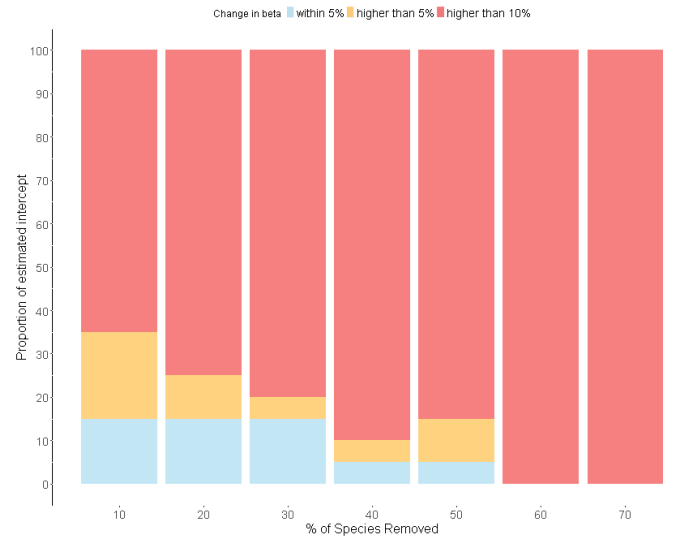




SensiPhy:

an R-package for sensitivity analysis of phylogenetic comparative methods

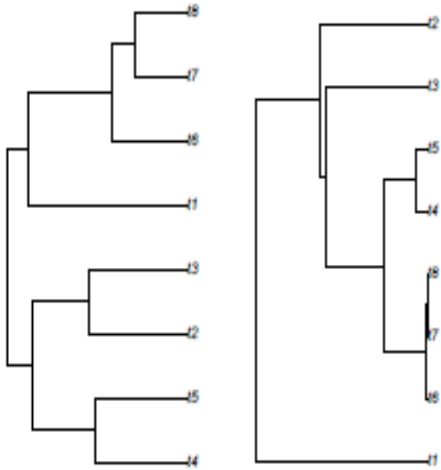


Gijsbert Werner, Caterina Penone and Gustavo Paterno



UNIVERSITY OF
OXFORD

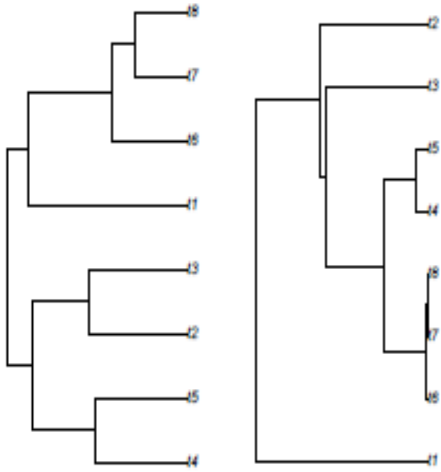
Sources of uncertainty in comparative methods



1. Phylogenetic



Sources of uncertainty in comparative methods



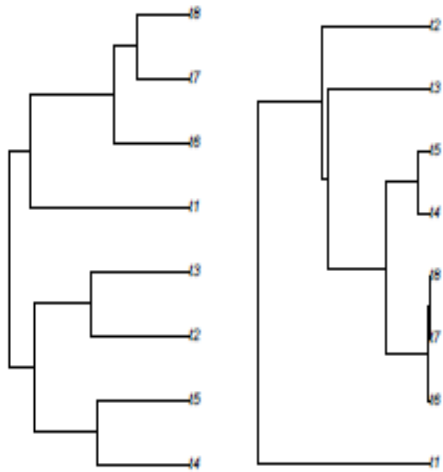
1. Phylogenetic



2. Data



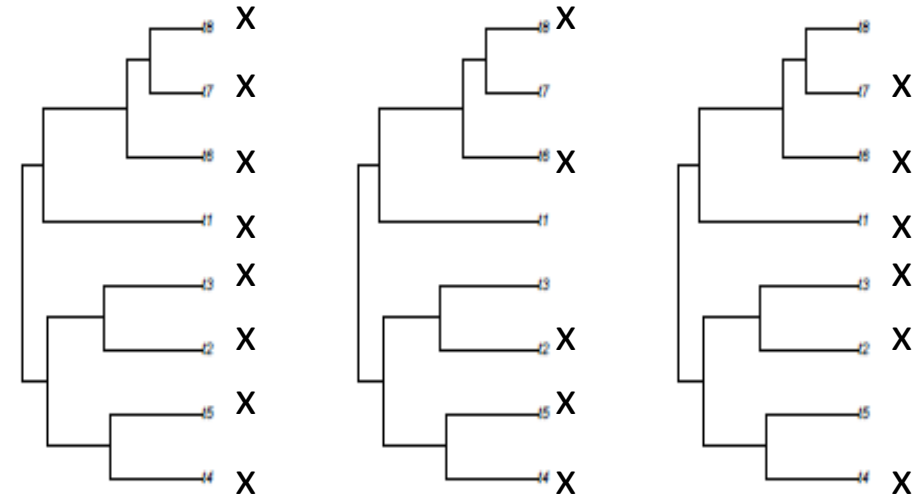
Sources of uncertainty in comparative methods



1. Phylogenetic



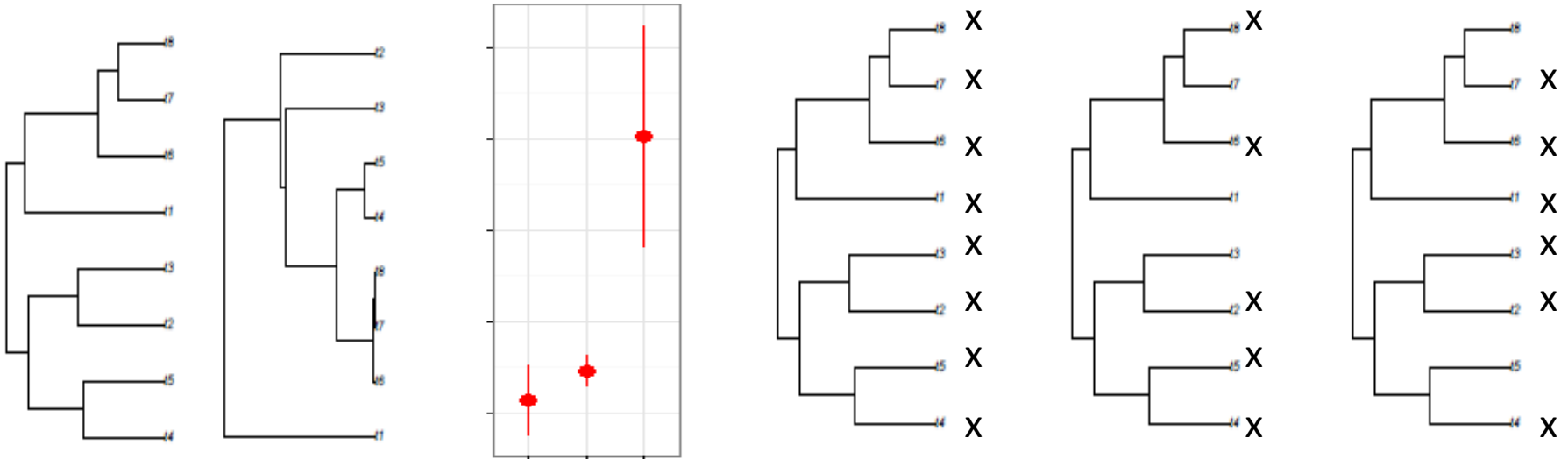
2. Data



3. Sampling / species



How robust are your results?



1. Phylogenetic

2. Data

3. Sampling / species

How robust are your conclusions?

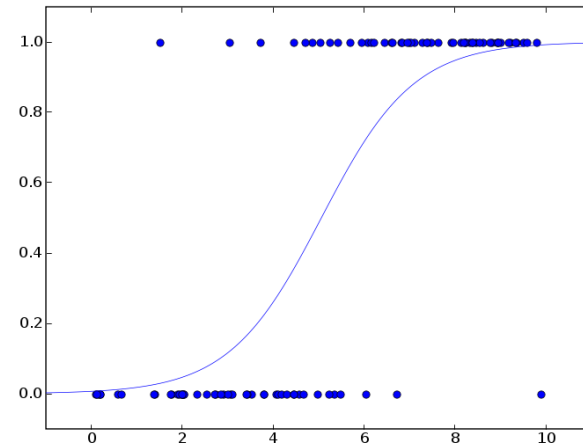
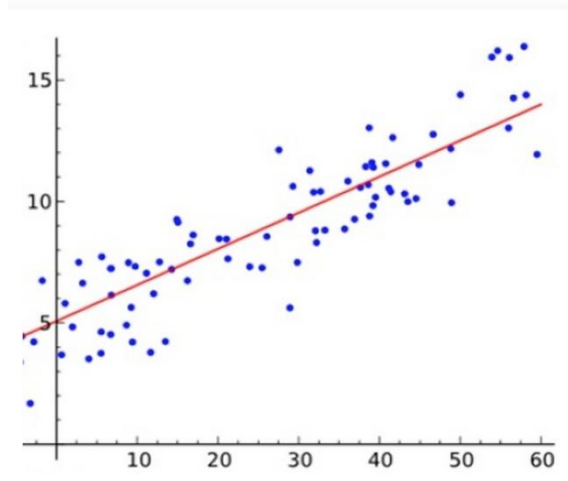
- Affected by phylogeny?
- By intraspecific variability?
- By sampling & influential species/clades?



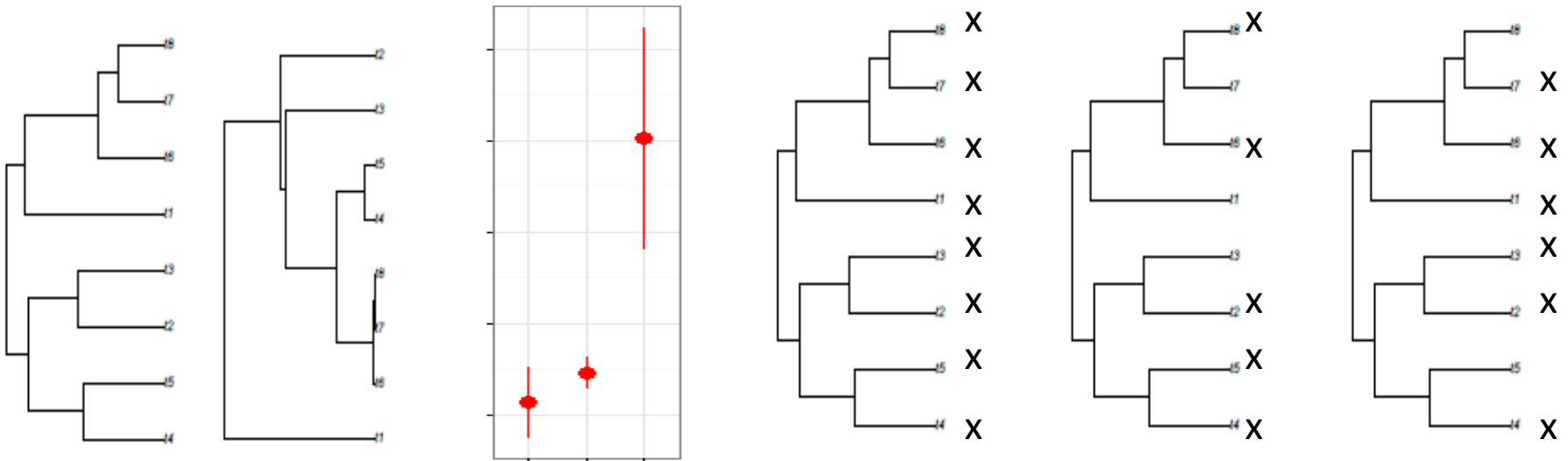
SensiPhy – R package to

Simulation approaches for phylogenetic linear and logistic regression

- Simple & Intuitive to use
- Quick check of robustness
- Using '*phylolm*' (Ho & Ane, 2014)



Functions to assess sensitivity



1. Phylogenetic

2. Data

3. Sampling / species

Prefix + Phylolm functions `phylolm` (linear) and `phyloglm` (logistic regression)

`tree`

`intra`

`samp`

`influ`

`clade`

i.e. `tree_phylm`

For visualisation `sensi_plot` (output)



Assess sampling effects

Randomly remove species -> refit model -> repeat
times: number of repeats, breaks: removal percentages

```
samp <- samp_phylm(log(mass) ~ log(gesta), phy = tree, data = dat,  
  model = "lambda", times = 20, breaks = seq(0.1, 0.7, 0.1) )
```



Assess sampling effects

Randomly remove species -> refit model -> repeat
times: number of repeats, breaks: removal percentages

```
samp <- samp_phylm(log(mass) ~ log(gesta), phy = tree, data = dat,  
  model = "lambda", times = 100, breaks = seq(0.1, 0.7, 0.1) )
```

	n.remov	n.percent	intercept	DFintercept	intercept.perc	pval.intercept	slope	DFslope	slope.perc	pval.slope	AIC	optpar
1	8	10	-1.0320955	-0.17210480	20.0	0.4792661	2.504122	0.025442900	1.0	7.268549e-12	222.9548	0.9597284
2	8	10	-1.0099905	-0.14999982	17.4	0.5021995	2.528157	0.049477693	2.0	2.614550e-11	220.9238	0.9627048
3	8	10	-0.9084772	-0.04848654	5.6	0.5533478	2.497190	0.018511117	0.7	7.556465e-11	219.0675	0.9690649
4	8	10	-0.9671268	-0.10713607	12.5	0.5629012	2.472081	-0.006597702	0.3	6.210014e-10	223.3468	0.9788589
5	8	10	-1.1843418	-0.32435111	37.7	0.3805827	2.537420	0.058741260	2.4	4.006307e-13	217.9968	0.9374231
6	8	10	-0.4992263	0.36076439	41.9	0.7484574	2.366226	-0.112453363	4.5	2.399611e-10	216.6453	0.9889346

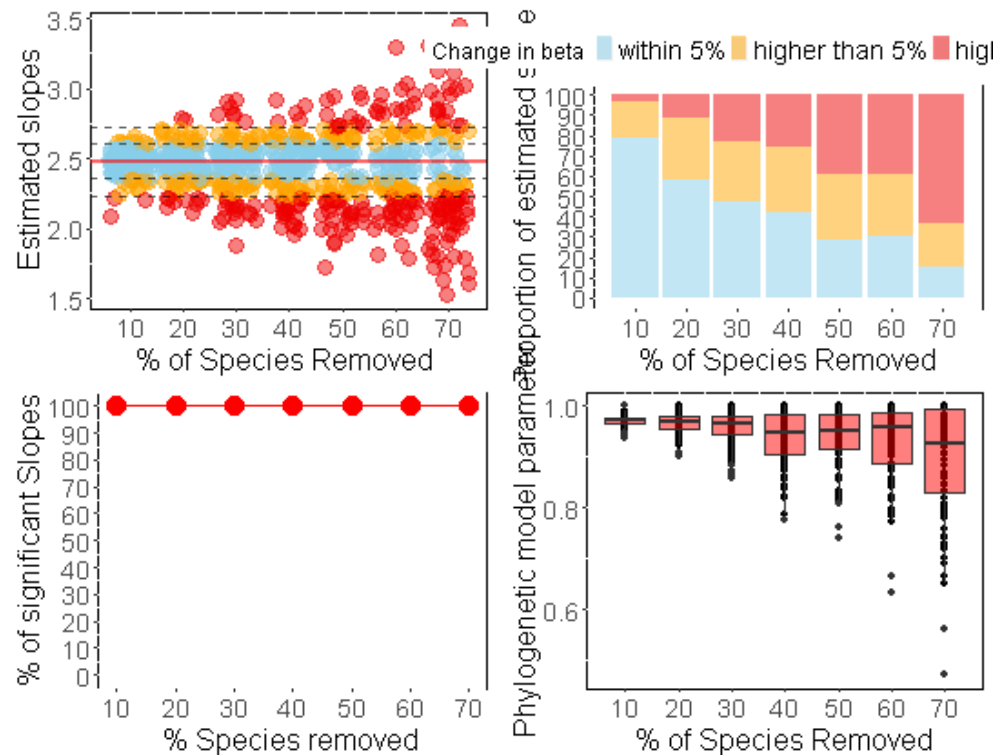


Assess sampling effects

Randomly remove species -> refit model -> repeat
times: number of repeats, breaks: removal percentages

```
samp <- samp_phylm(log(mass) ~ log(gesta), phy = tree, data = dat,  
  model = "lambda", times = 100, breaks = seq(0.1, 0.7, 0.1) )
```

```
sensi_plot(samp)
```

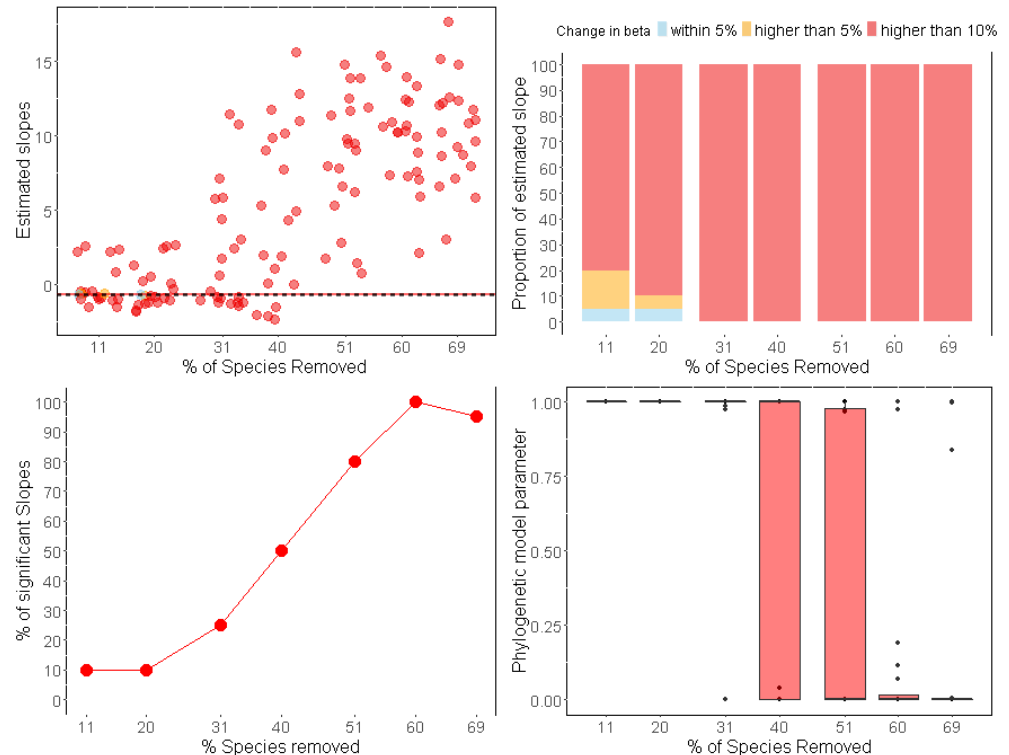


Assess sampling effects

Randomly remove species -> refit model -> repeat
times: number of repeats, breaks: removal percentages

```
samp <- samp_phylm(log(mass) ~ log(gesta), phy = tree, data = dat,  
  model = "lambda", times = 100, breaks = seq(0.1, 0.7, 0.1) )
```

```
sensi_plot(samp)
```



Detect influential species

Leave-one-out deletion -> refit model

```
influ <- influ_phylm(formula=log(mass) ~ log(gesta),  
  phy = tree, data = dat, model = "lambda", cutoff = 2, track = T)
```



Detect influential species

Leave-one-out deletion -> refit model

```
influ <- influ_phylm(formula=log(mass) ~ log(gesta),  
  phy = tree, data = dat, model = "lambda", cutoff = 2, track = T)
```

```
> summary(influ)
```

```
$`Influential species for the slope`
```

```
[1] "Sorex_cinereus"           "Mustela_sibirica"           "Daubentonia_madagascariensis"  
[4] "Mustela_erminea"
```

```
$`slope Estimates`
```

	Species removed	slope	DFslope	Change(%)	Pval
1	Sorex_cinereus	2.271780	-0.20689858	8.3	1.360877e-10
2	Mustela_sibirica	2.580444	0.10176484	4.1	5.242595e-12
3	Daubentonia_madagascariensis	2.564122	0.08544278	3.4	4.263868e-12
4	Mustela_erminea	2.558313	0.07963412	3.2	1.952081e-12

```
$`Influential species for the Intercept`
```

```
[1] "Sorex_cinereus" "Mustela_sibirica" "Mustela_erminea"
```

```
$`Intercept Estimates`
```

	Species removed	Intercept	DFintercept	Change(%)	Pval
1	Sorex_cinereus	-0.04665979	0.8133309	94.6	0.9743078
2	Mustela_sibirica	-1.21992463	-0.3599339	41.9	0.4212079
3	Mustela_erminea	-1.13910524	-0.2791145	32.5	0.4387456

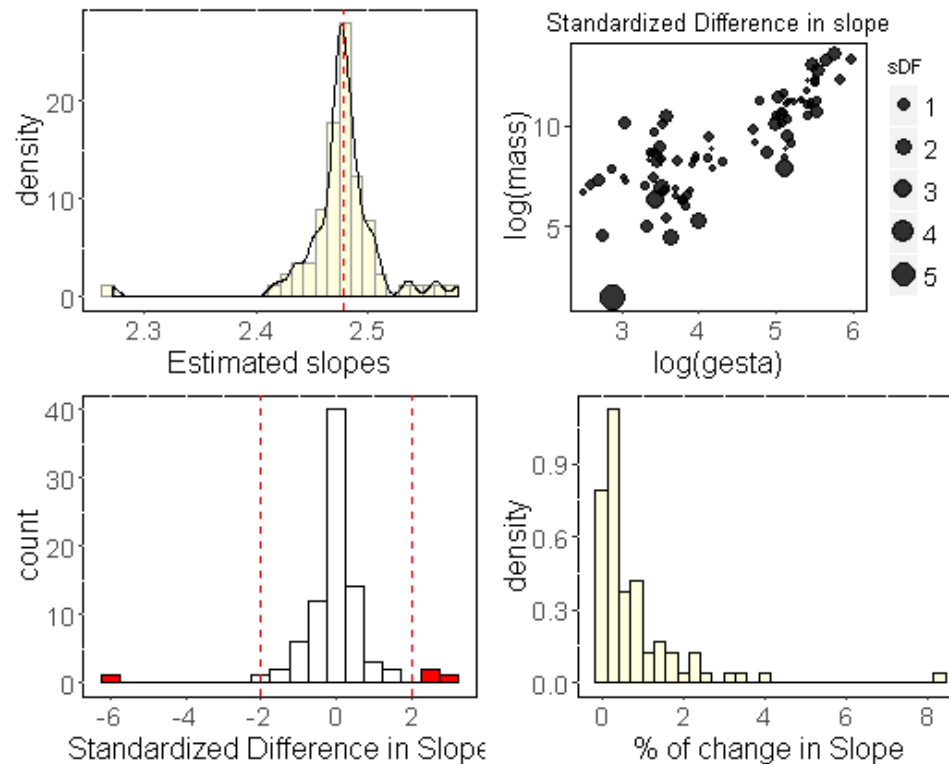


Detect influential species

Leave-one-out deletion -> refit model

```
influ <- influ_phylm(formula=log(mass) ~ log(gesta),  
  phy = tree, data = dat, model = "lambda", cutoff = 2, track = T)
```

```
sensi_plot(influ)
```



Where to find?



Install sensiPhy from github:

```
devtools::install_github("paternogbc/sensiPhy")
```

More information, tutorial, examples:

<https://github.com/paternogbc/sensiPhy>

gijsbert.werner@zoo.ox.ac.uk /

caterina.penone@gmail.com / paternogbc@gmail.com



Where to find?



Install sensiPhy from github:

```
devtools::install_github("paternogbc/sensiPhy")
```

More information, tutorial, examples:

<https://github.com/paternogbc/sensiPhy>

gijsbert.werner@zoo.ox.ac.uk /

caterina.penone@gmail.com / paternogbc@gmail.com



Caterina Penone
University of Bern
Switzerland



Gustavo Paterno
Universidade Federal do Rio
Grande do Norte, Brazil

