# Introduction

The package `lyne` implements likelihood calculation and parameter optimization for continous time Markov models for 5-mC DNA methylation. Methylation states are coded as integers $\{1, 2, 3\}$ for un-methylated, partially methlated and methylated CpGs, respectively. The following is a minimal example:

First we specify a lineage tree from which methylation data has been assayed and load some example data:

```
> require(lyne)
>
> #- Lineage tree from Bock et al. (2012)
> LT  = exampleLineageTree()
> plot(LT$graph,layout=LT$layout)
```
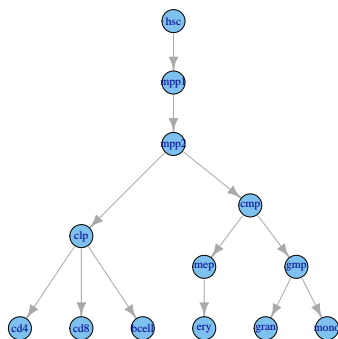


```
>
> #- Methylation data from Bock et al. (2012)
> load(system.file(      "extdata",
+                        "bock12-blood_small-ali.rdat",
+                        package="lyne"))
> print(ali[1:3,1:10])

       [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
bcell     3    2    3    3    3    3   NA    3    3     3
```

```
cd4      3   3   3   3   3   3   NA   3   3   3
cd8      3   3   3   3   3   3   NA   3   3   3

>
> #- Specify a tree traversal for likelihood calclation
> edgeMat = getEdgeMatrix(LT$graph)$traversal
>
> #- Make sure tree nodes and alignment match up
> nodeNms = V(LT$graph)$name
> ali     = ali[nodeNms,]
```

Next, we calculate the likelihood of the observed data, given a $\Gamma + I$ model specified by a parameter vector:

```
> #- Rate matrix: 3 equilibrium frequences, 3 rate parameters
> pars = c(.7,.1,.2,.45,.1,.45)
> Q    = pars2Q.3states.GTR(pars)
>
> #- Random branch lengths for the lineage tree
> bl = rgamma(nrow(edgeMat),1,1)
>
> #- Transition probabilites for each edge
> require(plyr)
> transProb = aperm(maply(bl,function(x) Q2P(Q,x)),perm=c(2,3,1))
>
> #- Likelihood of observation given the parameters
> pars = c(pars,bl)
> like = mcTreeLike.3states.GTR.bhom(pars,ali,mult=1,edgeMat)
```

And finally we obtain maximum likelihood estimates for the parameters of a $\Gamma + I$ model (for three rate categories). In addition to the parameters above this includes mixture coefficient and the α parameter for the Gamma distribution:

```
> #- Initial parameters: (mixture coeficient, rate categories, alpha)
> init.pars  = c(0.7,3,0.9,pars)
> #- Maximum likelihood:
> pars.ml = optimize.mcTreeLike.3states.GTR.bhom.gamma.inv(
```

```
+          pars    = init.pars,
+          ali     = ali,
+          mult    = 1,
+          edgeMat = edgeMat)
```