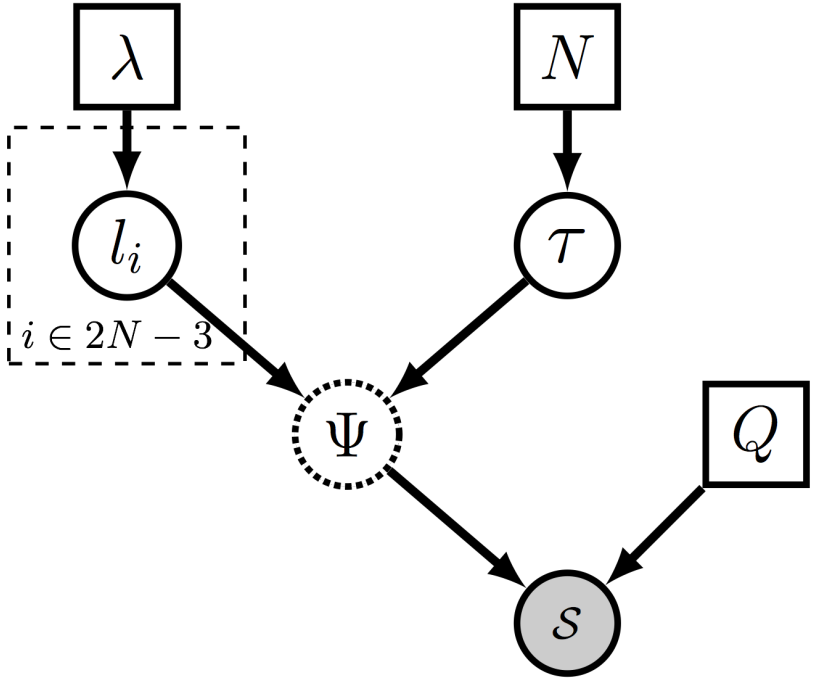
Tracy A. Heath

@trayc7
http://phyloworks.org

Probabilistic graphical models

Probabilistic Graphical Models

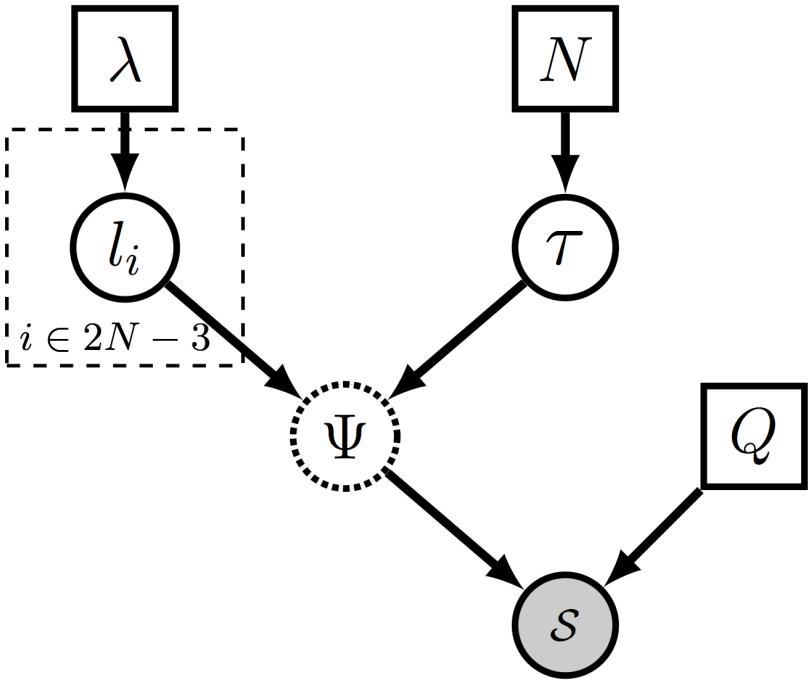
A representation of a model as a directed acyclic graph that exposes all of the assumptions and conditional dependence structure



Höhna et al. 2014. RevBayes: Probabilistic graphical model representation in phylogenetics. *Systematic Biology*. (doi: 10.1093/sysbio/syu039)

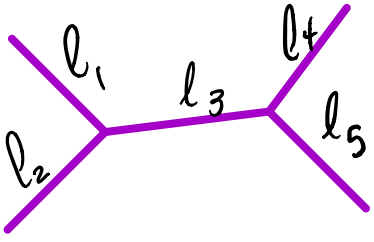
Probabilistic Graphical Models

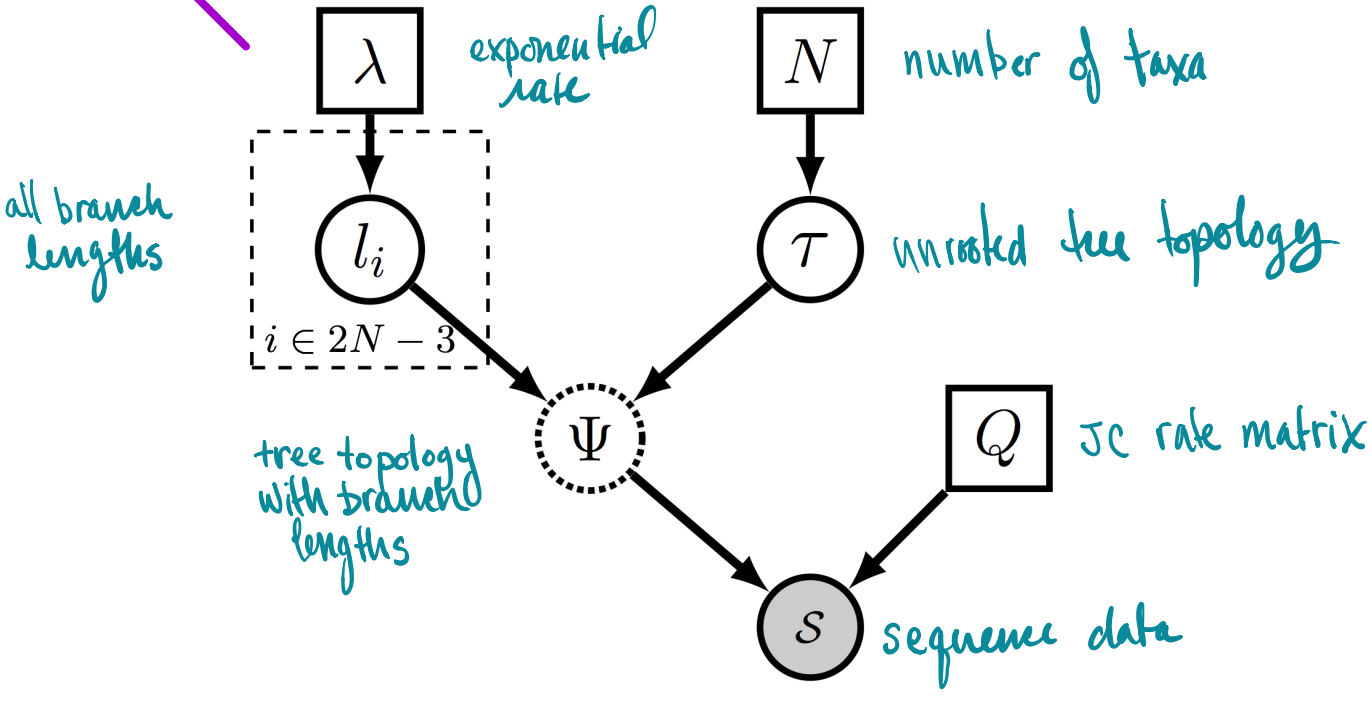
This model: observed sequence data are generated by a Jukes-Cantor substitution process, with a uniform distribution over topologies and branch lengths drawn from an exponential distribution



Höhna et al. 2014. RevBayes: Probabilistic graphical model representation in phylogenetics. *Systematic Biology*. (doi: 10.1093/sysbio/syu039)

Probabilistic Graphical Models

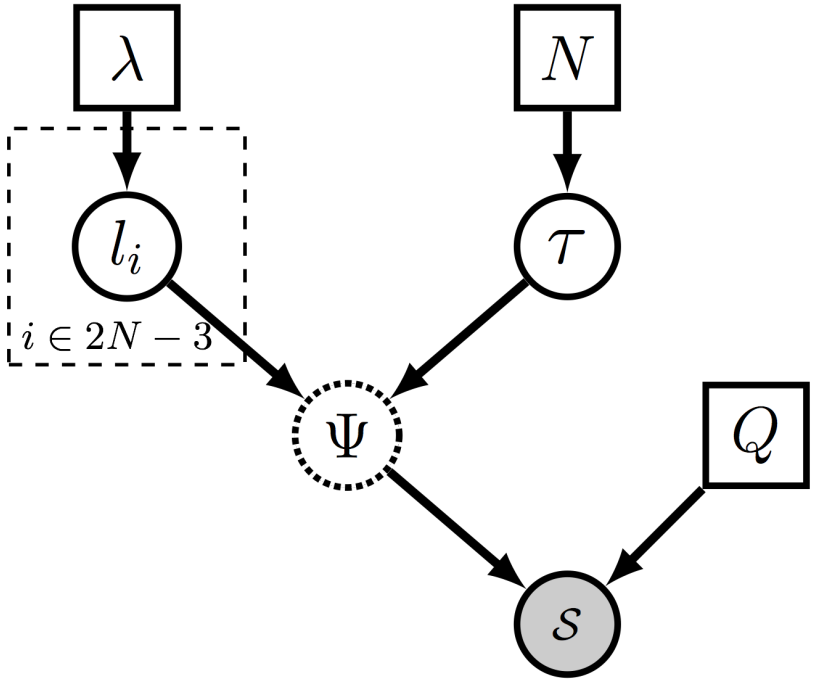




Components of a Graphical Model

A probabilistic graphical model is a visual representation of the model structure

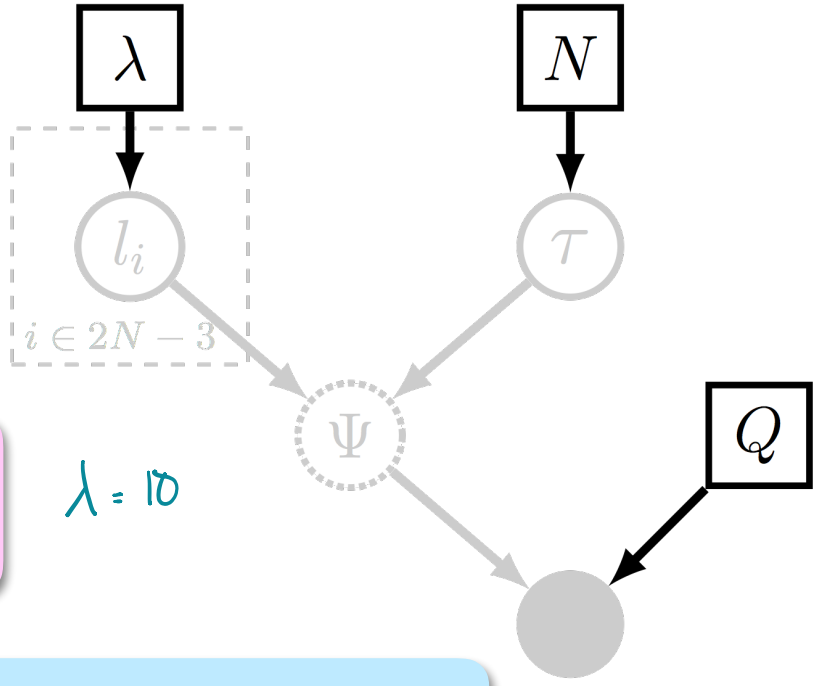
The elements of the graph denote the different kinds of variables in the model



Components of a Graphical Model

Constant node:
"equals"

represents a fixed value that is asserted or known



value of a prior parameter `λ <- 10`

$\lambda = 10$

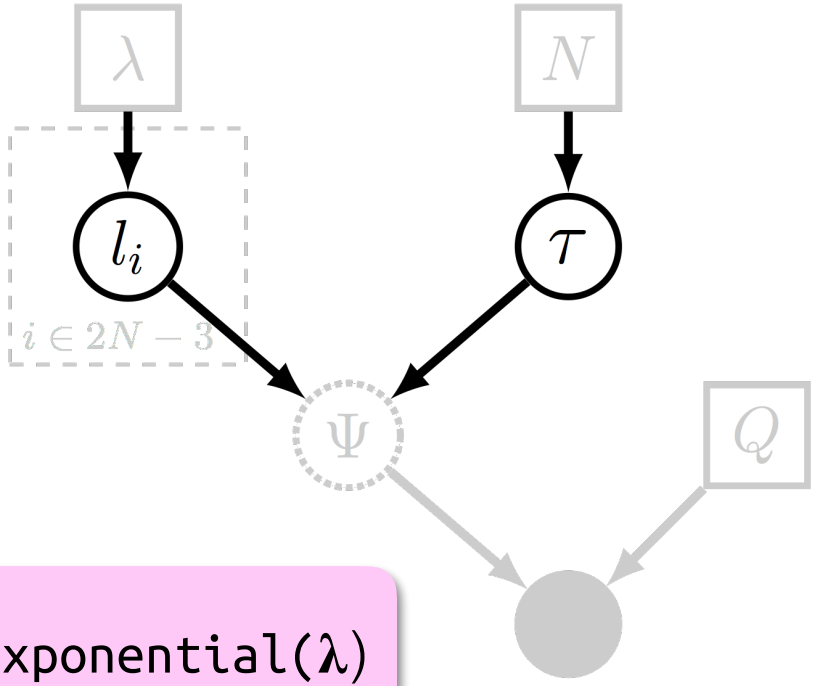
dimension of a parameter `N <- 127`

$N = 127$

Components of a Graphical Model

Stochastic node:
"distributed by"

represents a random variable that is unknown and estimated



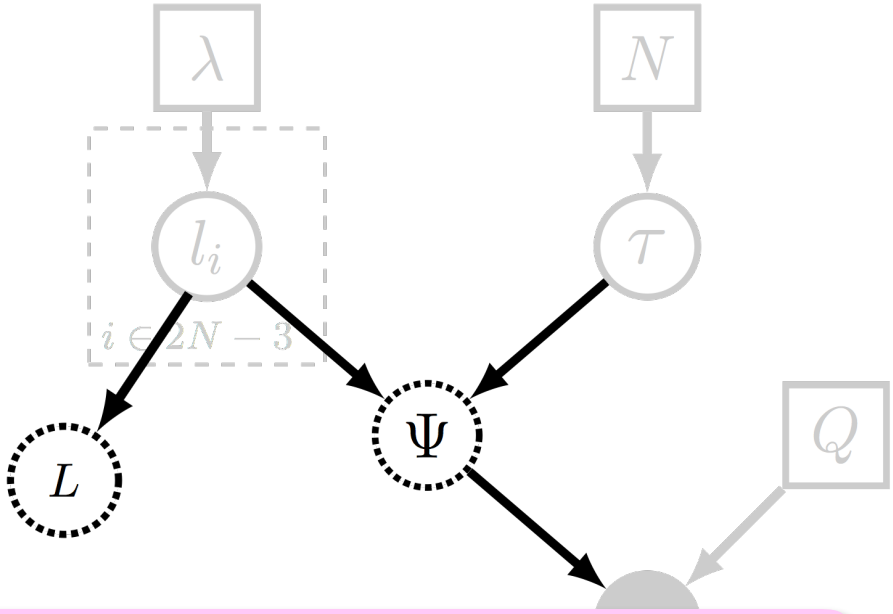
parameters that have a distribution

```
li ~ Exponential(λ)
```

Components of a Graphical Model

Deterministic node:
"determined by"

represents an unknown random variable estimated



$$L := \sum_{i=1}^{2N-3} l_i$$

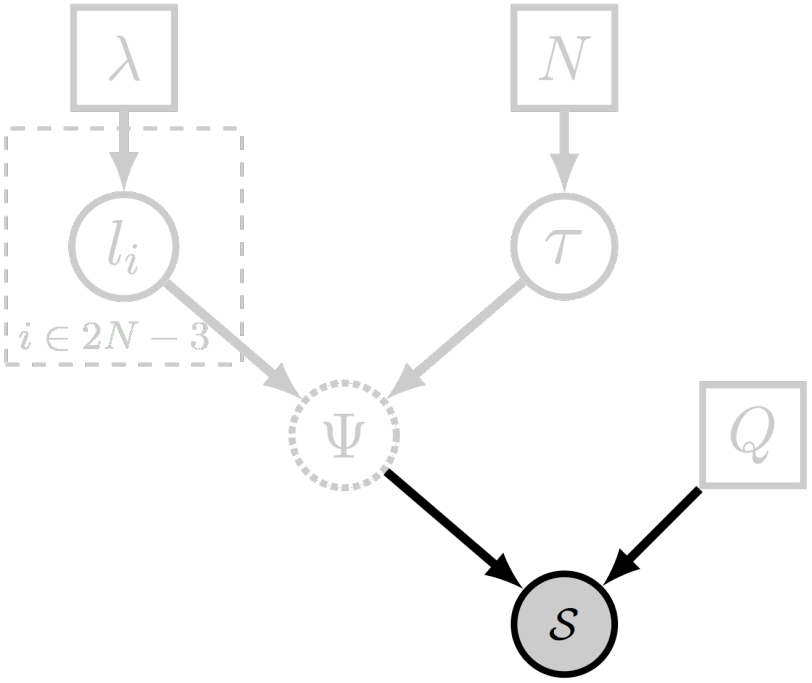
parameters that are a function of other parameters

$\Psi :=$ treeAssembly(l, τ)

Components of a Graphical Model

Observed stochastic node: "distributed by"

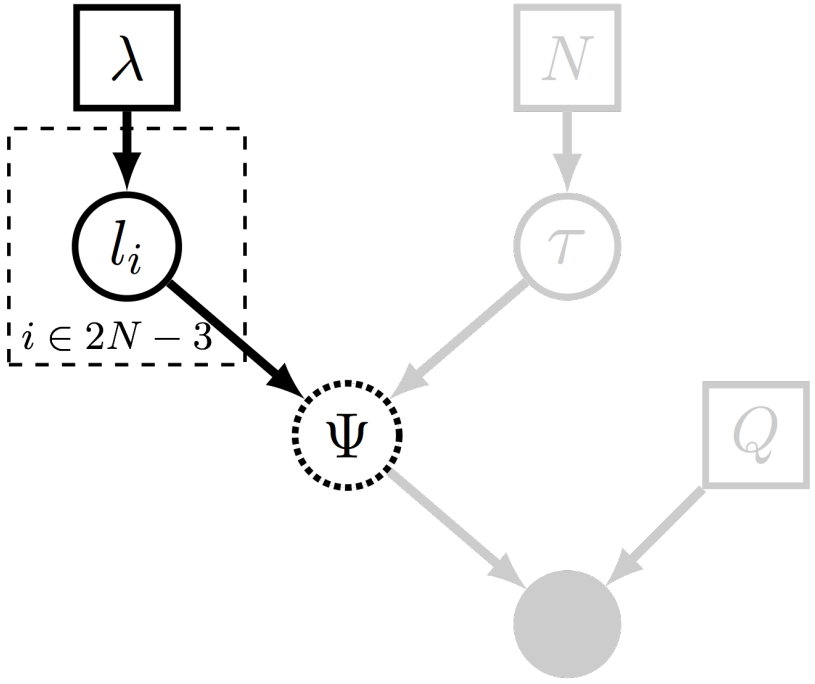
represents a random variable that is the observed outcome of the model and fixed to observed value



Components of a Graphical Model

Plate: repetition

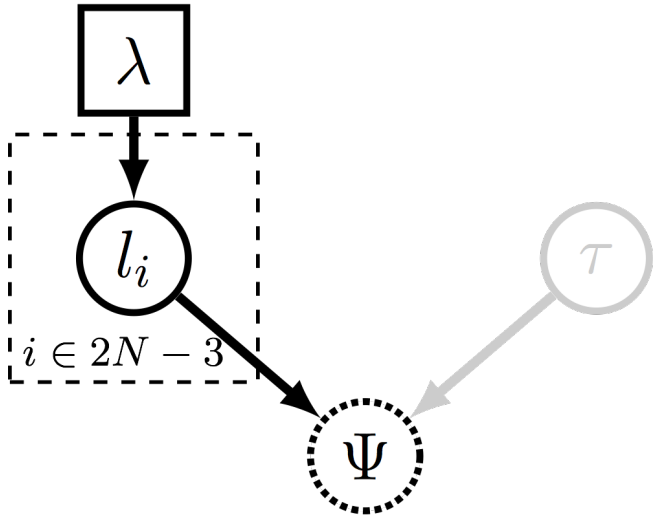
repeats model structure to simplify visualization



Components of a Graphical Model

Plate: repetition

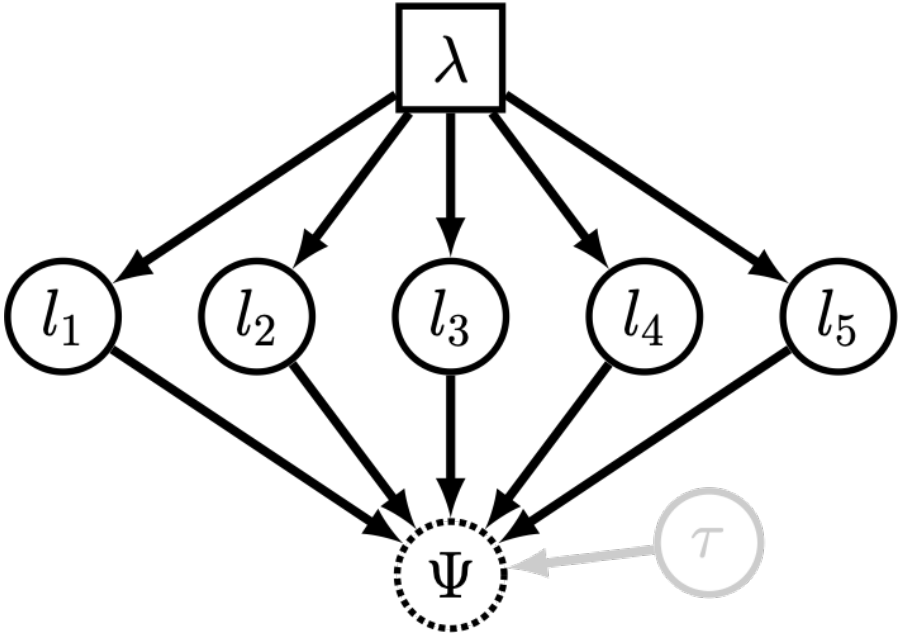
repeats model structure to simplify visualization



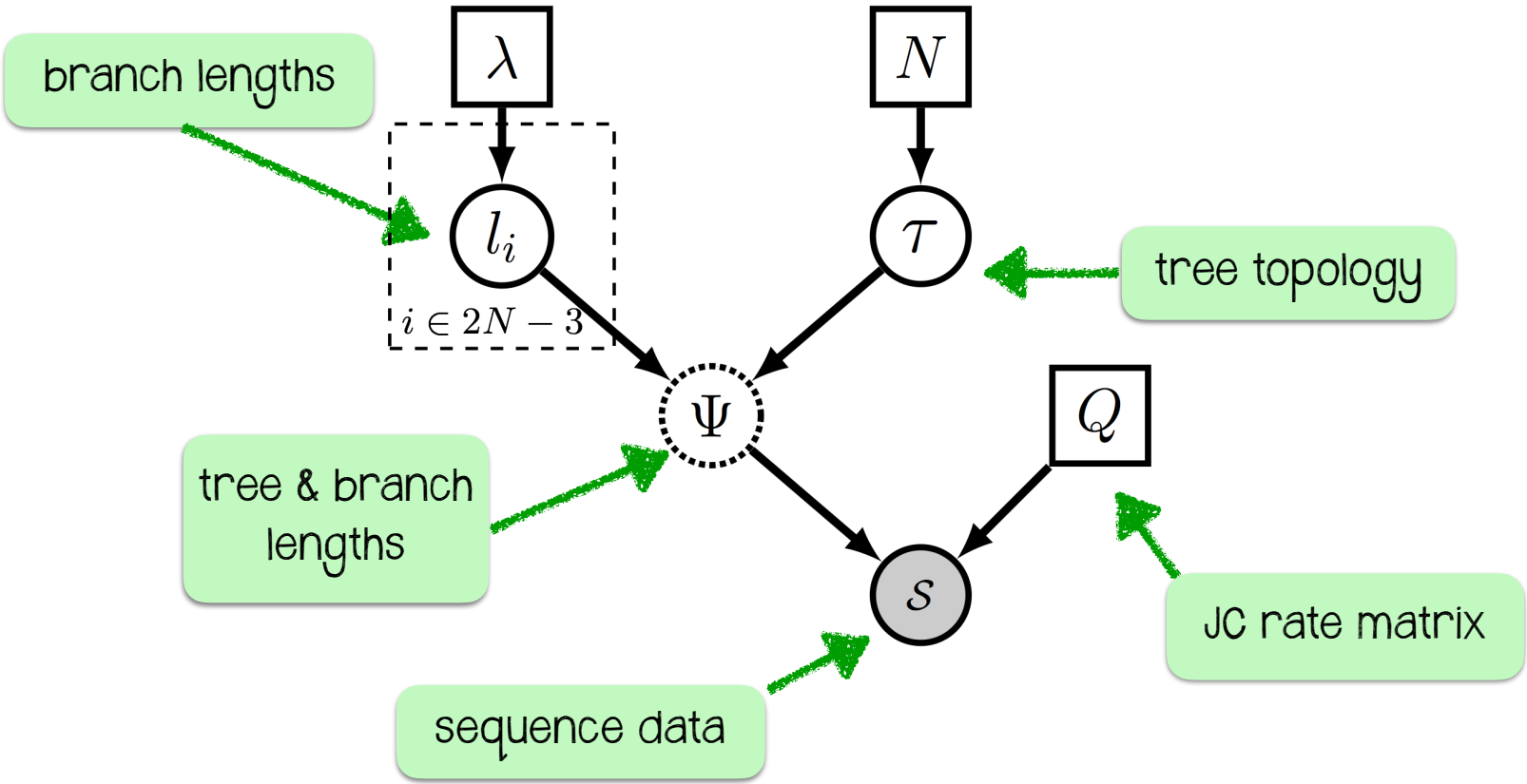
Components of a Graphical Model

Plate: repetition

repeats model structure to simplify visualization



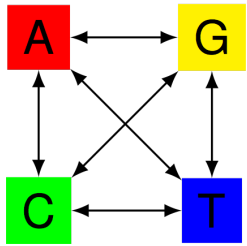
Probabilistic Graphical Models



Jukes-Cantor Model

Substitution model: with equal rates of substitution between bases and equal base frequencies

graphical model



$$Q_{JC} = \begin{pmatrix} - & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & - & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & - & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & - \end{pmatrix}$$

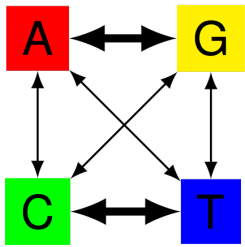
```
Q <- fnJC(4)
```


(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

Kimura 2-Parameter Model

Transitions happen at a higher rate than transversions and equal base frequencies

graphical model



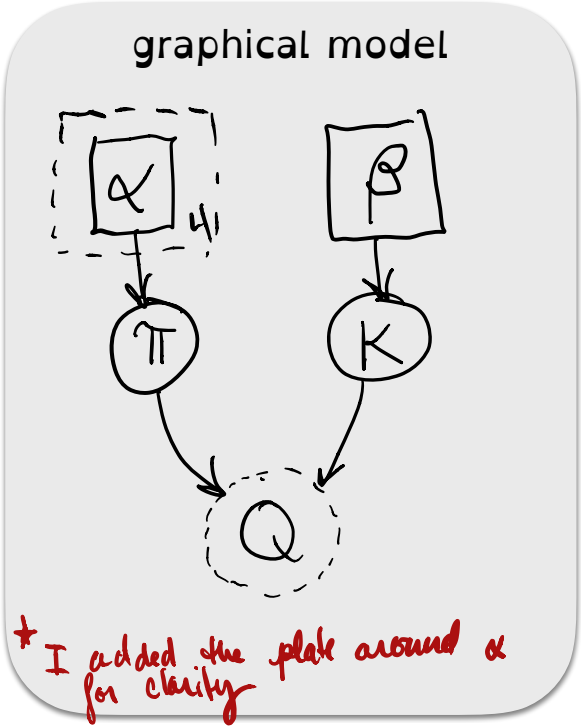
$$Q_{K80} = \begin{pmatrix} - & 1 & \kappa & 1 \\ 1 & - & 1 & \kappa \\ \kappa & 1 & - & 1 \\ 1 & \kappa & 1 & - \end{pmatrix}$$

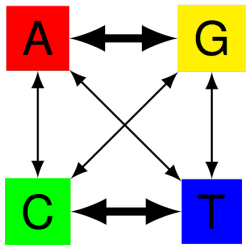
```
beta <- 1
kappa ~ dnGamma(beta,beta)

Q := fnK80(kappa)
```


(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

HKY Model

Transitions happen at a higher rate than transversions, with unequal base frequencies





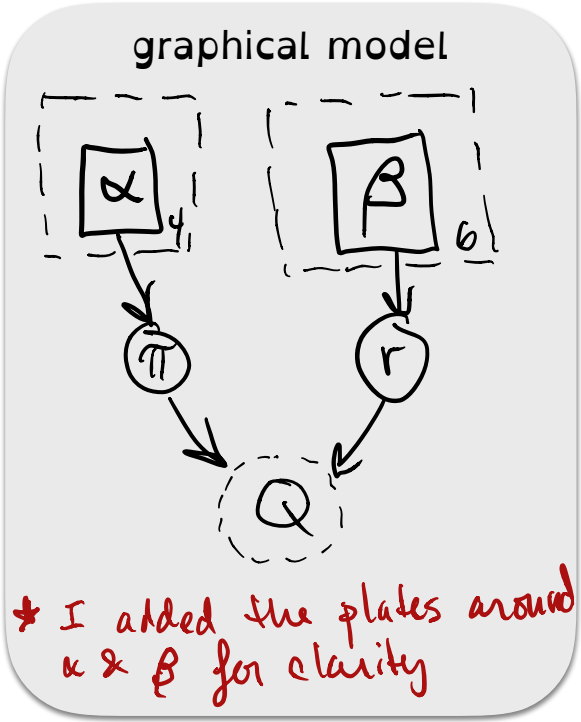
$$Q_{HKY} = \begin{pmatrix} - & \pi_C & \kappa\pi_G & \pi_T \\ \pi_A & - & \pi_G & \kappa\pi_T \\ \kappa\pi_A & \pi_C & - & \pi_T \\ \pi_A & \kappa\pi_C & \pi_G & - \end{pmatrix}$$

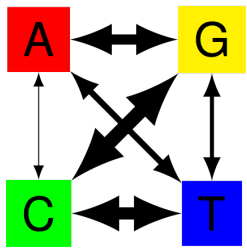
```
alpha <- [1, 1, 1, 1]
pi ~ dnDirichlet(alpha)

beta <- 1
kappa ~ dnGamma(beta, beta)

Q := fnHKY(pi, kappa)
```


(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

General Time Reversible Model

Unequal substitution rates and unequal base frequencies



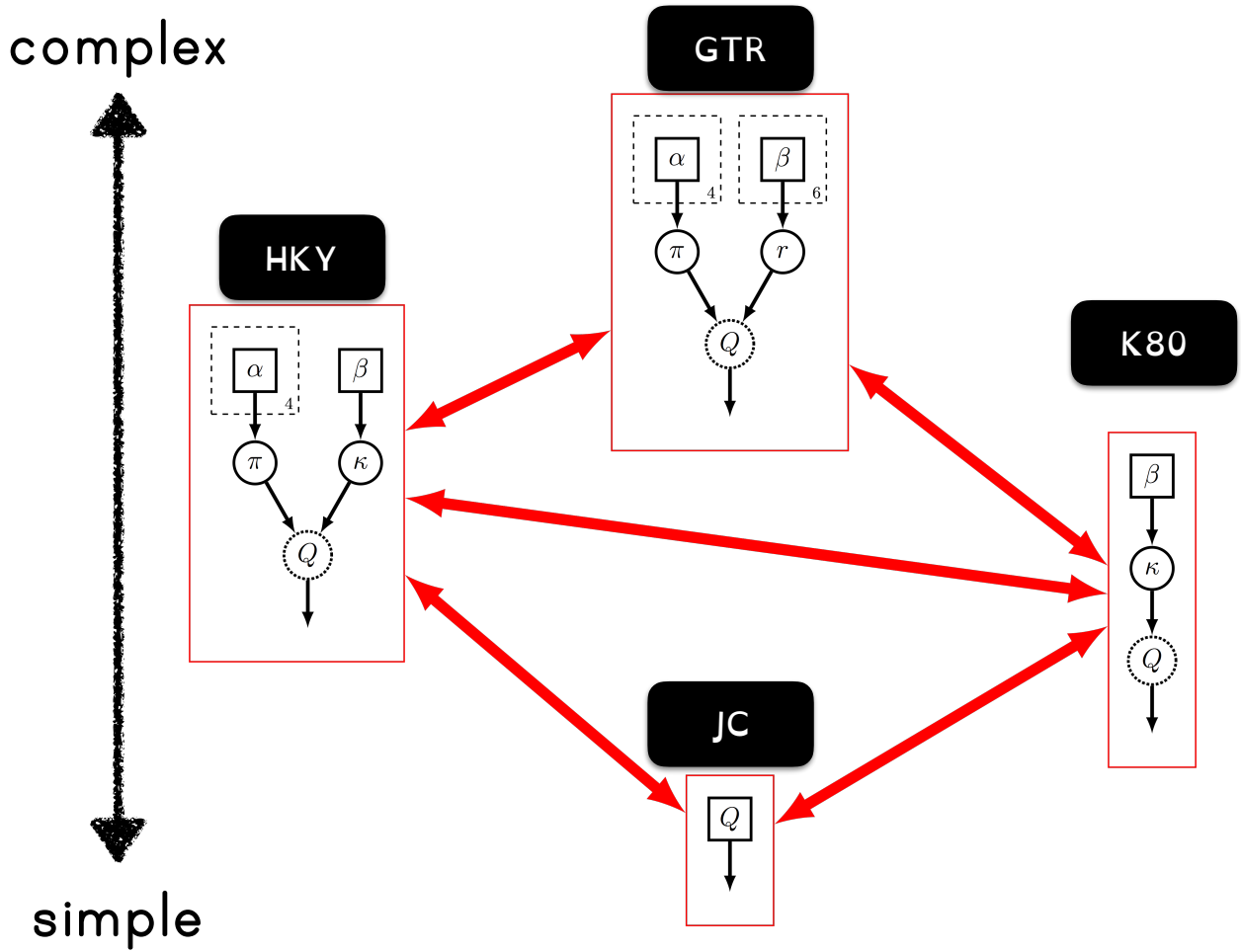


$$Q_{GTR} = \begin{pmatrix} - & r_{AC}\pi_C & r_{AG}\pi_G & r_{AT}\pi_T \\ r_{AC}\pi_A & - & r_{CG}\pi_G & r_{GT}\pi_T \\ r_{AG}\pi_A & r_{CG}\pi_C & - & r_{GT}\pi_T \\ r_{AT}\pi_A & r_{CT}\pi_C & r_{GT}\pi_G & - \end{pmatrix}$$

```
alpha <- [1, 1, 1, 1]
pi ~ dnDirichlet(alpha)

beta <- [1, 1, 1, 1, 1, 1]
r ~ dnDirichlet(beta)

Q := fnGTR(pi, r)
```

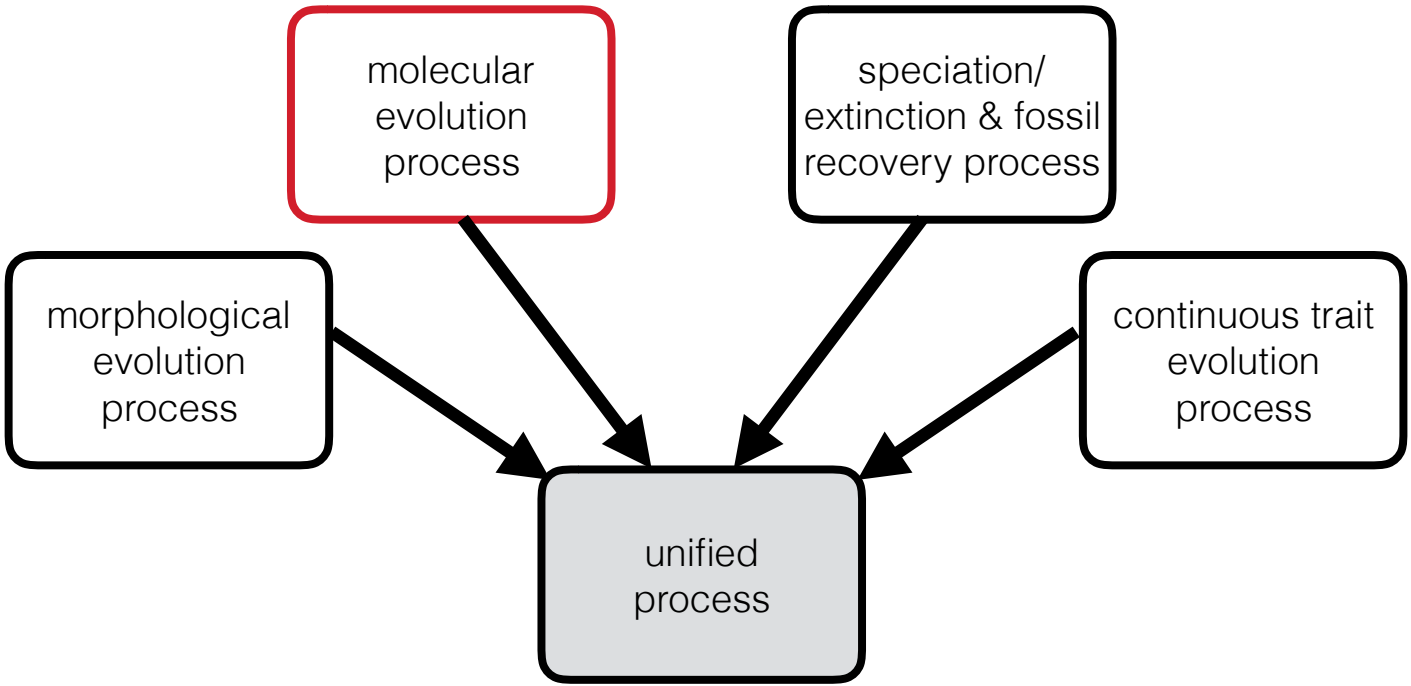

(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

Model Hierarchy



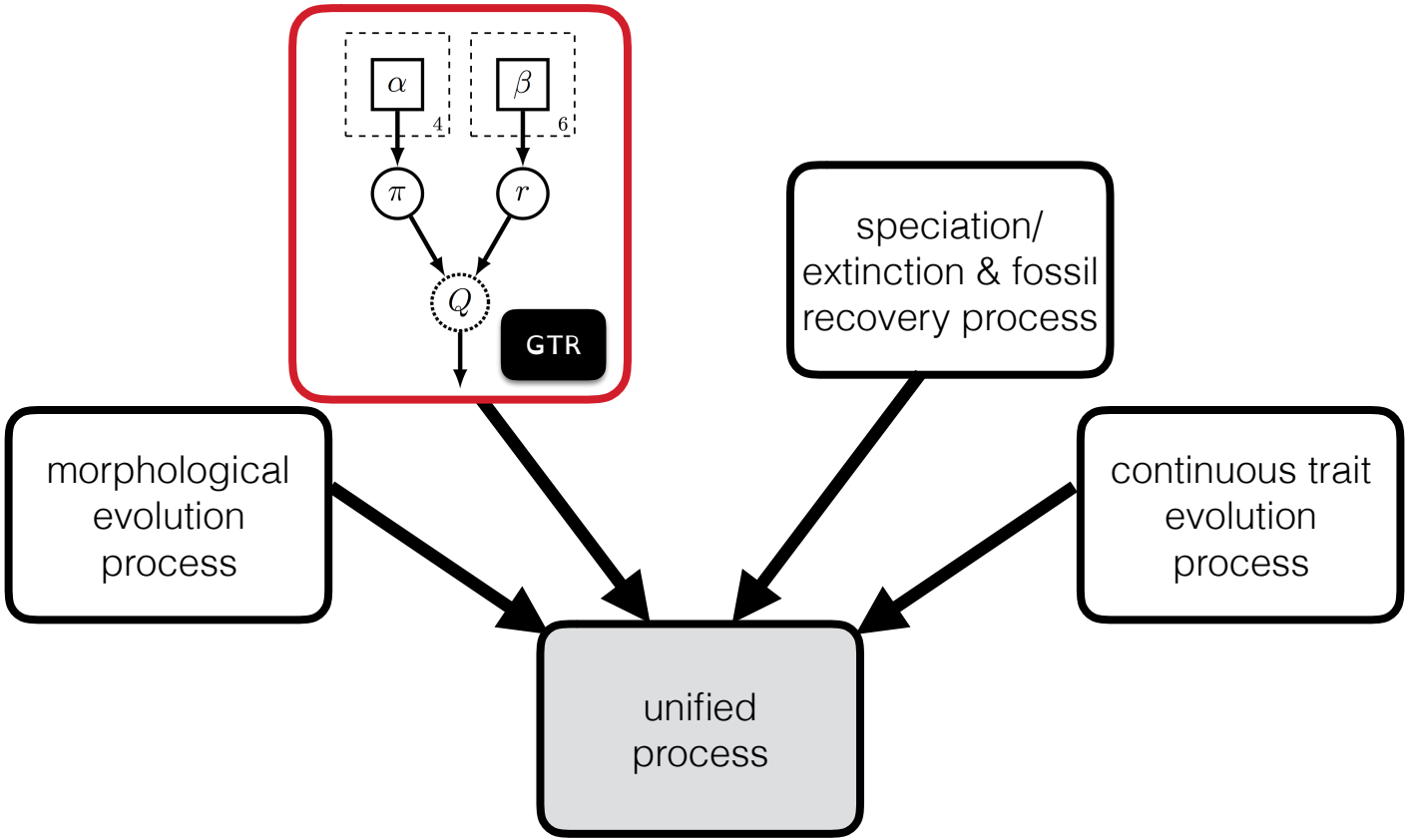
(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

Model Modularity



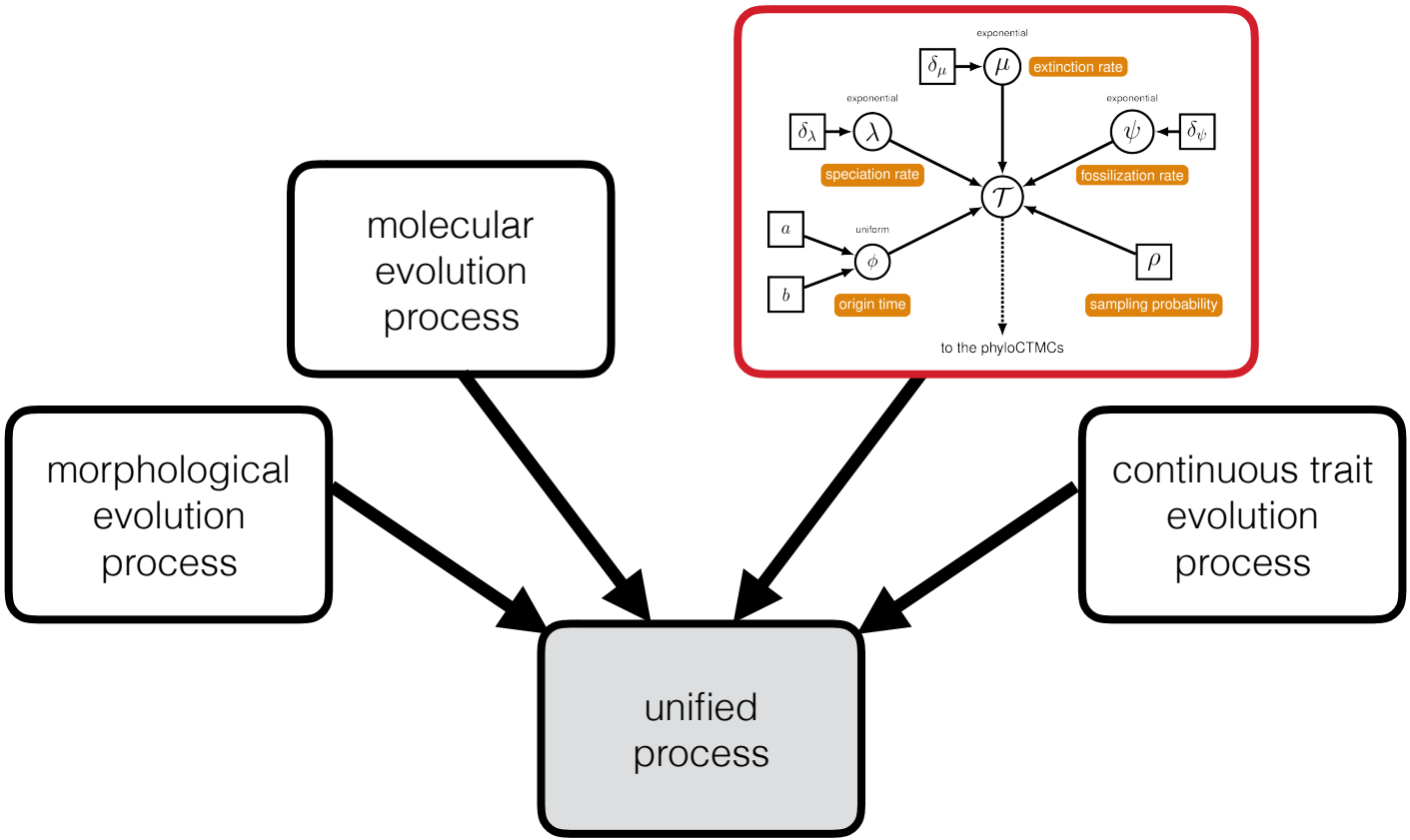
(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

Model Modularity



(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)

Model Modularity



(based on slides by M. Landis: http://bit.ly/2Hzr9Hu)