

ELETRONIC SUPPLEMENTARY MATERIAL

Background Noise as a Selective Pressure: Stream-breeding Anurans Call at Higher Frequencies

David Lucas Röhr; Gustavo B. Paterno; Felipe Camurugi; Flora A. Juncá; Adrian A. Garda;

December 17, 2014

Contents

| | |
|--|-----------|
| 1. Packages versions: | 2 |
| 2. Data structure: | 3 |
| 2.1 Species data | 3 |
| 2.2 Phylogentic tree | 3 |
| 2.3 Summary metrics for Dominant frequency and Sout-vent length | 5 |
| 3. Phylogenetic signal | 6 |
| 3.1 Domiant Frequency and Snout-vent length | 6 |
| 3.2 Checking the phylogenetic signal of the residuals from stantard OLS regression | 6 |
| 4. Data analysis | 7 |
| 4.1 Data preparation: | 7 |
| 4.2 Phylogenetic generalized least square model (PGLS) | 7 |
| 4.3 ANOVA table | 8 |
| 4.4 Confidence interval for lambda estimation | 8 |
| 4.5 Model diagnostic: | 9 |
| 4.5.1 Standard graphic methods for model diagnostics | 9 |
| 4.5.2 Phylogenetic signal of model residuals | 9 |
| 4.6 Model comparison: OSL vs PGLS | 10 |
| 5. References | 11 |

This document follows the principles of reproducible science (Peng, 2011). All Data and code required to repeat the analysis bellow are linked at [Github](#). To dowanload the **source code** used to generate all figures, tables and analysis in the paper, please see: [source code](#). This document was generated in **R studio** with **kintr** package.

1. Packages versions:

We used R version 3.1.2 (2014-10-31) and the following packages:

```
library(ape);library(caper);library(knitr)
library(dplyr);library(ggplot2);library(picante);library(gridExtra)
```

Please check the [Packages versions](#), for details.

2. Data structure:

2.1 Species data

- To download **raw data**: [link](#).
- To download **pdf table** with data and references used: [link](#)

The species dataset contains six variables (see Methods for detailed information on data collection).

| variable | discription |
|-------------|--|
| sp | species name |
| environment | Reproductive environment (lentic or lotic) |
| DF | Dominant frequency (hertz) |
| SVL | snout-vent length (mm) |
| logDF | log10 of dominant frequency (DF) |
| logSVL | log10 of snout vent length (SVL) |

Last six rows of the species dataset:

| | sp | environment | DF | SVL | logDF | logSVL |
|-----|---------------------|-------------|------|-----|----------|----------|
| 504 | Bufo_retiformis | still | 3113 | 47 | 3.493179 | 1.672098 |
| 505 | Bufo_houstonensis | still | 2151 | 77 | 3.332640 | 1.886491 |
| 506 | Pelophryne_misera | still | 4000 | 21 | 3.602060 | 1.322219 |
| 507 | Ansonia_longidigita | running | 3500 | 50 | 3.544068 | 1.698970 |
| 508 | Ansonia_hanitschi | running | 5700 | 32 | 3.755875 | 1.505150 |
| 509 | Ansonia_platysoma | running | 8000 | 25 | 3.903090 | 1.397940 |

2.2 Phylogenetic tree

The phylogenetic tree used in this paper was pruned from: [Pyron and Wiens \(2014\)](#) anura super tree. To dowanload the pruned tree with study species (509): Study Tree.

```
##
## Phylogenetic tree with 509 tips and 508 internal nodes.
##
## Tip labels:
## Hadromophryne_natalensis, Heleophryne_purcellii, Heleophryne_regis, Calyptocephallela_gayi, Neobatra
## Node labels:
## 209.59, 206.29, 195.63, 151.99, 47.93, 9.02, ...
##
## Rooted; includes branch lengths.
```

- running
- still

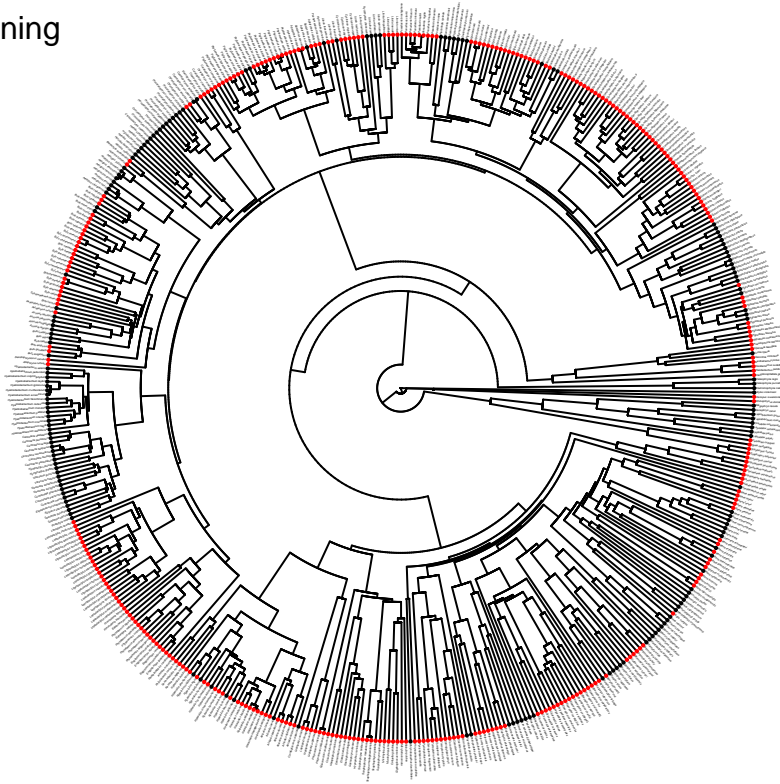


Figure S1: Phylogeny for 509 anuran species sampled in this study extracted from Pyron and Wiens (2011) original tree. Black circles represent pond-breeding species ($N = 332$) and red circles stream-breeding species ($N = 177$)

2.3 Summary metrics for Dominant frequency and Sout-vent length

| environment | meanDF | seDF | meanSVL | seSVL |
|-------------|----------|----------|----------|----------|
| running | 3377.322 | 2036.938 | 41.70932 | 20.03995 |
| still | 2180.557 | 1263.536 | 51.27440 | 29.35337 |

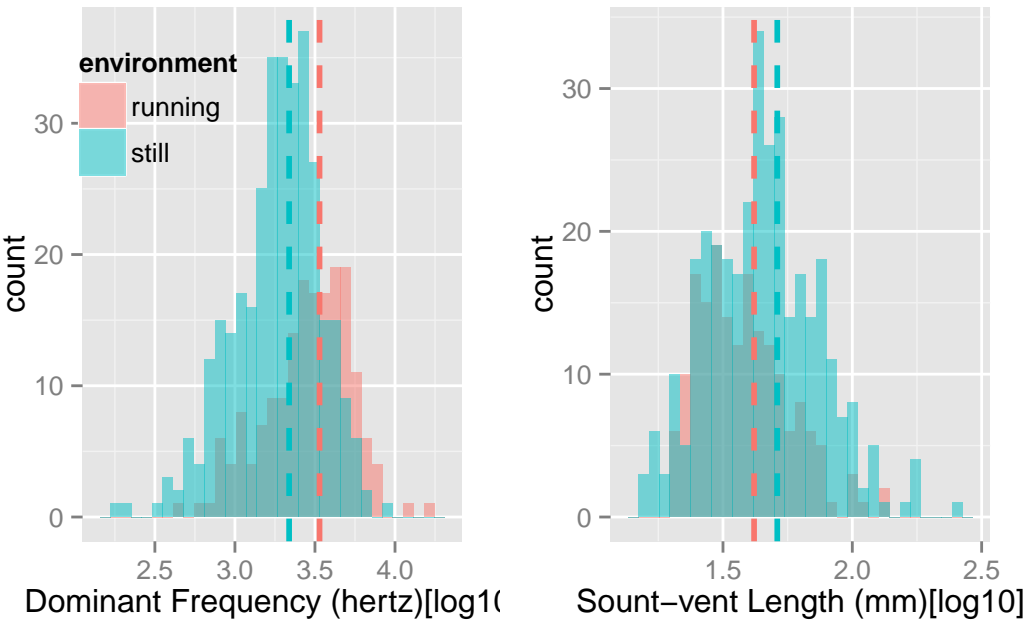


Figure S2: Distribution histograms for Dominant Frequency (log10) and Snout-vent length (log10)

3. Phylogenetic signal

3.1 Dominant Frequency and Snout-vent length

We used K statistics to test the phylogenetic signal for Dominant Frequency ($\log DF$) and Snout-vent length ($\log SVL$) (for details about the method, see [Blomberg et al \(2003\)](#)).

```
k.signal <- multiPhylosignal(select(comp.data$data, logDF, logSVL), comp.data$phy, reps=999)
kable(k.signal)
```

| | K | PIC.variance.obs | PIC.variance.rnd.mean | PIC.variance.P | PIC.variance.Z |
|--------|-----------|------------------|-----------------------|----------------|----------------|
| logDF | 0.3660099 | 0.0014224 | 0.0055201 | 0.001 | -4.411998 |
| logSVL | 0.4387480 | 0.0006061 | 0.0028041 | 0.001 | -4.527165 |

Dominant Frequency and Snout-Vent Length show significant phylogenetic signal, however, K values are low.

3.2 Checking the phylogenetic signal of the residuals from standard OLS regression

In order to check the need to include the phylogeny in our analysis, first it is important to check if there is phylogenetic signal in the residuals of an Ordinary Least Square regression (OLS) ([Kamilar & Cooper, 2013](#); [Freckleton, 2009](#)).

```
mod.osl <- lm(logDF ~ environment*logSVL, anura.data)
# Extracting residuals from the model:
comp.data$data$lm.res <- residuals(mod.osl)
osl.resi.sig <- phylosignal(comp.data$data$lm.res, reps=999, comp.data$phy)
kable(osl.resi.sig)
```

| | K | PIC.variance.obs | PIC.variance.rnd.mean | PIC.variance.P | PIC.variance.Z |
|--|-----------|------------------|-----------------------|----------------|----------------|
| | 0.1472979 | 0.0018287 | 0.0027465 | 0.003 | -1.721673 |

Because the residuals from OLS regression show phylogenetic signal $k = 0.15$, it is necessary to correct for phylogenetic non-independence in data.

4. Data analysis

We used a phylogenetic generalized least square model (PGLS) with dominant frequency as the response variable and reproduction habitat (lentic/lotic) and SVL as the explanatory variables to test if dominant frequency was affected by reproduction environment. Dominant frequencies and body sizes were log transformed (log10) before the analysis. To optimize branch length transformation, the lambda value was set by maximum likelihood (see [Freckleton et al., 2002](#); [Orme et al., 2013](#) for details). PGLS analysis were performed with the function `pgls` from the package `caper`.

4.1 Data preparation:

Using the function `comparative.data` we combined our phylogenie with the species dataset

```
## Comparative dataset of 509 taxa:
## Phylogeny: study.tree
##   509 tips, 508 internal nodes
## chr [1:509] "Hadromophryne_natalensis" "Heleophryne_purcelli" ...
## VCV matrix present:
## VCV.array [1:509, 1:509, 1:25] 47.9 3.3 3.3 3.3 3.3 ...
## Data: anura.data
## $ environment: Factor w/ 2 levels "running","still": 1 1 1 2 2 2 2 2 1 2 ...
## $ DF          : int [1:509] 1500 2250 1800 866 816 1509 1300 2676 2735 2260 ...
## $ SVL         : num [1:509] 45 47 43 120 39.6 40 55 27 25.2 35 ...
## $ logDF       : num [1:509] 3.18 3.35 3.26 2.94 2.91 ...
## $ logSVL      : num [1:509] 1.65 1.67 1.63 2.08 1.6 ...
```

4.2 Phylogenetic generalized least square model (PGLS)

Fitting `pgls` model with with lambda adjusted by maximum likelihood:

```
mod.pgls <- pgls(logDF ~ environment*logSVL, data=comp.data,lambda="ML")
summary(mod.pgls)
```

```
##
## Call:
## pgls(formula = logDF ~ environment * logSVL, data = comp.data,
##       lambda = "ML")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.073524 -0.014980 -0.001757  0.010458  0.060338
##
## Branch length transformations:
##
## kappa [Fix] : 1.000
## lambda [ ML] : 0.889
## lower bound : 0.000, p = < 2.22e-16
## upper bound : 1.000, p = < 2.22e-16
## 95.0% CI    : (0.823, 0.933)
## delta [Fix] : 1.000
##
```

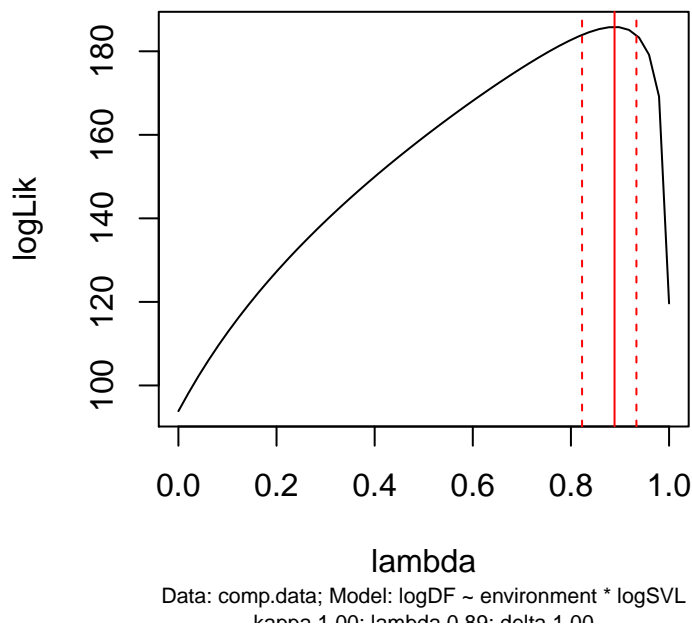
```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.874025   0.190792  25.5462  <2e-16 ***
## environmentstill -0.182756   0.173309  -1.0545   0.2922
## logSVL         -0.918595   0.093136  -9.8630  <2e-16 ***
## environmentstill:logSVL 0.060128   0.104687   0.5744   0.5660
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02088 on 505 degrees of freedom
## Multiple R-squared:  0.3825, Adjusted R-squared:  0.3788
## F-statistic: 104.3 on 3 and 505 DF, p-value: < 2.2e-16
```

4.3 ANOVA table

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|--------------------|-----|-----------|-----------|-------------|-----------|
| environment | 1 | 0.0144038 | 0.0144038 | 33.0318903 | 0.0000000 |
| logSVL | 1 | 0.1218527 | 0.1218527 | 279.4422042 | 0.0000000 |
| environment:logSVL | 1 | 0.0001428 | 0.0001428 | 0.3274965 | 0.5673916 |
| Residuals | 505 | 0.2202086 | 0.0004361 | NA | NA |

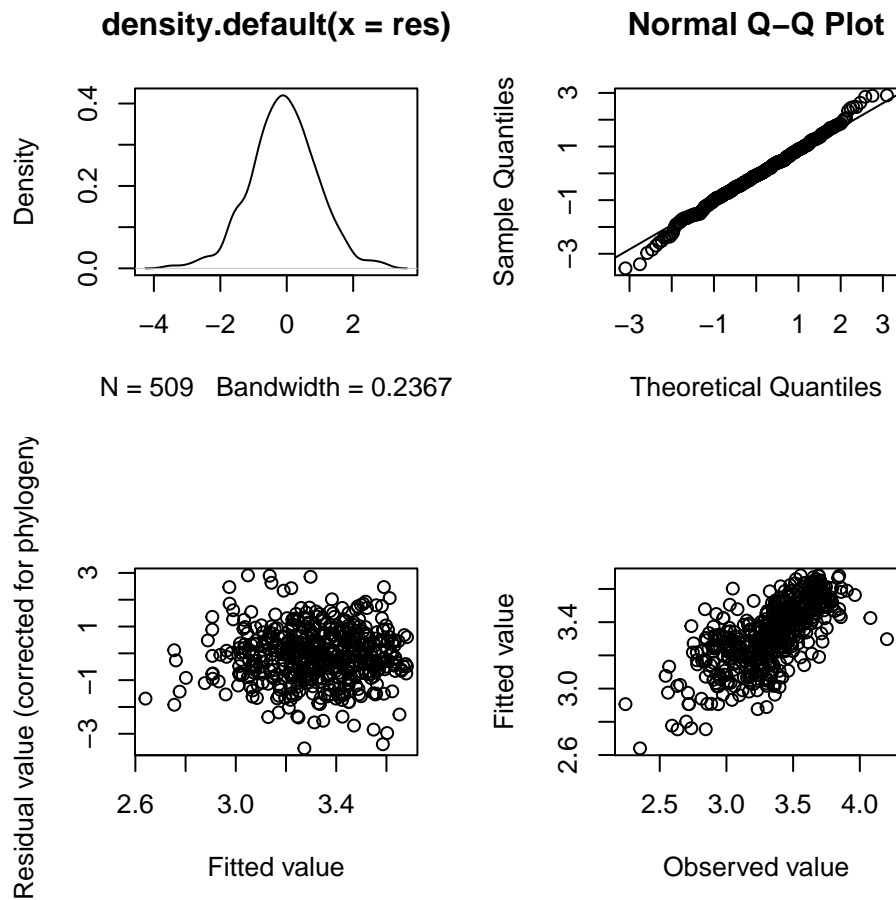
4.4 Confidence interval for lambda estimation

```
profile.lambda <- pgls.profile(mod.pgls)
plot(profile.lambda)
```



4.5 Model diagnostic:

4.5.1 Standard graphic methods for model diagnostics



Residuals do not show any tendency.

4.5.2 Phylogenetic signal of model residuals

After performing PGLS analysis it is important to check the phylogenetic signal of model residuals.

```
k.residuals <- phylosignal(mod.pgls$phyres, reps=999, comp.data$phy)
kable(k.residuals)
```

| K | PIC.variance.obs | PIC.variance.rnd.mean | PIC.variance.P | PIC.variance.Z |
|-----------|------------------|-----------------------|----------------|----------------|
| 0.1119063 | 2.39e-05 | 2.94e-05 | 0.105 | -1.052582 |

Results above shows that the residuals do not present significant phylogenetic signal.

4.6 Model comparison: OSL vs PGLS

```
kable(AIC(mod.osl,mod.pgl))
```

| | df | AIC |
|---------|----|-----------|
| mod.osl | 5 | -177.6885 |
| mod.pgl | 4 | -363.6674 |

AIC comparison shows that PGLS model has much lower AIC value $\text{round}(-363.6674412)$ than OSL model $\text{round}(-177.6884567)$. Thus, PGLS model is a better fit for the data.

5. References

1. Pyron, R.A. & Wiens, J.J. 2011 A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* 61, 543-583.
2. Blomberg, S. P., T. Garland, Jr., and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717-745.
3. David Orme, Rob Freckleton, Gavin Thomas, Thomas Petzoldt, Susanne Fritz, Nick Isaac and Will Pearse (2013). caper: Comparative Analyses of Phylogenetics and Evolution in R. R package version 0.5.2. <http://CRAN.R-project.org/package=caper>
4. Kamilar JM, Cooper N. 2013 Phylogenetic signal in primate behaviour, ecology and life history. *Phil Trans R Soc B* 368: 20120341. <http://dx.doi.org/10.1098/rstb.2012.0341>
5. Freckleton, R.P. (2009) The seven deadly sins of comparative analysis. *J Evol Biol*, 22, 1367-1375.

When using the **data available** in this paper, please cite the original publication.
Contact davidlucasr@yahoo.com.br for any further information.