Introduction to Phylogenetic Inference in RevBayes

Tracy A. Heath

Iowa State University

@trayc7

Michael J. Landis

University of California, Berkeley

@landismj

2015 Workshop on Molecular Evolution

Woods Hole, MA

Outline

Overview – Heath

Introduction to RevBayes

- Motivation
- Probabilistic graphical models
- The Rev language and demo

short break

Demo & Tutorial – Landis

Phylogenetic reconstruction in RevBayes

- Demo: tree reconstruction using MCMC under JC
- Tutorial (on your own): specify the HKY model, sample using MCMC, summarize the tree

beer(s)

Challenges of Statistical Programming

Prior options in MrBayes v3.2

```
Parameter        Options                      Current Setting
------------------------------------------------------------------
Tratiopr         Beta/Fixed                   Beta(1.0,1.0)
Revmatpr         Dirichlet/Fixed              Dirichlet(1.0,1.0,1.0,1.0,1.0,1.0)
Aamodelpr        Fixed/Mixed                  Fixed(Poisson)
Aarevmatpr       Dirichlet/Fixed              Dirichlet(1.0,1.0,...)
Omegapr          Dirichlet/Fixed              Dirichlet(1.0,1.0)
Ny98omega1pr     Beta/Fixed                   Beta(1.0,1.0)
Ny98omega3pr     Uniform/Exponential/Fixed    Exponential(1.0)
M3omegapr        Exponential/Fixed            Exponential
Codoncatfreqs    Dirichlet/Fixed              Dirichlet(1.0,1.0,1.0)
Statefreqpr      Dirichlet/Fixed              Dirichlet(1.0,1.0,1.0,1.0)
Shapepr          Uniform/Exponential/Fixed    Exponential(2.0)
Ratecorrpr       Uniform/Fixed                Uniform(-1.0,1.0)
Pinvarpr         Uniform/Fixed                Uniform(0.0,1.0)
Covswitchpr      Uniform/Exponential/Fixed    Uniform(0.0,100.0)
Symdirihyperpr   Uniform/Exponential/Fixed    Fixed(Infinity)
Topologypr       Uniform/Constraints/Fixed    Uniform
Brlenspr         Unconstrained/Clock/Fixed    Unconstrained:Exp(10.0)
Treeagepr        Gamma/Uniform/Fixed/         Gamma(1.00,1.00)
                 Truncatednormal/Lognormal/
                 Offsetlognormal/Offsetgamma/
                 Offsetexponential
```

Challenges of Statistical Programming

Prior options in MrBayes v3.2

Parameter	Options	Current Setting
Tratiopr	Beta/Fixed	Beta(1.0,1.0)
Revmatpr	Dirichlet/Fixed	Dirichlet(1.0,1.0,1.0,1.0,1.0,1.0)
Aamodelpr	Fixed/Mixed	Fixed(Poisson)
Aarevmatpr	Dirichlet/Fixed	Dirichlet(1.0,1.0,...)
Omegapr	Dirichlet/Fixed	Dirichlet(1.0,1.0)
Ny98omega1pr	Beta/Fixed	Beta(1.0,1.0)
Ny98omega3pr	Uniform/Exponential/Fixed	Exponential(1.0)
M3omegapr	Exponential/Fixed	Exponential
Codoncatfreqs	Dirichlet/Fixed	Dirichlet(1.0,1.0,1.0)
Statefreqpr	Dirichlet/Fixed	Dirichlet(1.0,1.0,1.0,1.0)
Shapepr	Uniform/Exponential/Fixed	Exponential(2.0)
Ratecorrpr	Uniform/Fixed	Uniform(-1.0,1.0)
Pinvarpr	Uniform/Fixed	Uniform(0.0,1.0)
Covswitchpr	Uniform/Exponential/Fixed	Uniform(0.0,100.0)
Symdirihyperpr	Uniform/Exponential/Fixed	Fixed(Infinity)
Topologypr	Uniform/Constraints/Fixed	Uniform
Brlenspr	Unconstrained/Clock/Fixed	Unconstrained:Exp(10.0)
Treeagepr	Gamma/Uniform/Fixed/ Truncatednormal/Lognormal/ Offsetlognormal/Offsetgamma/ Offsetexponential	Gamma(1.00,1.00)

Modular Bayesian Phylogenetic Software

Several software packages in phylogenetics are moving toward a more modular framework

- reuse code
- easier to extend existing models and implement new models through a rich, language-based interface
- provides a unified framework for analyses under complex models

RevBayes

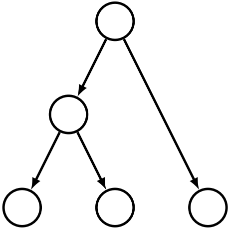
Bali-Phy

BEAST2

RevBayes

Fully integrative Bayesian inference of phylogenetic parameters

http://revbayes.com



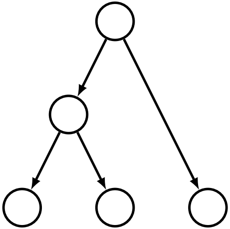
Development team

Höhna	Landis	Heath	Boussau
Lartillot	Huelsenbeck	Ronquist	& others...

Graphical Models in RevBayes

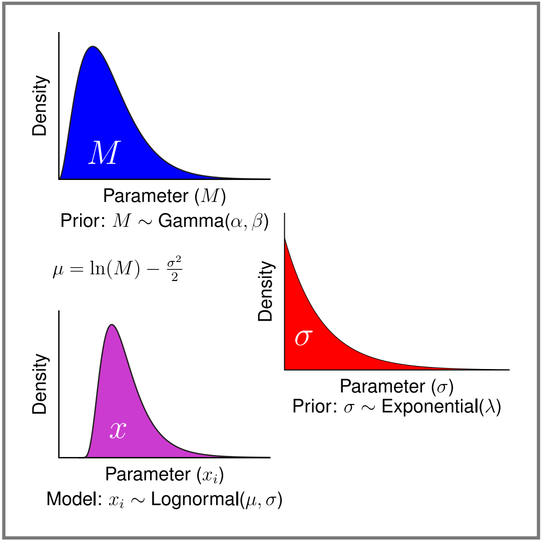
Graphical models provide tools for visually & computationally representing complex, parameter-rich probabilistic models

We can depict the conditional dependence structure of various parameters and other random variables

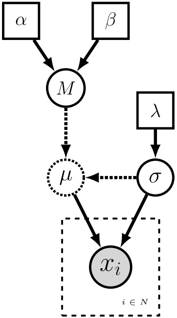


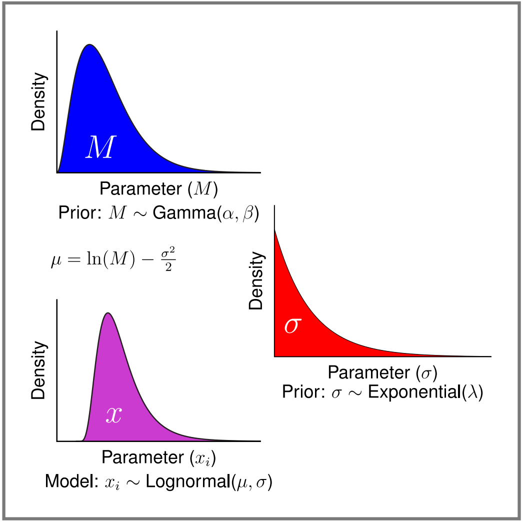
Höhna, Heath, Boussau, Landis, Ronquist, Huelsenbeck. 2014.
Probabilistic Graphical Model Representation in Phylogenetics.
Systematic Biology. (doi: 10.1093/sysbio/syu039)

Graphical Models in RevBayes

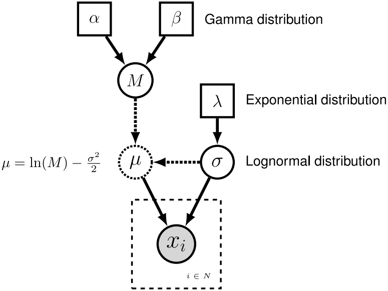


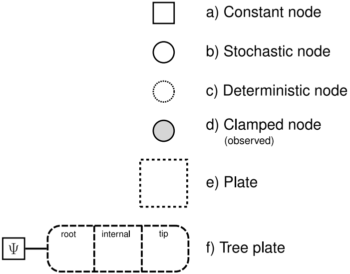
Graphical Models in RevBayes



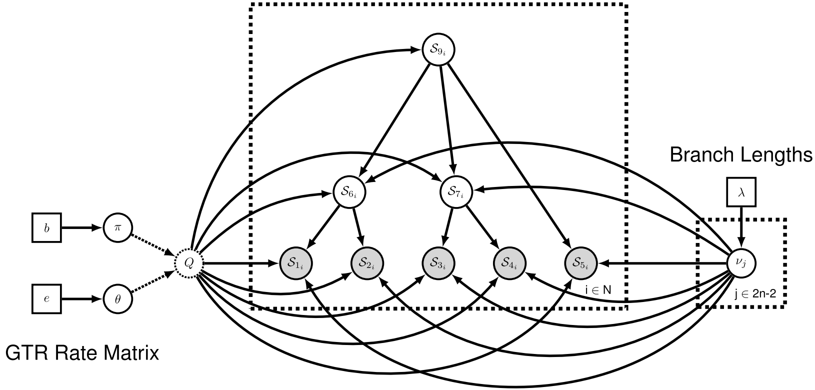


Graphical Models in RevBayes

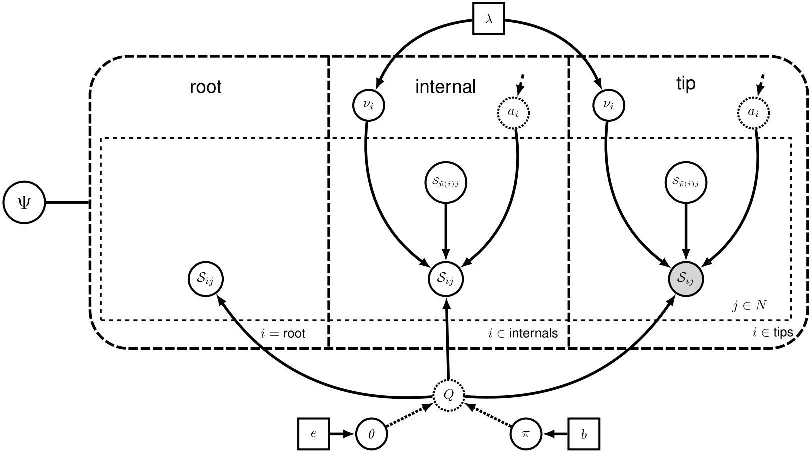




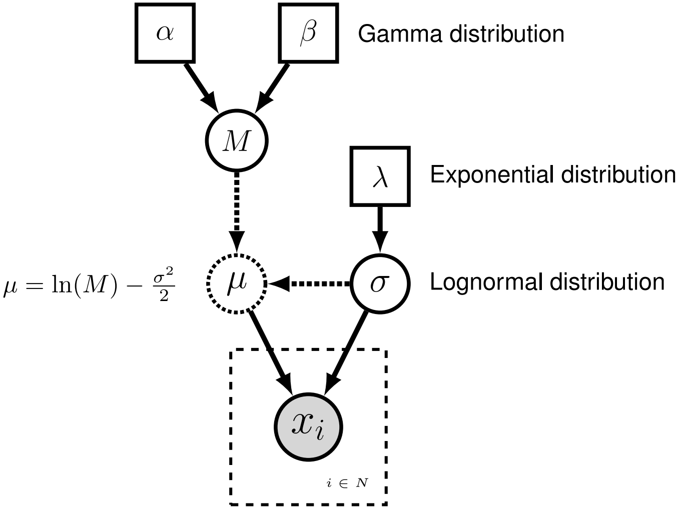
Graphical Models in RevBayes



Graphical Models in RevBayes



Graphical Models in RevBayes



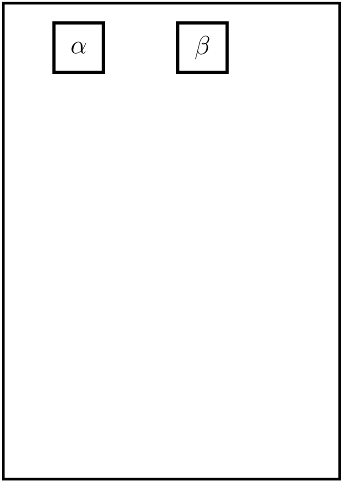
Graphical Models in RevBayes

Defining the model in the Rev language

```
observations <- [<your data go here>]
```

Graphical Models in RevBayes

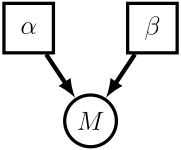
Defining the model in the Rev language




```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
```

Graphical Models in RevBayes

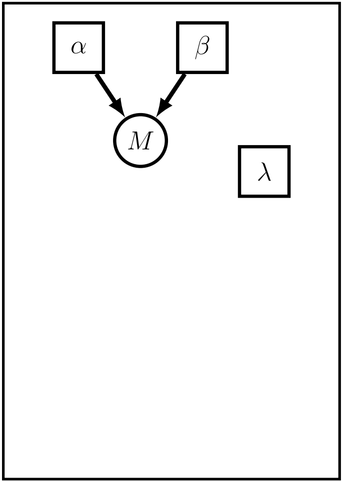
Defining the model in the Rev language




```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
M ~ dnGamma(alpha, beta)
```

Graphical Models in RevBayes

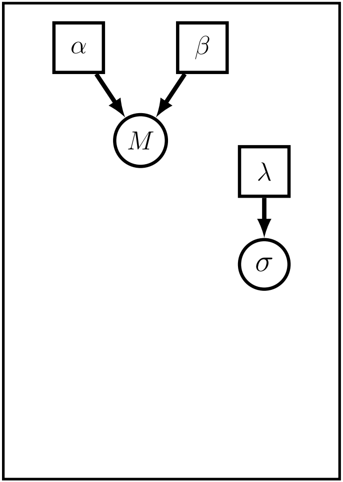
Defining the model in the Rev language




```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
M ~ dnGamma(alpha, beta)


lambda <- 1.0
```

Graphical Models in RevBayes

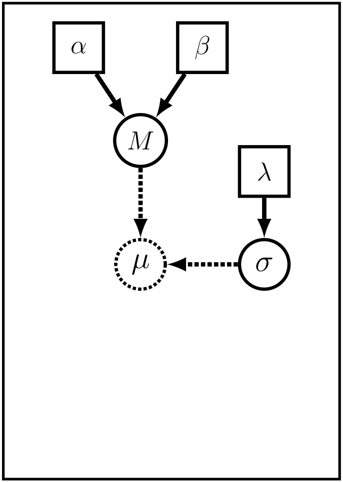
Defining the model in the Rev language




```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
M ~ dnGamma(alpha, beta)


lambda <- 1.0
sigma ~ dnExponential(lambda)
```

Graphical Models in RevBayes

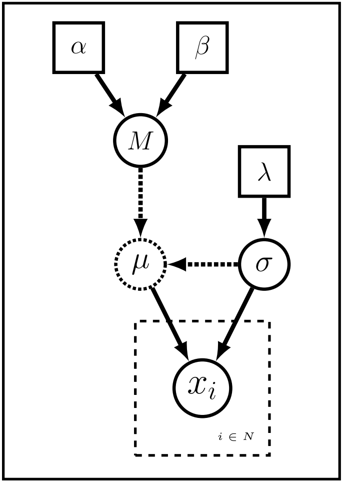
Defining the model in the Rev language




```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
M ~ dnGamma(alpha, beta)

lambda <- 1.0
sigma ~ dnExponential(lambda)

mu := ln(M) - (power(sigma, 2.0) / 2.0)
```

Graphical Models in RevBayes

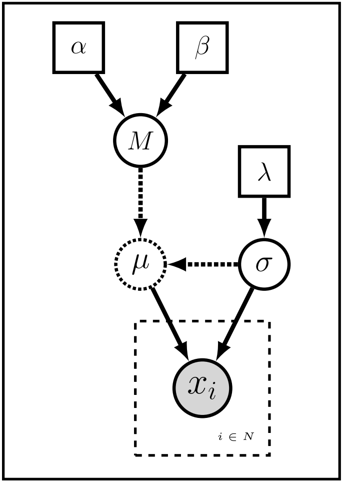
Defining the model in the Rev language




```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
M ~ dnGamma(alpha, beta)


lambda <- 1.0
sigma ~ dnExponential(lambda)


mu := ln(M) - (power(sigma, 2.0) / 2.0)


N <- observations.size()
for( i in 1:N ){
  x[i] ~ dnLnorm(mu, sigma)

}
```

Graphical Models in RevBayes

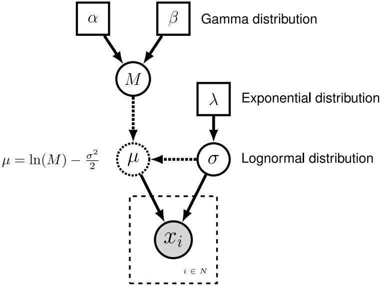
Defining the model in the Rev language

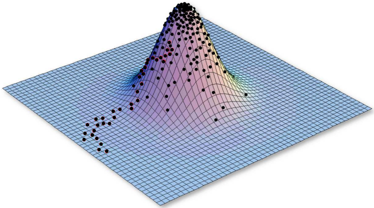



```
observations <- [<your data go here>]

alpha <- 3.0
beta <- 1.0
M ~ dnGamma(alpha, beta)

lambda <- 1.0
sigma ~ dnExponential(lambda)

mu := ln(M) - (power(sigma, 2.0) / 2.0)

N <- observations.size()
for( i in 1:N ){
  x[i] ~ dnLnorm(mu, sigma)
  x[i].clamp(observations[i])
}
```

RevBayes Demo: A Simple Model

Use MCMC to approximate the posterior distributions of stochastic and deterministic variables





RevBayes

http://revbayes.com

- Downloads
- Tutorials
- Help documentation
- User forum
- Source code:
  https://github.com/revbayes/revbayes

Integrative Bayesian Modeling in RevBayes