Introduction to FestR

Kevin Healy, Thomas Guillerme & Andrew L Jackson

6 April 2016

This package is based on the `MCMCglmm` package and runs a MCMCglmm analysis on multiple trees using the `MulTree` package.

Installation

To install `FestR` its dependancy `MulTree` package must first be installed directly from GitHub using the following:

```r
if(!require(devtools)) install.packages("devtools")
install_github("TGuillerme/mulTree", ref = "master")
```

Follwong this you can then install `FestR` directly from GitHub using the following:

```r
install_github("healyke/FestR", ref = "master")
```

And load in the data

```r
library(mulTree)
library(FestR)
```

Read in the data

```r
#read in the data
mydata<-read.csv(file=system.file("extdata", "FestR_data.csv", package = "FestR"),
 header=T, stringsAsFactors = F)

# re-oprder it into species
mydata <- mydata[order(mydata$species),]

#read in the phylogentic information
#first the mammal trees
mammal_trees <- read.tree(system.file("extdata", "3firstFritzTrees.tre", package = "FestR"))
#than the bird trees
bird_trees   <- read.tree(system.file("extdata", "3firstJetzTrees.tre", package = "FestR"))

#combine them together using the tree.bind function from the mulTree package
combined_trees<-tree.bind(x = mammal_trees, y = bird_trees, sample = 2, root.age = 250)
```

Testing the new data: `set.tef.est`

In order to estimate a trophic enrichment factor for a new species we need to check that the species is already present in our phylogeny and check what data is available for the new species.

This data set.tef.est checks for includes:

tissue type ("blood", "claws", "collagen", "feather", "hair", "kidney", "liver", "milk", "muscle"),

habitat ("terrestial", "marine"),

diet.type ("carnivore", "herbivore", "omnivore", "pellet"),

and the isotopic ratios of the food sources for carbon

(source.iso.13C)

and nitrogen

(source.iso.15N).

```
#####function that checks the dat for some species we want to estimate TEF for
new.data.test <- setTefEst(species = "Meles_meles",
                           habitat = "terrestrial",
                           taxonomic.class = "mammalia",
                           tissue = "blood",
                           diet.type = "omnivore",
                           source.iso.13C = c(-24.1),
                           source.iso.15N = c(7.0),
                           tree = combined_trees)
```

```
## Meles_meles present in phylogeny
```

If the species is not in the phylogeny already such as the Komodo dragon (Varanus komodoensis), or we are missing on of the values, in this case diet.type we get warning messages to indicate what is missing from our data.

```
## Varanus_komodoensis not present in phylogeny
```

Formating the new data: `tef.mul.clean`

We now need to format the data by combining both the isotpic data already avialbe within the package and the data from the new species and matching these species to the included phylogeny. We also include what isotope we want to estimate a trophic discrimination value for (either "carbon" or "nitrogen").

```
tef_data_c <- tefMulClean(new.data = new.data.test, data = mydata, species.col.name = "species", tree =
```

```
## The following taxa were dropped from the analysis:
##  Tachyglossus_aculeatus Zaglossus_bruijni Zaglossus_bartoni Ornithorhynchus_anatinus Anomalurus_beec:
```

We now have a mulTree class object, which is required by the imputation analysis, that contains the matched phylogeny and a dataset containting isotpic data with the new species for which you want to estimate a trophic enrichment factor at the top with a NA for either delta13C or delta15N depending on isotope.

```
tef_data_c$phy
```

```
## 2 phylogenetic trees
```

```r
head(tef_data_c$data)
```

```
##                   sp.col     habitat taxonomic.class tissue diet.type
## 1           Meles_meles terrestrial        mammalia  blood  omnivore
## 130          Bos_taurus terrestrial        mammalia   hair herbivore
## 94   Callorhinus_ursinus     marine        mammalia  blood carnivore
## 95   Callorhinus_ursinus     marine        mammalia  blood carnivore
## 132         Capra_hircus terrestrial        mammalia   hair herbivore
## 31       Enhydra_lutris     marine        mammalia   hair carnivore
##      source.iso.13C delta13C              animal
## 1             -24.1       NA         Meles_meles
## 130           -27.0      2.7          Bos_taurus
## 94            -18.9      0.2 Callorhinus_ursinus
## 95            -18.9      1.0 Callorhinus_ursinus
## 132           -27.0      3.2         Capra_hircus
## 31            -17.8      2.8       Enhydra_lutris
```

Running the analysis: `tef.mul.clean`

With the data formated as a mulTree object we now need to set up the model. In this case we will run the full model to esstimate delta13C with the fixed factors of source isotpic ratio, diet type and habitat type.

```r
formula.c <- delta13C ~ source.iso.13C + diet.type + habitat
```

and random terms that includes the "animal" term which is required to include phylogeny into the analysis, sp.col to allow multiple species entries into the analysis and account for variation within a species, and tissue type.

```r
random.terms = ~ animal + sp.col + tissue
```

As we rely on Bayesian imputation to estimate the missing value we also need to speciafy a prior, in this case we follow the standard prior from MCMCglmm

```r
prior_tef <- list(R = list(V = 1/4, nu=0.002), G = list(G1=list(V = 1/4, nu=0.002),
                 G2=list(V = 1/4, nu=0.002), G3=list(V = 1/4, nu=0.002)))
```

finally we can run the analysis

```r
Tef_est.c <- tefMcmcglmm(mulTree.data = tef_data_c,
                  formula = formula.c,
                  random.terms = random.terms,
                  prior = prior_tef,
                  output = "test_c_run",
                  nitt = c(1200000),
                  thin = c(500),
                  burnin = c(200000),
                  no.chains = c(2),
                  convergence =  c(1.1),
                  ESS = c(1000))
```

```
## Output chain name "test_c_run" already exists!
## Press [enter] if you wish to overwrite the models or [esc] to cancel.
```

```
## Models will be overwritten...

##
## 2016-04-07 - 10:30:09: MCMCglmm performed on tree 1
## Convergence diagnosis:
## Effective sample size is > 1000: TRUE
## 2000; 2000; 1745.184; 2000; 2000; 2000; 2165.34; 2184.666; 2198.561; 2000; 2388.605; 1931.708
## All levels converged < 1.1: TRUE
## 1.001197; 1.000068; 1.000801; 1.000428; 1.000435; 1.000134
## Individual models saved as: test_c_run-tree1_chain*.rda
## Convergence diagnosis saved as: test_c_run-tree1_conv.rda
##
## 2016-04-07 - 10:34:57: MCMCglmm performed on tree 2
## Convergence diagnosis:
## Effective sample size is > 1000: TRUE
## 2230.65; 2000; 2073.291; 1664.955; 2000; 2358.108; 1822.325; 1792.695; 2177.565; 2149.792; 2000; 2000
## All levels converged < 1.1: TRUE
## 1.00063; 1.001684; 1.000417; 0.9996571; 1.001278; 0.9996844
## Individual models saved as: test_c_run-tree2_chain*.rda
## Convergence diagnosis saved as: test_c_run-tree2_conv.rda
##
## 2016-04-07 - 10:34:57: MCMCglmm successfully performed on 2 trees.
## Total execution time: 9.273651 mins.
## Use read.mulTree() to read the data as 'mulTree' data.
## Use summary.mulTree() and plot.mulTree() for plotting or summarizing the 'mulTree' data.
```

explain what is going on here

```r
hist(Tef_est.c$tef_global)
```

Histogram of Tef_est.c$tef_global
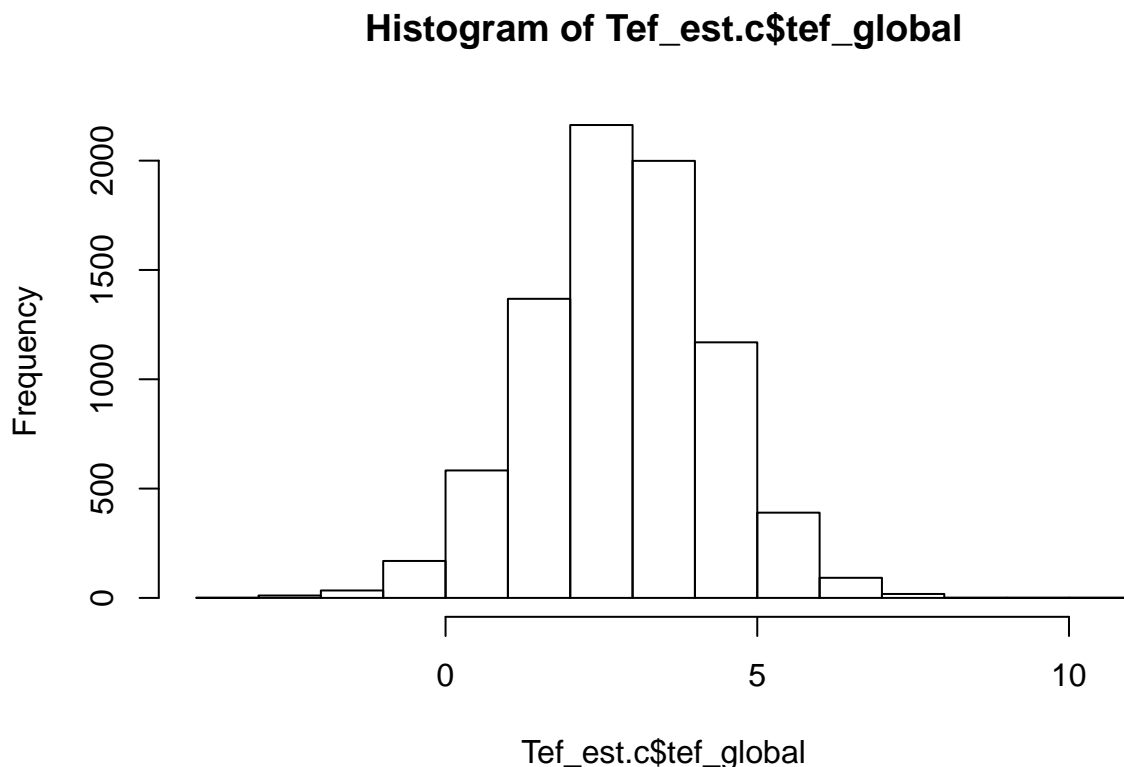
```r
median(Tef_est.c$tef_global)
```

```
## [1] 2.845063
```