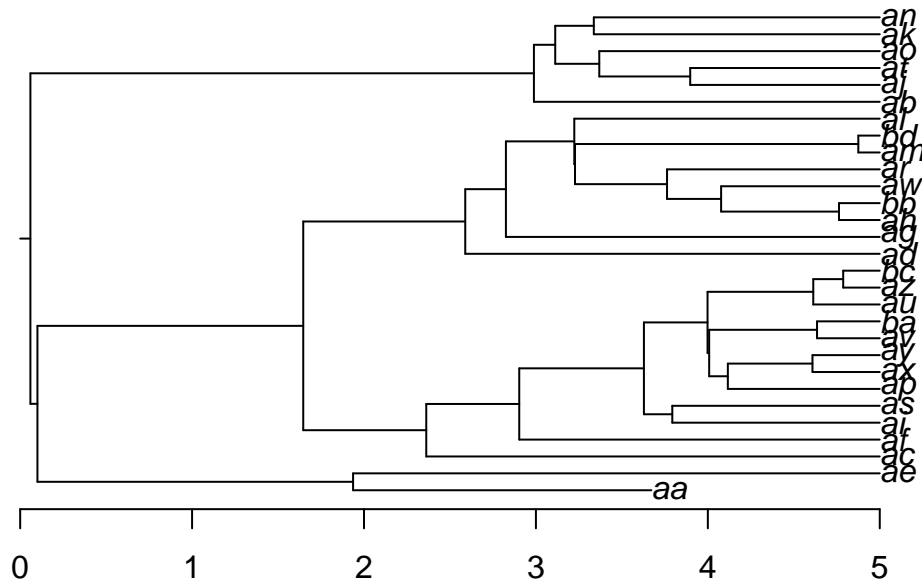


# Tests

## 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	0.00000000	1	aa
## 2	0.05941906	1	ab
## 3	0.10036685	1	ac
## 4	1.64658012	1	ad
## 5	1.93651764	1	ae
## 6	2.36215640	1	af
## 7	2.58888433	1	ag
## 8	2.82541383	1	ah
## 9	2.90302593	1	ai
## 10	2.98839385	1	aj
## 11	3.11224072	1	ak
## 12	3.22320552	1	al
## 13	3.22736388	1	am
## 14	3.33672544	1	an
## 15	3.36859206	1	ao
## 16	3.62815398	1	ap
## 30	3.67145112	0	aa
## 17	3.76273458	1	ar
## 18	3.79324519	1	as

```
## 19 3.89753396 1 at
## 20 3.99834609 1 au
## 21 4.00779412 1 av
## 22 4.07725076 1 aw
## 23 4.11635308 1 ax
## 24 4.60872189 1 ay
## 25 4.61342138 1 az
## 26 4.63519402 1 ba
## 27 4.76285972 1 bb
## 28 4.78733699 1 bc
## 29 4.87538453 1 bd
```

```
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  0.00000000 1  aa  0
## 2  0.05941906 1  ab  1
## 3  0.10036685 1  ac  2
## 4  1.64658012 1  ad  3
## 5  1.93651764 1  ae  4
## 6  2.36215640 1  af  5
## 7  2.58888433 1  ag  6
## 8  2.82541383 1  ah  7
## 9  2.90302593 1  ai  8
## 10 2.98839385 1  aj  9
## 11 3.11224072 1  ak 10
## 12 3.22320552 1  al 11
## 13 3.22736388 1  am 12
## 14 3.33672544 1  an 13
## 15 3.36859206 1  ao 14
## 16 3.62815398 1  ap 15
## 30 3.67145112 0  aa 16
## 17 3.76273458 1  ar 15
## 18 3.79324519 1  as 16
## 19 3.89753396 1  at 17
## 20 3.99834609 1  au 18
## 21 4.00779412 1  av 19
## 22 4.07725076 1  aw 20
## 23 4.11635308 1  ax 21
## 24 4.60872189 1  ay 22
## 25 4.61342138 1  az 23
## 26 4.63519402 1  ba 24
## 27 4.76285972 1  bb 25
## 28 4.78733699 1  bc 26
## 29 4.87538453 1  bd 27
```

```
st$br
```

```
## [1] 0.05941906 0.10036685 1.64658012 1.93651764 2.36215640 2.58888433
## [7] 2.82541383 2.90302593 2.98839385 3.11224072 3.22320552 3.22736388
## [13] 3.33672544 3.36859206 3.62815398 3.67145112 3.76273458 3.79324519
## [19] 3.89753396 3.99834609 4.00779412 4.07725076 4.11635308 4.60872189
## [25] 4.61342138 4.63519402 4.76285972 4.78733699 4.87538453 5.00000000
```

```

st$E

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

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

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

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

##      brtimes E spec  n
## 1  0.000000 1   aa  0
## 30 3.671451 0   aa 16

```

## Phylo2p and update\_\_tree

First, to warm up, we plot a simple tree

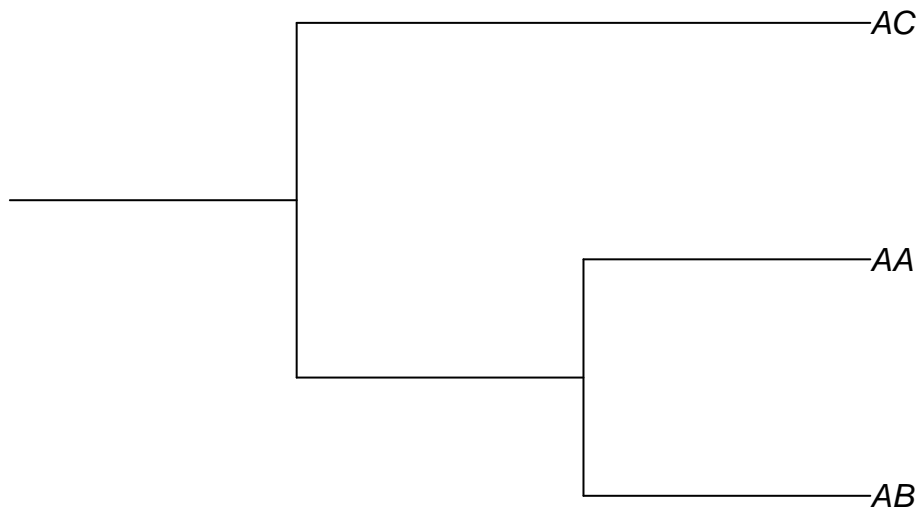
```

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)
```

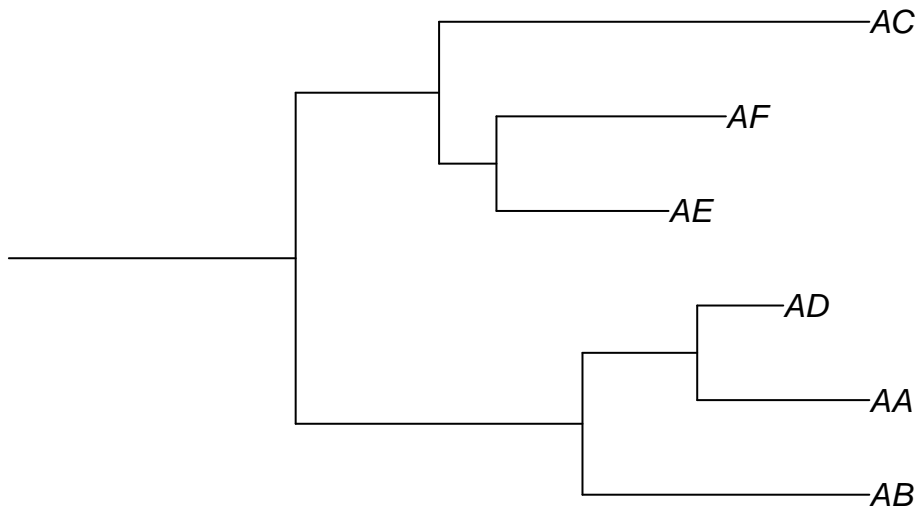
```
## 4 5
## 2 1
```

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

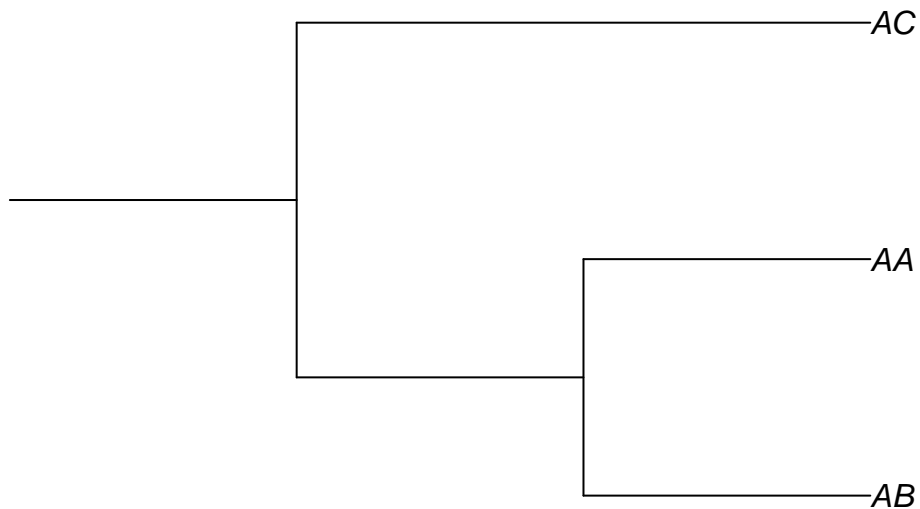
```
incomp = drop.fossil(comp)
write.tree(incomp)
```

```
## [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$t
```

```
## NULL
```

```
in_dmea$E
```

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

```
in_dmea$n
```

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

```
update_tree(wt=in_dmea$t,t_spe=0.5,t_ext = 1.5 ,E=in_dmea$E,n=in_dmea$n)
```

```
## $wt
```

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

```
##
```

```
## $E
```

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

```
##
```

```
## $n
```

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

Now , after a little modification on the drop.fossil function, it works properly

```
in_dmea = phylo2p(obs)
in_dmea$t
```

```
## NULL
```

```
in_dmea$E
```

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

```
in_dmea$n
```

```
## [1] 1 2 3
update_tree(wt=in_dmea$t,t_spe=0.5,t_ext = 1.5 ,E=in_dmea$E,n=in_dmea$n)

## $wt
## [1] 0.5 1.0 NA NA NA 0.5
##
## $E
## [1] 1 0 1 1
##
## $n
## [1] 1 2 1 2 3
```

## Create\_L

```
st <- sim_phyl()
L <- create_L(t=st$t, E=st$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))

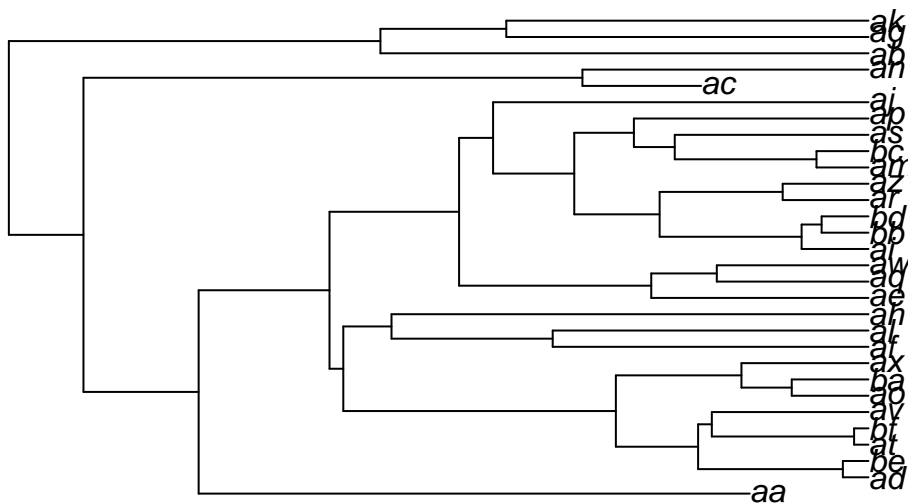
## [1] TRUE

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
```

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

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

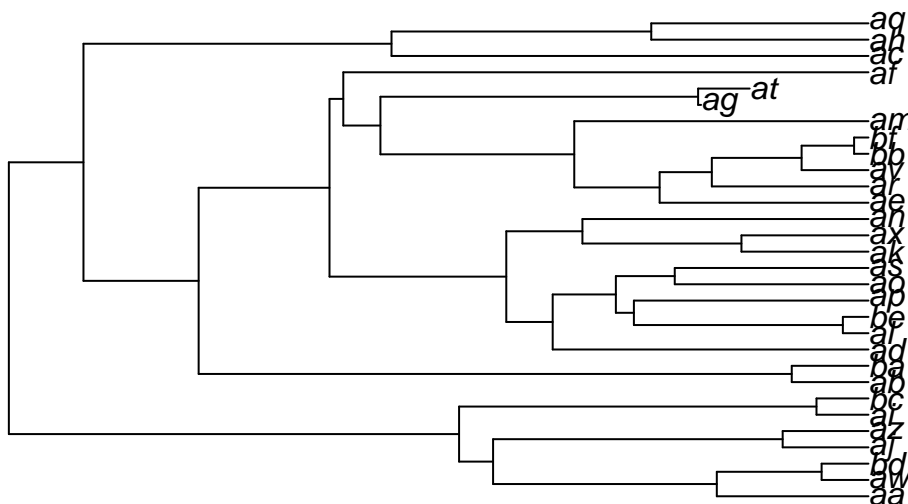
```
#p$wt
```

```
#s$t
```

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

```
## [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
```

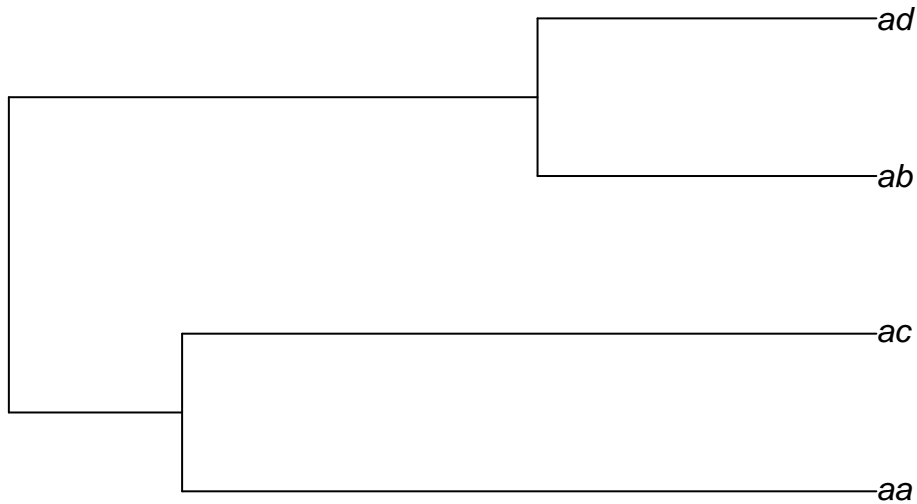
```
s2 = p2phylo(p)
plot(s2)
```



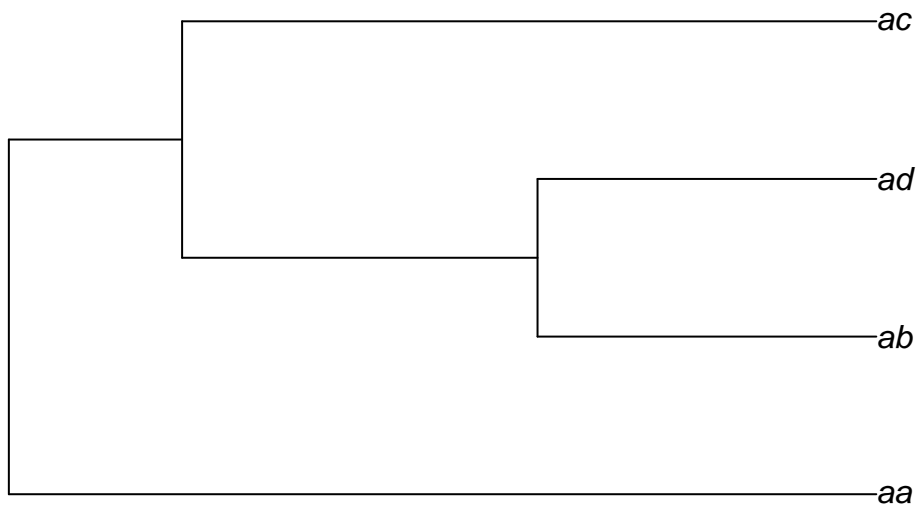
voilà!, we found two things:

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)
```

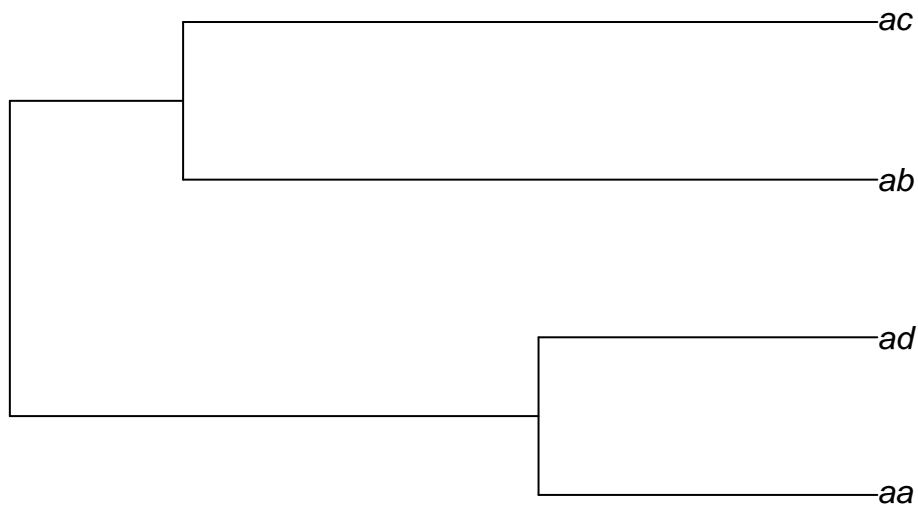
```
##          5          6          7
## 1.377397 1.102269 0.537950
```

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

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

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





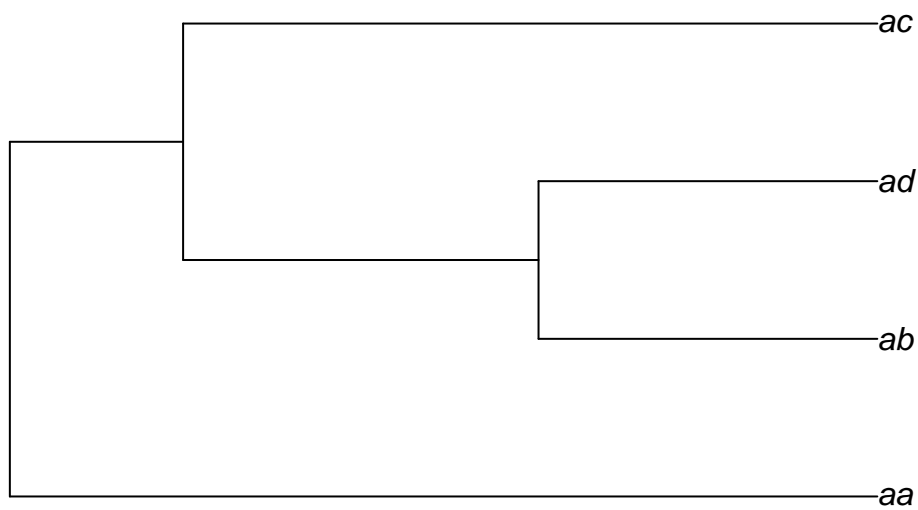
```
branching.times(s2)
```

```
##          5          6          7
## 1.377397 0.537950 1.102269
```

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

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

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



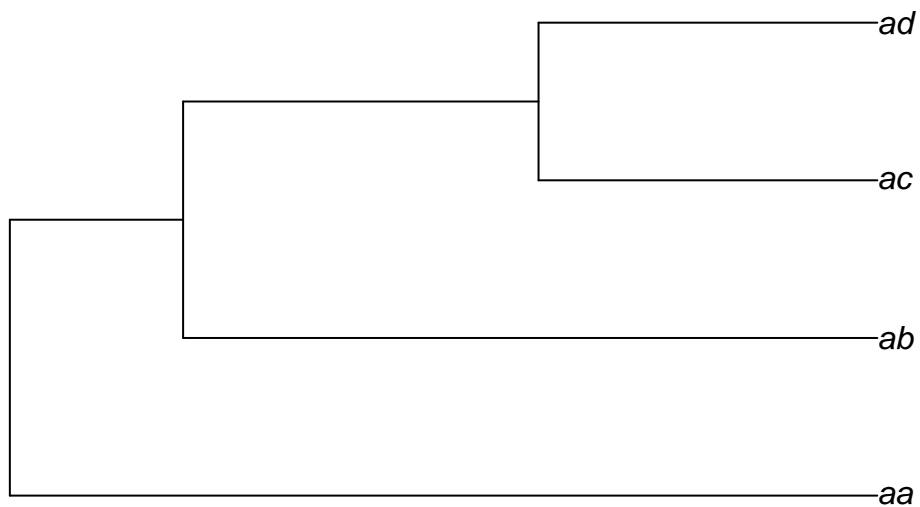
```
branching.times(s2)
```

```
##          5          6          7
## 1.377397 1.102269 0.537950
```

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

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

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



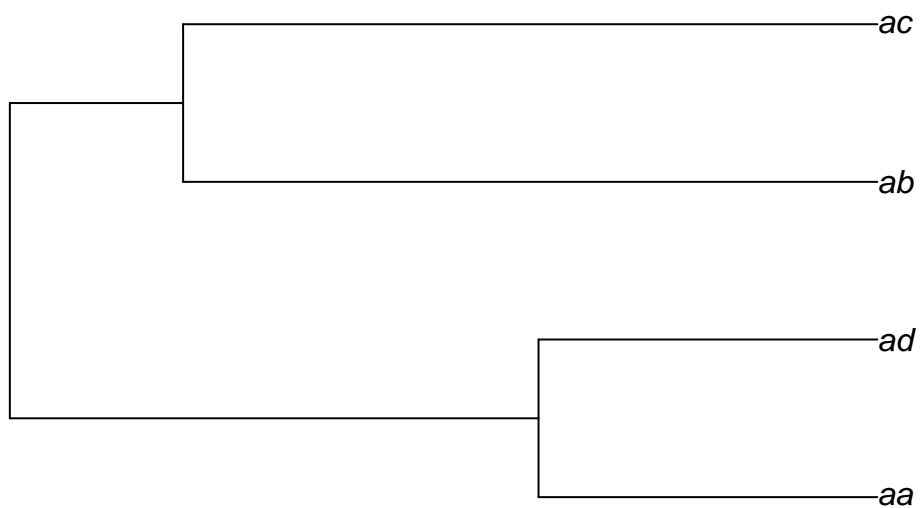
```
branching.times(s2)
```

```
##          5          6          7
## 1.377397 1.102269 0.537950
```

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

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

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



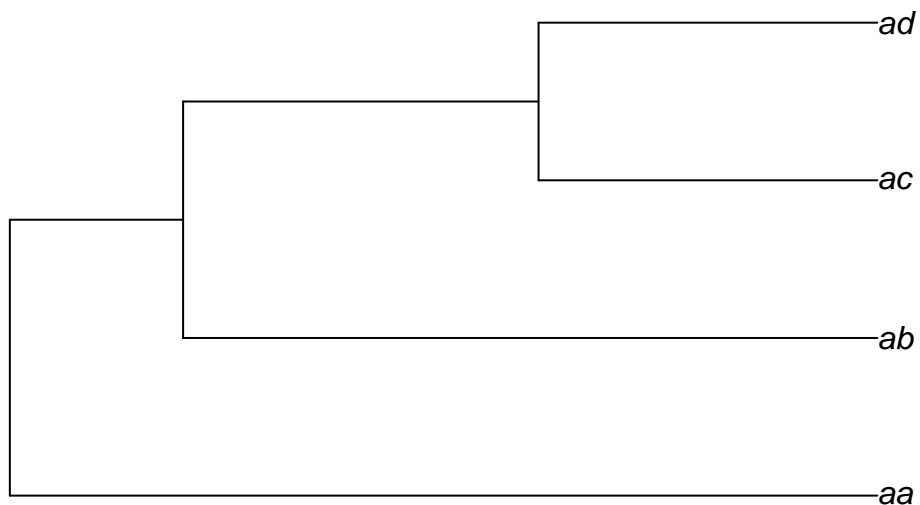
```
branching.times(s2)
```

```
##          5          6          7
## 1.377397 0.537950 1.102269
```

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

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

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



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

```
##          5          6          7
## 1.377397 1.102269 0.537950
```

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

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

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

```
plot(s2)
```

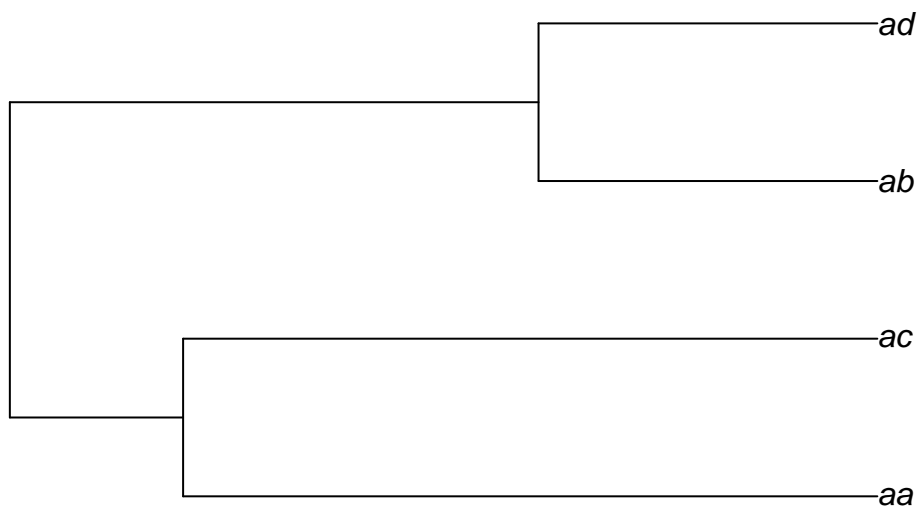
```
branching.times(s2)
```

```
##          5          6          7
## 1.377397 1.102269 0.537950
```

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

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

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



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

```
##          5          6          7
```

## 1.377397 1.102269 0.537950

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