Readme

Alice Ledda

June 22, 2016

What you can do so far...

1 Required packages

You need to install the R package `ape` and `phangorn`

```
> library(ape)
> library(phangorn)
```

The files needed to run the whole thing are in the following files. The required file fore each command is specified in the command example.

The files containing the scripts can be found github at the following address

GITHUB ADDRESS

2 Simulate genealogies

We have successfully implemented the simulation of genealogies. We can simulate genealogies with constant population size, with expanding population size and with one constant population with one expanding subpopulation (just one of each).

2.1 Constant population size

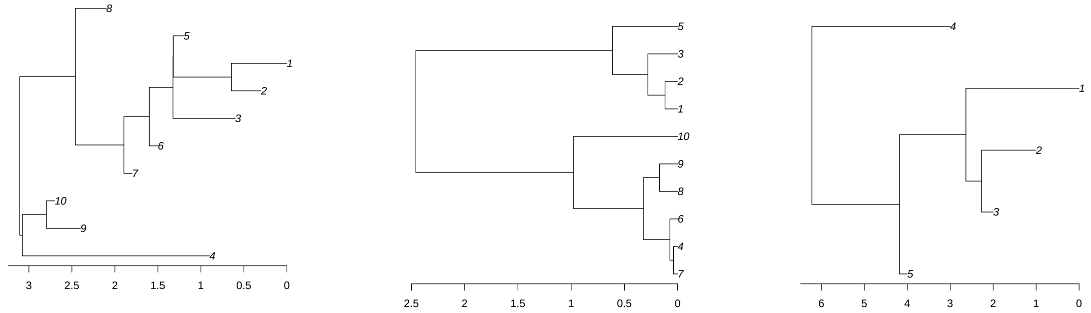
Required files :

```
> source("NonExpFunc1.R")
```

Command:

```
> treeNon<-MakeTreeNon(timesNon=timesNon)
```

Input: nonexp *timesNon*, a vector with the times at which the new leaves are to appear. (For an ultrametric tree they have to be all the same time...). In the end there will be as many leaves as the vector length.



(a) distance between leaves =0.3

(b) ultrametric tree

(c) non ultrametric tree

Figure 1: nonexpanding trees

Examples: Example1

```
> source("NonExpFunc1.R")
> times0<-rep(0,10)+0.3*c(1:10)
> tree<-MakeTreeNon(timesNon=times0)
> plot.phylo(tree);axisPhylo()
```

Example2: ultrametric tree

```
> source("NonExpFunc1.R")
> times0<-rep(0,10)
> tree<-MakeTreeNon(timesNon=times0)
> plot.phylo(tree);axisPhylo()
```

Example3:

```
> plot(MakeTreeNon(seq(0,4)));axisPhylo()
```

2.2 Expanding population size

Required files:

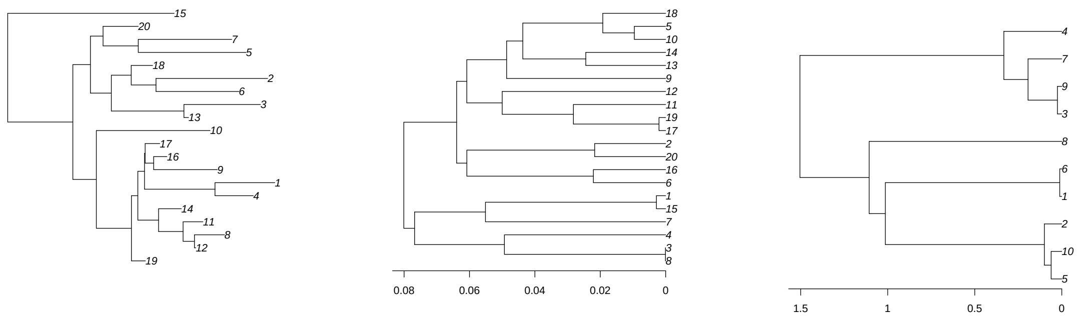
```
> source("ExpFunc3.R")
```

Command:

```
> tree<-MakeTreeExp(alphaExp=alphaExp,timesExp=timesExp)
```

Input: *timesExp* -a vector- of times at whnonexpich the new lineages are to appear. To get an ultrametric tree the vector should contain n times the same time. The number of leaves in the tree is the length of the

alphaExp -a number- the rate at which the population is expanding.



(a) distances between leaves 0.005, expansion rate =20

(b) ultrametric, expansion rate=50

(c) ultrametric, no expansion

Figure 2: expanding trees

Example Example 1:

```
> source("ExpFunc3.R")
> timesExp<-rep(0,20)+0.005*c(1:20)
> alphaExp<-20
> tree<-MakeTreeExp(alphaExp=alphaExp,timesExp=timesExp)
> plot.phylo(tree)
```

Example 2: ultrametric tree

```
> source("ExpFunc3.R")
> timesExp<-rep(0,20)
> alphaExp<-50
> tree<-MakeTreeExp(alphaExp=alphaExp,timesExp=timesExp)
> plot.phylo(tree);axisPhylo()

> plot(MakeTreeExp(0,rep(0,10)));axisPhylo()
```

2.3 Constant pupulation with expanding subpopulation

Required files:

```
> source("coalescentBothFuncNested4.R")
> source("ExpFunc2Both2.R")
```

Command:

```
> tree<-MakeTreeBoth(timesExp,timesNon,alpha,Ts,Texp)
```

Input: *timesExp* -a vector- the vector of times at which the new lineages of the expanding subpopulation are to appear. To get an ultrametric tree the vector should contain n times the same time. The number of leaves in this subtree is the length of the *timesEnonexpxp* vector.

timesNon, -a vector- the vector of times at which the new leaves of the non expanding subpopulation are to appear. (For an ultrametric tree they have to be all the same time...). In the end there will be as many leaves in the nonexpanding subpopulation as the length of the vector *alpha* -a number-: the rate at which the population is expanding.

Ts -a number- the time at which the two subpopulations separate. For any $t > Ts$ (the zero of time is at the more recent leave) the two subpopulations will be merged and treated as one (any lineage can coalesce with any other).

Texp -a number- the time at which one of the expanding subpopulations starts to expand. If we set this time to 0 we just have two identical subpopulations that expand at the same rate $Texp \leq Ts$.

Example Example 1: different leaf times, expansion rate =25

```
> timesExp<-c(seq(0.5,0.51,0.001))
> timesNon<-c(seq(0.1,0.2,0.01))
> alpha<-25
> Ts<-2
> Texp<-2
> tree<-MakeTreeBoth(timesExp,timesNon,alpha,Ts,Texp)
> plot.phylo(tree)
> h<-max(node.depth.edgelength(tree));axisPhylo()
> abline(v=(h-Ts), col=2, lwd=5)
> abline(v=(h-Texp), col=4, lwd=2)
```

nonexp

Example 2: ultrametric treenonexp

```
> timesExp<-rep(0,20)
> timesNon<-rep(0,20)
> alphaExp<-50
> Ts<-10
> Texp<-10
> tree<-MakeTreeBoth(timesExp,timesNon,alpha,Ts,Texp)
> plot.phylo(tree);axisPhylo()
> h<-max(node.depth.edgelength(tree))
> abline(v=(h-Ts), col=2, lwd=5)
> abline(v=(h-Texp), col=4, lwd=2)
>
```

Example 3: two subpopulations, none expanding

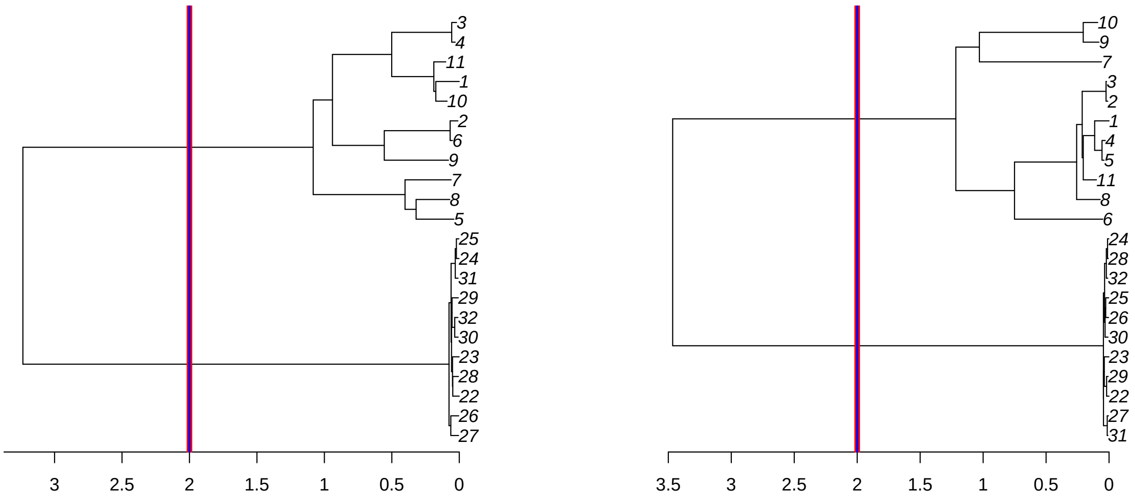
```
> Ts<-10
> Texp<-10
> plot(MakeTreeBoth(rep(0,20),rep(0,10),0.001,10,10));axisPhylo()
> h<-max(node.depth.edgelength(tree))
> abline(v=(h-Ts), col=2, lwd=5)
> abline(v=(h-Texp), col=4, lwd=2)
>
```

2.4 Likelihood

3 Log Likelihood

Required files:

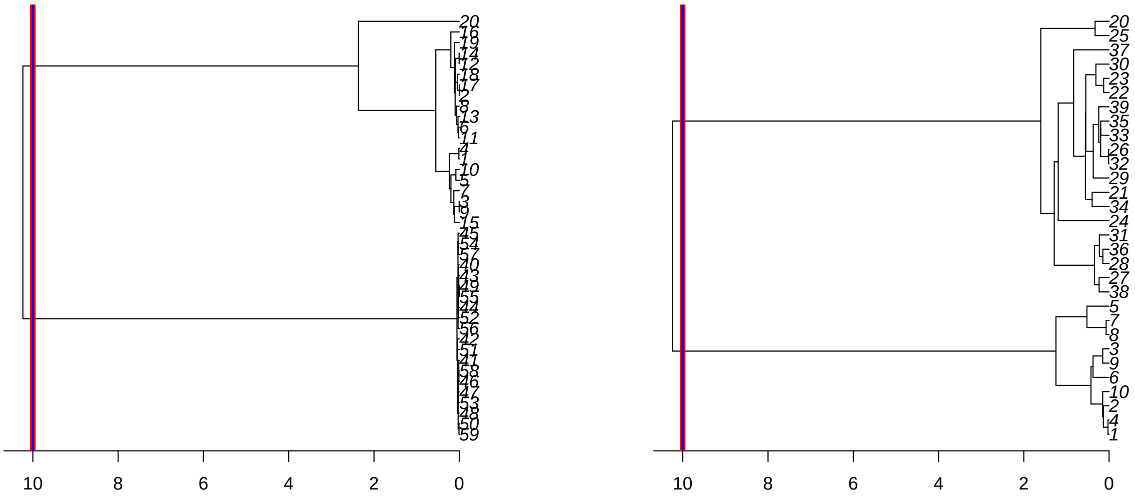
```
> source("like0.5Sum.R")
```



(a) expansion rate =25

(b) expansion rate=50

Figure 3: Example1. Same parameters, different expansion trees. The red and blue line represent the Ts and Texp parameters.



(a) expansion rate 50

(b) no expansion

Figure 4: ultrametric trees. The red and blue line do not appear because Ts and Texp are out of the plot range.

3.1 Log likelihood of a non expanding tree

Command:

```
> Lnon<-LogLikeliHood(NTree, theta)
```

Input: *Ntree* -a tree, phylo object- the tree we want to compute the likelihood of

theta -a number- the conversion factor between calendar units and coalescent units. If you are working in coalescent units, set it to 1.

Example:

```
> times0<-rep(0,10)+0.3*c(1:10)
> NTree<-MakeTreeNon(timesNon=times0)
> theta<-1
> Lnon<-LogLikeliHood(NTree, theta)
```

3.2 Log likelihood of an expanding tree

Command:

```
> Lexp<-LogLikeliHoodExp(ETree,theta,alpha)
```

Input: *Etree* -a tree, phylo object- the tree we want to compute the likelihood of, under the assumption at it is expanding at a given rate

theta -a number- the conversion factor between calendar units and coalescent units. If you are working in coalescent units, set it to 1.

alpha -a number- the given expansion rate

Example:

```
> timesExp<-rep(0,20)+0.005*c(1:20)
> alphaExp<-20
> ETree<-MakeTreeExp(alphaExp=alphaExp,timesExp=timesExp)
> theta<-1
> Lexp<-LogLikeliHoodExp(ETree,theta,alpha)
```

3.3 Log likelihood of a non expanding tree with an expanding tree

Required files:

```
> source("like0.5Sum.R")
> source("MCMCII.2.R")
```

Command:

```
> best<-LogLikeliHoodBoth(tree, theta, alpha,branch)
```

Input: *tree* -a tree, phylo object- the tree we want to compute the likelihood of, under the assumption that part of it is expanding at a given rate

theta -a number- the conversion factor between calendar units and coalescent units. If you are working in coalescent units, set it to 1.

alpha -a number- the given expansion rate

branch -a number- the branch where the expansion started

Example:

```
> source("like0.5Sum.R")
> source("MCMCII.2.R")
> timesExp<-rep(0,20)
> timesNon<-rep(0,20)
> alphaExp<-50
> Ts<-10
> Texp<-10
> tree<-MakeTreeBoth(timesExp,timesNon,alpha,Ts,Texp)
> write.tree(tree, file="tmp.tre")
> tree<-read.tree("tmp.tre")
> nodes<-AvailableNodes(tree,4)
> Ta<-matrix(nrow=length(nodes),ncol=3)
> colnames(Ta)<-c("branch","likelihood","expRate")
> for (i in 1:length(nodes)){
+         branch<-nodes[i]
+         alpha<-20
+         theta<-1
+         best<-LogLikeliHoodBoth(tree, theta, alpha,branch)
+         Ta[i,1]<-branch
+         Ta[i,2]<-best$likelihood
+         Ta[i,3]<-best$rate
+
+ }
> Ta
```

```
      branch    likelihood expRate
 [1,]     42 -4.721459e+01      20
 [2,]     43 -5.301594e+01      20
 [3,]     47 -5.728002e+01      20
 [4,]     48 -5.863279e+01      20
 [5,]     53 -5.669485e+01      20
 [6,]     56 -5.855356e+01      20
 [7,]     61 -7.051217e+10      20
 [8,]     62 -6.614904e+06      20
 [9,]     63 -2.308923e+02      20
[10,]     64 -5.088251e+01      20
[11,]     71 -8.118319e+02      20
[12,]     72 -5.503020e+01      20

>
```

3.4 MCMC

Required files:

```
> source("MCMCII.2.R")
```

Command:

```
> tab<-MCMC(tree,trials,burnon)
```

Input: *tree* -a tree, phylo object- the tree we want to compute the ...., under the assumption that part of it is expanding at a given rate

trials -a number- the total numbers of trials

burnon -a number- the number of burnon trials

Example:

```
> timesExp<-rep(0,20)
> timesNon<-rep(0,20)
> alphaExp<-20
> alpha<-20
> Ts<-10
> Texp<-10
> tree<-MakeTreeBoth(timesExp,timesNon,alpha,Ts,Texp)
> write.tree(tree, file="tmp.tre")
> tree<-read.tree("tmp.tre")
> trials<-100
> burnon<-3
> tab<-MCMC(tree,trials,burnon)
```