample | Sex | Age | Lat | Long |
| **w15** | W15 | m | A | 320843 | 7555958 |
| **w33** | W33 | m | A | 320656 | 7555703 |
| **w43** | W43 | m | A | 320843 | 7555958 |
| **w45** | W45 | f | A | NA | NA |
| **w53** | W53 | m | A | 320656 | 7555703 |
| **w64** | W64 | f | A | NA | NA |

Since one MP (FH47040) was sampled twice during brooding, we average the two samples, thereby we avoid pseudo-replication.

samples <- which(factors$Ring == "FH47040") # corresponding sample rows  
scent[samples[1], ] <- unlist(lapply(1:ncol(scent), function(x) {  
 # average the samples  
 if (any(scent[samples, x] > 0)) {  
 scent[samples[1], x] <- mean(scent[samples, x][scent[samples, x] > 0]) # do not average zeros  
 } else {  
 scent[samples[1], x] <- 0  
 }  
}))  
scent <- scent[-samples[2], ] # remove the second sample   
factors <- factors[-samples[2], ]

## Now we can start to analyse patterns using Nonmetric multidimensional scaling (NMDS)

At first we want to load the package **vegan** which provides some useful functionalities

library(vegan) # for metaMDS, adonis, betadisper and simper

#> Loading required package: permute

#> Loading required package: lattice

#> This is vegan 2.3-5

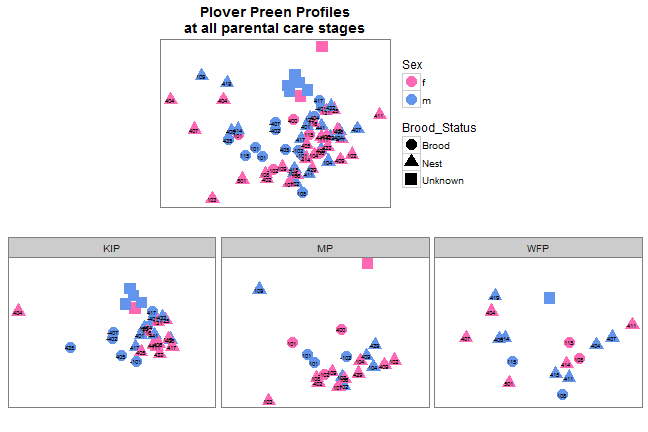
source("R/nmds\_calculator.R") # nmds scaling formatting for plotting using vegan::metaMDS   
library(ggplot2)  
source("R/nmds\_plotter.R") # little function for plotting using ggplot2

Now we can start to do some nmds plots

# We use vegan::metaMDS specifying a bray-curtis dissimilarity matrix and  
# subset the data set a bit  
m1 <- nmds\_calculator(scent = scent, factors = factors, sub = list(Age = "A")) # we focus on adults for now

#> Run 0 stress 0.1511905   
#> Run 1 stress 0.1512651   
#> ... procrustes: rmse 0.002001026 max resid 0.01214789   
#> Run 2 stress 0.1512651   
#> ... procrustes: rmse 0.001996094 max resid 0.01214727   
#> Run 3 stress 0.151279   
#> ... procrustes: rmse 0.002593255 max resid 0.01213925   
#> Run 4 stress 0.1511915   
#> ... procrustes: rmse 0.0001990456 max resid 0.001442473   
#> \*\*\* Solution reached

nmds\_m1 <- m1$nmds # nmds-coordinates + factors  
scent\_m1 <- m1$scent # scent matrix, subsetted  
factors\_m1 <- m1$factors # factors, subsetted   
nmds\_plotter(nmds = nmds\_m1, main = "Plover Preen Profiles\nat all parental care stages") # plot the results



beta1 <- vegan::betadisper(vegan::vegdist(scent\_m1, method = "bray"), factors\_m1$Species)  
anova(beta1) # There is a dispersion effect, i.e. variance is not equal among Species

#> Analysis of Variance Table  
#>   
#> Response: Distances  
#> Df Sum Sq Mean Sq F value Pr(>F)   
#> Groups 2 0.10296 0.05148 3.6357 0.03169 \*  
#> Residuals 67 0.94869 0.01416   
#> ---  
#> Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

TukeyHSD(beta1) # post-hoc test

#> Tukey multiple comparisons of means  
#> 95% family-wise confidence level  
#>   
#> Fit: aov(formula = distances ~ group, data = df)  
#>   
#> $group  
#> diff lwr upr p adj  
#> MP-KIP 0.03286214 -0.04618490 0.1119092 0.5816599  
#> WFP-KIP 0.09736855 0.01078487 0.1839522 0.0238139  
#> WFP-MP 0.06450642 -0.02671858 0.1557314 0.2146520

vegan::adonis(scent\_m1 ~ factors\_m1$Brood\_Status \* factors\_m1$Species)

#>   
#> Call:  
#> vegan::adonis(formula = scent\_m1 ~ factors\_m1$Brood\_Status \* factors\_m1$Species)   
#>   
#> Permutation: free  
#> Number of permutations: 999  
#>   
#> Terms added sequentially (first to last)  
#>   
#> Df SumsOfSqs MeanSqs F.Model  
#> factors\_m1$Brood\_Status 2 2.4516 1.22579 5.2106  
#> factors\_m1$Species 2 2.0532 1.02658 4.3638  
#> factors\_m1$Brood\_Status:factors\_m1$Species 4 1.3398 0.33494 1.4238  
#> Residuals 61 14.3503 0.23525   
#> Total 69 20.1948   
#> R2 Pr(>F)   
#> factors\_m1$Brood\_Status 0.12140 0.001 \*\*\*  
#> factors\_m1$Species 0.10167 0.001 \*\*\*  
#> factors\_m1$Brood\_Status:factors\_m1$Species 0.06634 0.032 \*   
#> Residuals 0.71059   
#> Total 1.00000   
#> ---  
#> Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1