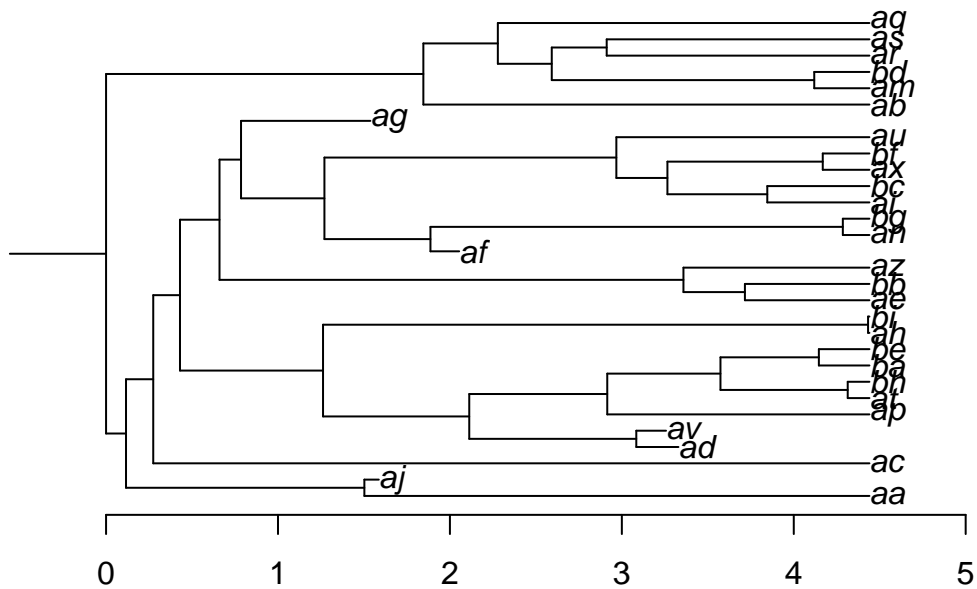


Tests

sim_phyl

```
st <- sim_phyl(ct=5)
plot(st$newick,root.edge = T)
axisPhylo(backward = F,root.time = st$t[1])
```



```
miss <- st$L[st$L[,3]!=(-1),]
miss_spe <- miss$spec
time <- rbind(data.frame(brtimes = st$L[,2], E=1, spec = st$L[,1]),data.frame(brtimes = miss[,3], E=0, spec = miss_spe))
time <- time[order(time$brtimes),]
time
```

##	brtimes	E	spec
## 1	0.0000000	1	aa
## 2	0.5597850	1	ab
## 3	0.6757521	1	ac
## 4	0.8340456	1	ad
## 5	0.9896804	1	ae
## 6	1.2195815	1	af
## 7	1.3446902	1	ag
## 8	1.8217114	1	ah
## 9	1.8298370	1	ai
## 10	2.0621950	1	aj
## 33	2.0969893	0	ag
## 34	2.1460336	0	aj
## 11	2.4052388	1	am
## 12	2.4465383	1	an
## 32	2.6136134	0	af
## 13	2.6722658	1	ap
## 14	2.8385040	1	aq
## 15	3.1521301	1	ar
## 16	3.4717304	1	as

```
## 17 3.4748887 1 at
## 18 3.5278725 1 au
## 19 3.6439110 1 av
## 35 3.8167714 0 av
## 20 3.8248819 1 ax
## 31 3.8889449 0 ad
## 21 3.9180702 1 az
## 22 4.1334003 1 ba
## 23 4.2756353 1 bb
## 24 4.4058079 1 bc
## 25 4.6789673 1 bd
## 26 4.7057839 1 be
## 27 4.7284592 1 bf
## 28 4.8452701 1 bg
## 29 4.8736083 1 bh
## 30 4.9919446 1 bi
```

```
time$n <- cumsum(time$E)-cumsum(1-time$E)
time$n <- c(0,time$n[1:length(time$n)-1])
time
```

```
##      brtimes E spec n
## 1 0.0000000 1 aa 0
## 2 0.5597850 1 ab 1
## 3 0.6757521 1 ac 2
## 4 0.8340456 1 ad 3
## 5 0.9896804 1 ae 4
## 6 1.2195815 1 af 5
## 7 1.3446902 1 ag 6
## 8 1.8217114 1 ah 7
## 9 1.8298370 1 ai 8
## 10 2.0621950 1 aj 9
## 33 2.0969893 0 ag 10
## 34 2.1460336 0 aj 9
## 11 2.4052388 1 am 8
## 12 2.4465383 1 an 9
## 32 2.6136134 0 af 10
## 13 2.6722658 1 ap 9
## 14 2.8385040 1 aq 10
## 15 3.1521301 1 ar 11
## 16 3.4717304 1 as 12
## 17 3.4748887 1 at 13
## 18 3.5278725 1 au 14
## 19 3.6439110 1 av 15
## 35 3.8167714 0 av 16
## 20 3.8248819 1 ax 15
## 31 3.8889449 0 ad 16
## 21 3.9180702 1 az 15
## 22 4.1334003 1 ba 16
## 23 4.2756353 1 bb 17
## 24 4.4058079 1 bc 18
## 25 4.6789673 1 bd 19
## 26 4.7057839 1 be 20
## 27 4.7284592 1 bf 21
## 28 4.8452701 1 bg 22
```

```
## 29 4.8736083 1    bh 23
## 30 4.9919446 1    bi 24

st$br

## [1] 0.5597850 0.6757521 0.8340456 0.9896804 1.2195815 1.3446902 1.8217114
## [8] 1.8298370 2.0621950 2.0969893 2.1460336 2.4052388 2.4465383 2.6136134
## [15] 2.6722658 2.8385040 3.1521301 3.4717304 3.4748887 3.5278725 3.6439110
## [22] 3.8167714 3.8248819 3.8889449 3.9180702 4.1334003 4.2756353 4.4058079
## [29] 4.6789673 4.7057839 4.7284592 4.8452701 4.8736083 4.9919446 5.0000000

st$E

## [1] 1 1 1 1 1 1 1 1 1 0 0 1 1 0 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1

st$E == time$E[2:length(time$E)]

## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [15] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [29] TRUE TRUE TRUE TRUE TRUE TRUE TRUE

st$br[1:(length(st$br)-1)] == time$brtimes[2:length(time$brtimes)]

## [1] TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE
## [12] FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE
## [23] FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [34] FALSE

missing = time[which(is.element(time$spec,miss_spe)),]
missing

##      brtimes E spec  n
## 4  0.8340456 1   ad  3
## 6  1.2195815 1   af  5
## 7  1.3446902 1   ag  6
## 10 2.0621950 1   aj  9
## 33 2.0969893 0   ag 10
## 34 2.1460336 0   aj  9
## 32 2.6136134 0   af 10
## 19 3.6439110 1   av 15
## 35 3.8167714 0   av 16
## 31 3.8889449 0   ad 16
```

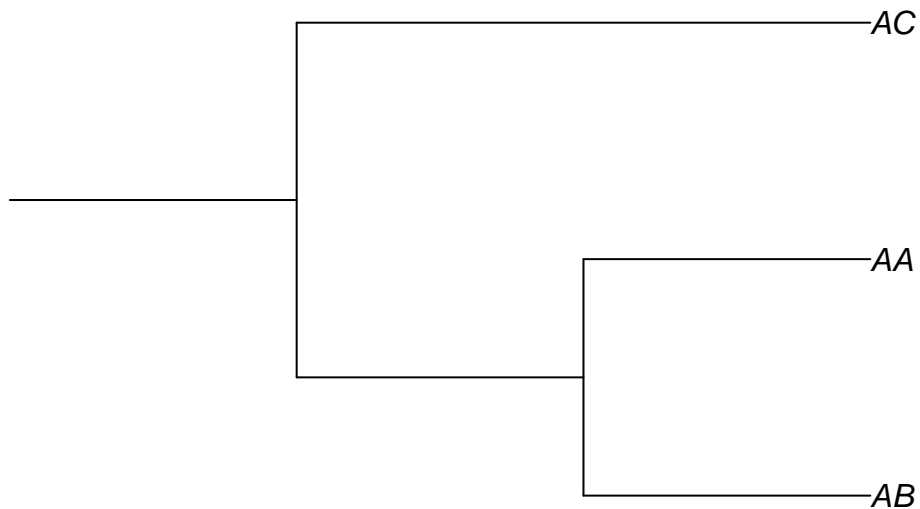
Phylo2p and update_tree

First, to warm up, we plot a simple tree

```
obs = '((AB:1,AA:1):1,AC:2):1;'
#library(ape)
obs = read.tree(text=obs)
is.rooted(obs)

## [1] TRUE

plot(obs,show.tip.label = T, root.edge = TRUE)
```



```
branching.times(obs)
```

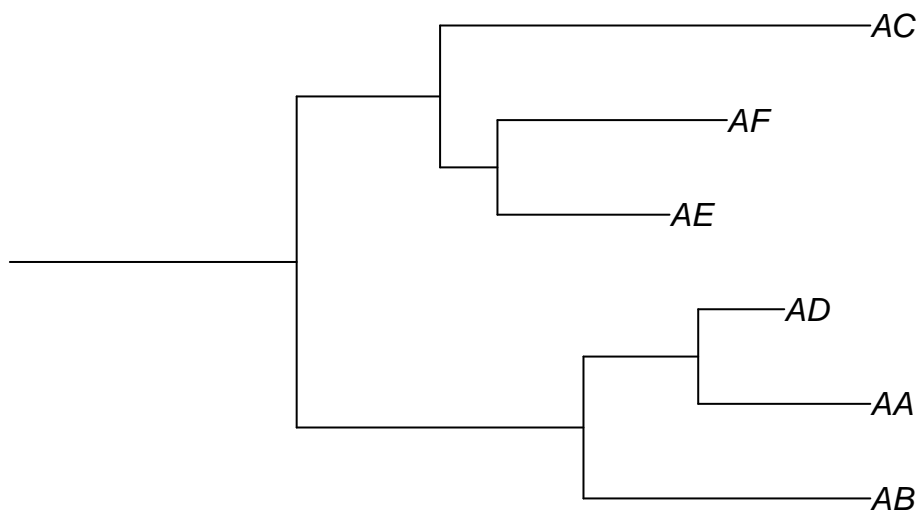
```
## 4 5
## 2 1
```

Then a longer one

```
# complete/incomplete trees
comp = '((AB:1,(AA:0.6,AD:0.3):0.4):1,((AE:0.6,AF:0.8):0.2,AC:1.5):0.5):1;'
comp= read.tree(text = comp)
is.rooted(comp)
```

```
## [1] TRUE
```

```
plot(comp,show.tip.label = T, root.edge = TRUE)
```



```
branching.times(comp)
```

```
## 7 8 9 10 11
## 2.0 1.0 0.6 1.5 1.3
```

Then we drop extinct species

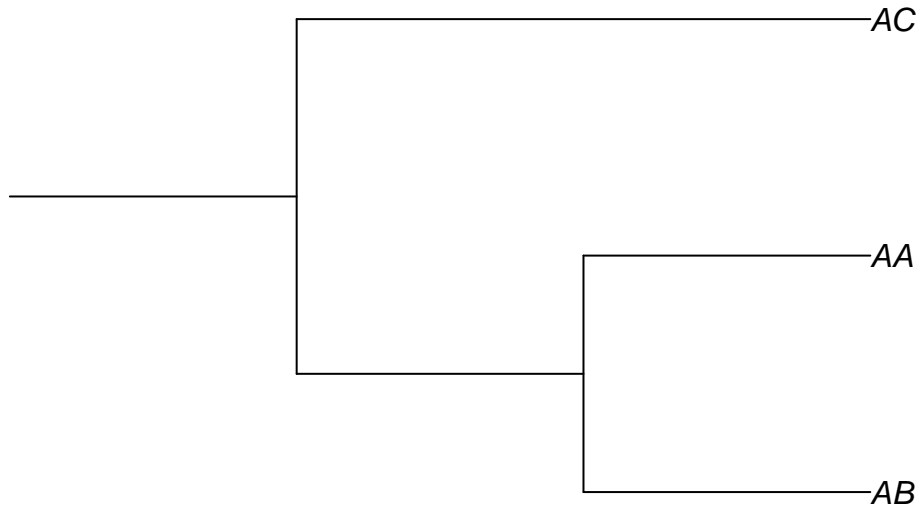
```
incomp = drop.fossil(comp)
write.tree(incomp)
```

```
## [1] "((AB:1,AA:1):1,AC:2):1;"
```

```
is.rooted(incomp)
```

```
## [1] TRUE
```

```
plot(incomp,show.tip.label = T, root.edge = TRUE)
```



```
branching.times(incomp)
```

```
## 4 5
```

```
## 2 1
```

Now we would like to be able to add an extinct species

```
in_dmea = phylo2p(incomp)
```

```
in_dmea$wt
```

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

```
in_dmea$E
```

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

```
in_dmea$n
```

```
## [1] 1 2 3
```

```
update_tree(in_dmea,t_spe=0.5,t_ext = 1.5)
```

```
## $wt
```

```
## [1] 0.5 0.5 0.5 0.5 1.0
```

```
##
```

```
## $E
```

```
## [1] 1 1 0 1
```

```
##
```

```
## $n
```

```
## [1] 1 2 3 2 3
```

Create_L

```
st <- sim_phyl()
L <- create_L(t=st$wt, E=st$E)
miss <- L[L[,3] != (-1),]
miss_spe <- miss$spec
time <- rbind(data.frame(brtimes = L[,2], E=1, spec = L[,1]), data.frame(brtimes = miss[,3], E=0, spec =
time <- time[order(time$brtimes),]
time$n <- cumsum(time$E)-cumsum(1-time$E)
all.equal(st$n,time$n)
```

```
## [1] TRUE
```

```
all.equal(st$n,cumsum(time$E)-cumsum(1-time$E))
```

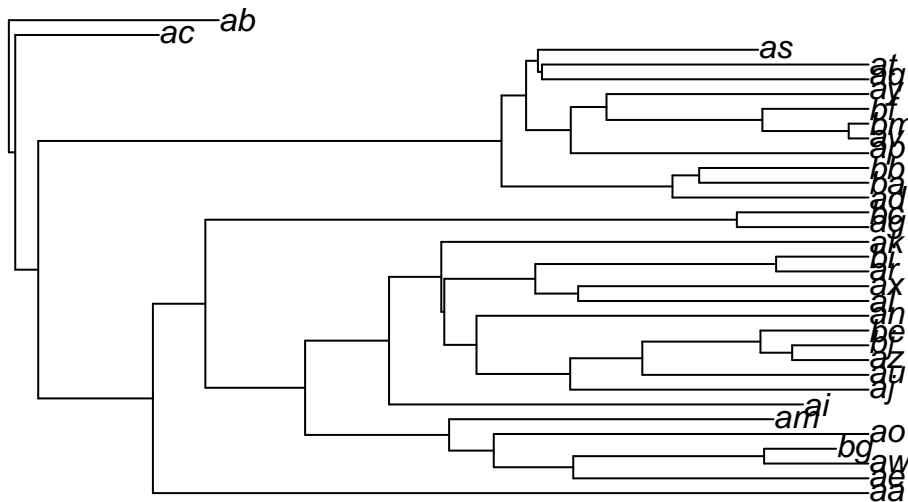
```
## [1] TRUE
```

```
all.equal(c(0,st$br),c(time$brtimes,15))
```

```
## [1] TRUE
```

p2phylo

```
s = sim_phyl(ct=6)
plot(s$newick)
```



```
p = phylo2p(s$newick)
p$E == s$E
```

```
## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

```
## [15] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

```
## [29] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

```
p$wt == s$t
```

```
## logical(0)
```

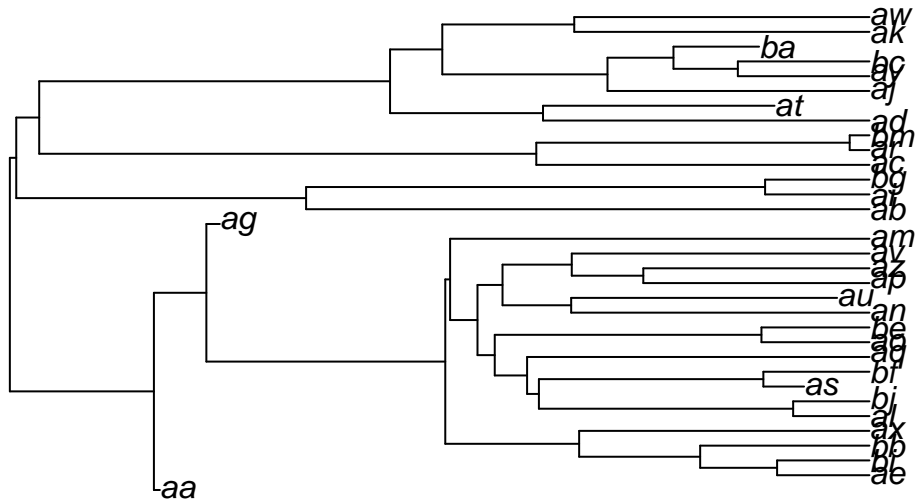
```
#p$wt
```

```
#s$t
```

```
p$n == s$n
```

```
## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [15] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [29] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

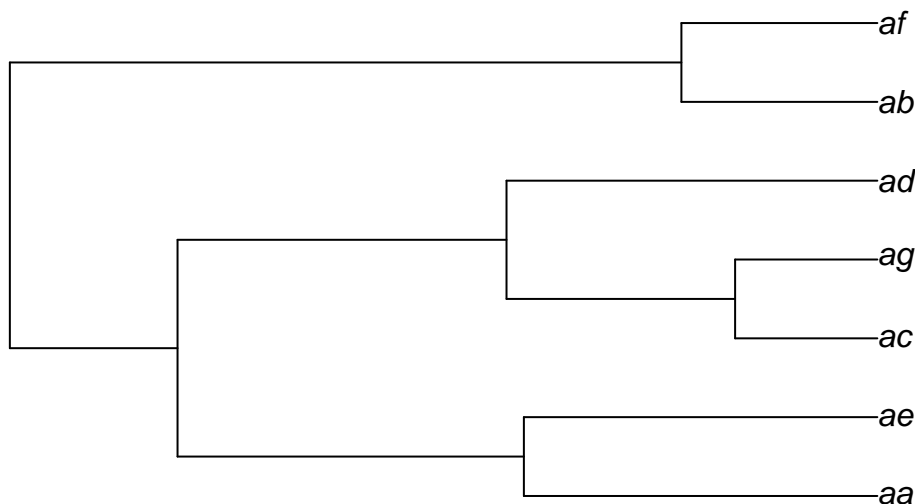
```
s2 = p2phylo(p)
plot(s2)
```



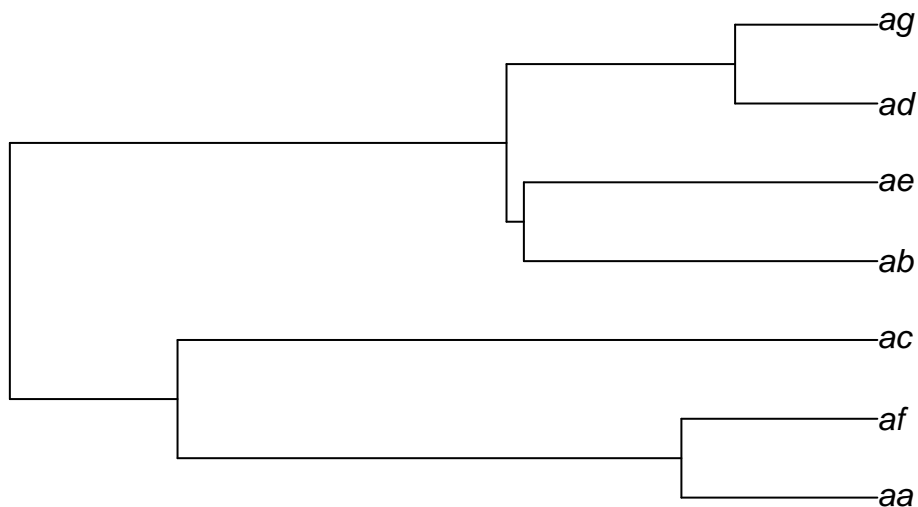
voilà!, we found two things:

1. This algorithm does not label species in a general way yet, but only to models where all species has same probability to speciate.
2. We found a nice way to illustrate 'equivalent' trees.

```
s = sim_phyl(ct=2)
plot(s$newick)
```



```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```



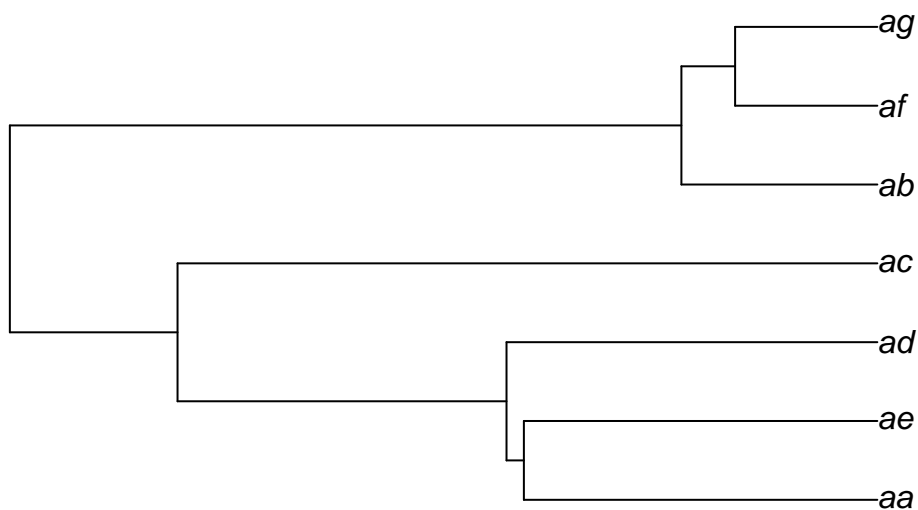
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.8910153 0.2494834 0.4721464 0.4500574 0.1810921
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 3.172066e-17
```

```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```



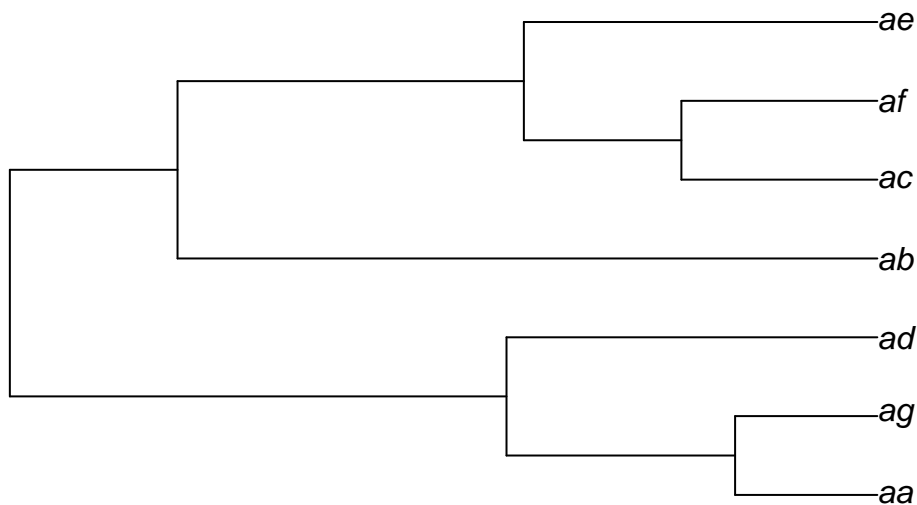
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.8910153 0.4721464 0.4500574 0.2494834 0.1810921
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 0
```

```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```

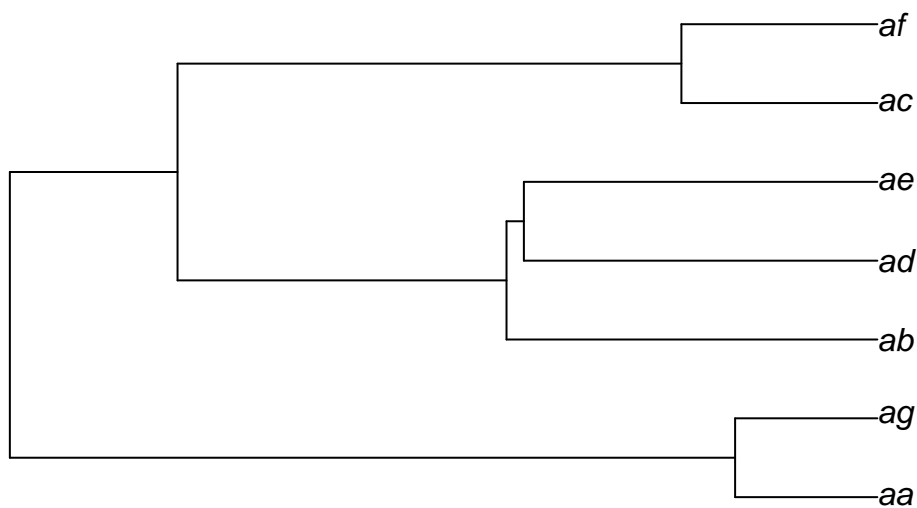
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.4721464 0.1810921 0.8910153 0.4500574 0.2494834
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 4.758099e-17
```

```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```



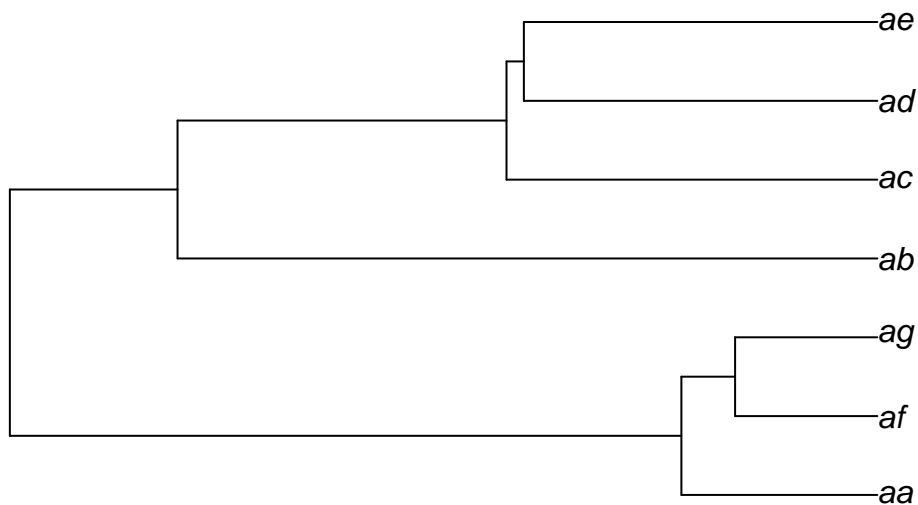
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.1810921 0.8910153 0.4721464 0.4500574 0.2494834
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 1.586033e-17
```

```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```



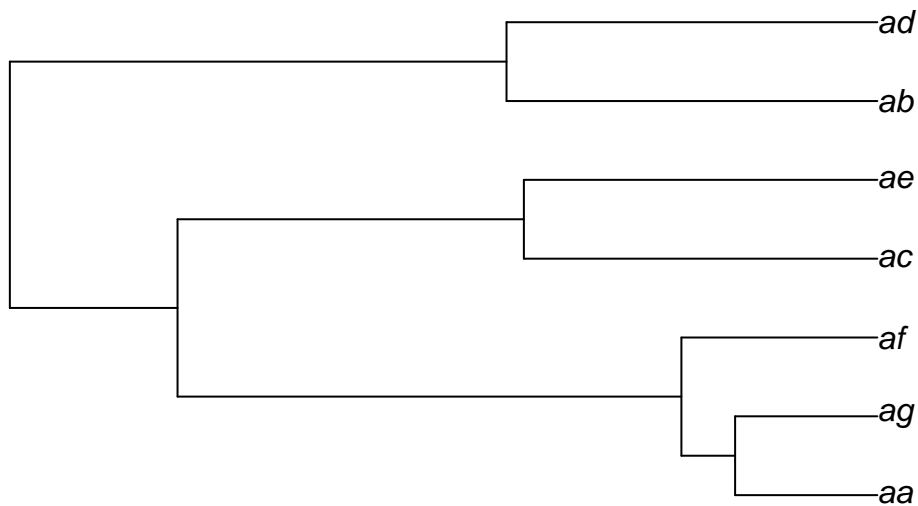
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.2494834 0.1810921 0.8910153 0.4721464 0.4500574
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 0
```

```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```



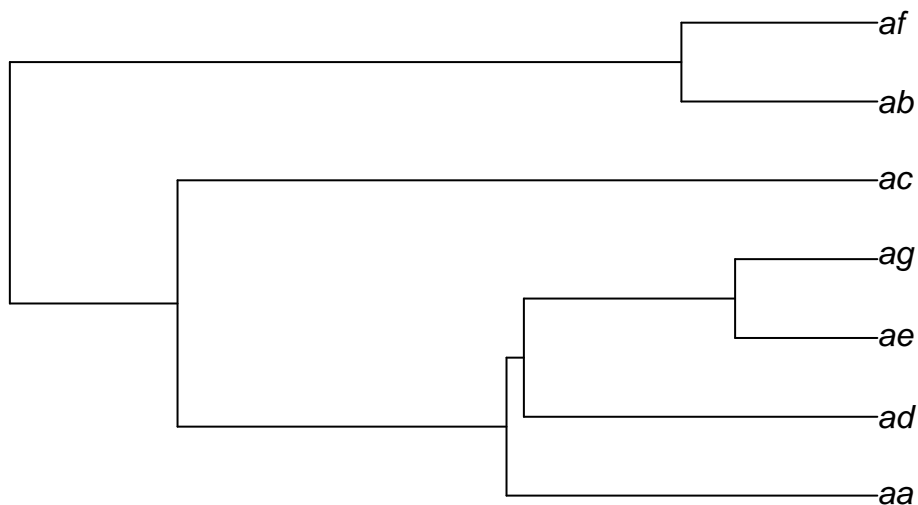
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.8910153 0.2494834 0.1810921 0.4500574 0.4721464
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 1.586033e-17
```

```
s2=p2phylo(phylo2p(s$newick))
plot(s2)
```



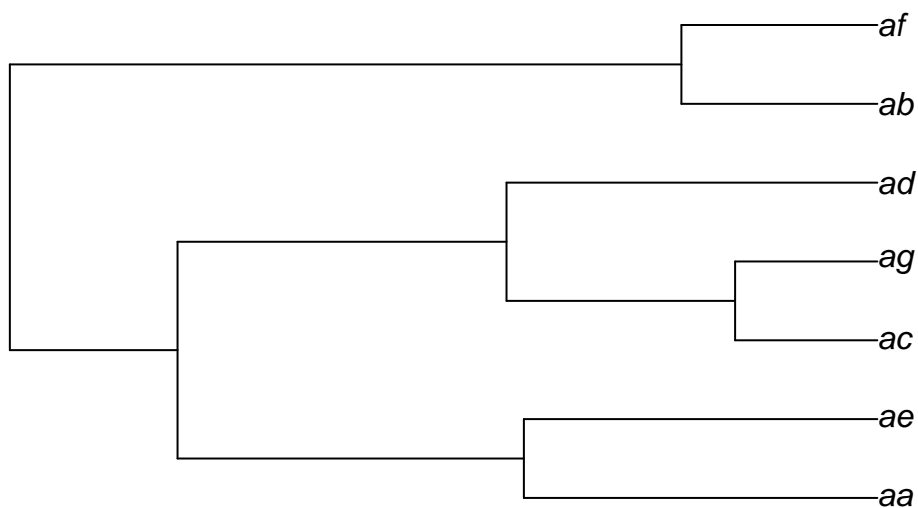
```
branching.times(s2)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.8910153 0.4721464 0.4500574 0.1810921 0.2494834
```

```
nLTTstat(s2,s$newick)
```

```
## [1] 3.172066e-17
```

```
plot(s$newick)
```



```
branching.times(s$newick)
```

```
##           8           9           10           11           12           13
## 1.1045501 0.8910153 0.4500574 0.4721464 0.1810921 0.2494834
```

- Why the branching times are different ??
- Why on those cases the nLTT > 0 ??