mutSOMA

1) Loading data and determine nucleotide frequency

```
## Reading in the poplar fasta file and determining the nucleotide frequency
poplar <- readDNAStringSet(paste(input.data.dir, "PtrichocarpaStettler14_532_v1.0.fa", sep=""))
freqBASE <- letterFrequency(poplar, letters = c("A", "C", "T", "G"), as.prob=FALSE)
freqBASE <- colSums(freqBASE[1:19,])
probBASE <- freqBASE/sum(freqBASE)

freqBASE
```

```
##         A         C         T         G
## 129631802  65849970 129443927  65791587
```

```
probBASE
```

```
##         A         C         T         G
## 0.3317790 0.1685361 0.3312982 0.1683867
```

2) Constructing pedigree

```
pedigree <- makeVCFpedigreeTEMP(genome.size=sum(freqBASE), input.dir = input.data.dir)
```

```
## Scanning file to determine attributes.
## File attributes:
##   meta lines: 34
##   header_line: 35
##   variant count: 2889
##   column count: 13
##
Meta line 34 read in.
## All meta lines processed.
## gt matrix initialized.
## Character matrix gt created.
##   Character matrix gt rows: 2889
##   Character matrix gt cols: 13
##   skip: 0
##   nrows: 2889
##   row_num: 0
##
Processed variant 1000
Processed variant 2000
Processed variant: 2889
## All variants processed
## Scanning file to determine attributes.
## File attributes:
##   meta lines: 34
##   header_line: 35
##   variant count: 308
##   column count: 13
##
Meta line 34 read in.
## All meta lines processed.
## gt matrix initialized.
## Character matrix gt created.
##   Character matrix gt rows: 308
##   Character matrix gt cols: 13
```

```
##    skip: 0
##    nrows: 308
##    row_num: 0
##
Processed variant: 308
## All variants processed
```

```
pedigree<-pedigree[[1]]
pedigree<-pedigree[-21,]
head(pedigree)
```

```
##   time0 time1 time2      D.value
## 1     0   297   287 2.385356e-06
## 2     0   297   324 3.396318e-06
## 3     0   327   287 2.442943e-06
## 4     0   297   287 2.904914e-06
## 5     0   328   287 2.504368e-06
## 6     0   328   287 2.385356e-06
```

3) Run models

```
out <- mutSOMA(pedigree.data = pedigree,
            p0aa= probBASE[1],
            p0cc= probBASE[2],
            p0tt= probBASE[3],
            p0gg= probBASE[4],
            Nstarts=20,
            out.dir = out.data.dir,
            out.name = "test_mutSOMA")
```

```
## Progress:  0.05
## Progress:  0.1
## Progress:  0.15
## Progress:  0.2
## Progress:  0.25
## Progress:  0.3
## Progress:  0.35
## Progress:  0.4
## Progress:  0.45
## Progress:  0.5
## Progress:  0.55
## Progress:  0.6
## Progress:  0.65
## Progress:  0.7
## Progress:  0.75
## Progress:  0.8
## Progress:  0.85
## Progress:  0.9
## Progress:  0.95
## Progress:  1
```

```
summary(out)
```

```
##                    Length Class      Mode
## estimates            12   data.frame list
## estimates.flagged    12   data.frame list
## pedigree              7   data.frame list
## settings              2   data.frame list
## model                 1   -none-     character
## for.fit.plot       3275   -none-     numeric
```

```r
head(out$estimates)
```

```
##                      gamma      intercept        value fevals gevals niter
## Nelder-Mead   4.788956e-09 7.761467e-10 6.186237e-12    117     NA    NA
## Nelder-Mead7  4.793970e-09 1.564170e-09 6.186263e-12     67     NA    NA
## Nelder-Mead12 4.784769e-09 3.475965e-09 6.186263e-12     73     NA    NA
## Nelder-Mead14 4.785504e-09 2.680097e-09 6.186263e-12     85     NA    NA
## Nelder-Mead18 4.798799e-09 5.850340e-11 6.186390e-12     83     NA    NA
## Nelder-Mead10 4.649408e-09 5.575924e-08 6.218592e-12     71     NA    NA
##               convcode kkt1  kkt2  xtime      g.start intercept.start
## Nelder-Mead          0 TRUE FALSE 23.609 1.031526e-08    3.507600e-06
## Nelder-Mead7         0 TRUE FALSE 13.435 3.606427e-09    3.155548e-06
## Nelder-Mead12        0 TRUE FALSE 14.669 1.326495e-11    3.486424e-06
## Nelder-Mead14        0 TRUE FALSE 18.090 4.538783e-10    2.363383e-07
## Nelder-Mead18        0 TRUE FALSE 16.706 9.748578e-08    2.735923e-06
## Nelder-Mead10        0 TRUE FALSE 14.450 6.503170e-07    2.833372e-06
```

```r
head(out$pedigree)
```

```
##   time0 time1 time2      div.obs delta.t     div.pred      residual
## 1     0   297   287 2.385356e-06     584 2.797521e-06 -4.121645e-07
## 2     0   297   324 3.396318e-06     621 2.974712e-06  4.216060e-07
## 3     0   327   287 2.442943e-06     614 2.941189e-06 -4.982462e-07
## 4     0   297   287 2.904914e-06     584 2.797521e-06  1.073927e-07
## 5     0   328   287 2.504368e-06     615 2.945978e-06 -4.416097e-07
## 6     0   328   287 2.385356e-06     615 2.945978e-06 -5.606216e-07
```