I. Actual Studies

- Conventional Time Estimation
- Bayesian Time Estimation

II. MCMCTREE Analysis

- Theory
- Tutorial

Characteristics of MCMCTREE

- Command line.
  (Windows: command prompt; Mac: terminal).
- Two steps:
  Branch lengths/variance-covariance estimation.
  Bayesian analysis.
- DNA and amino acid data
- Compatible with genome-wide data.
- Partitioning and several models.
- Topology should be estimated before analysis.

Bayesian Method: Lack of Acceptance

Calculation of posterior requires high-dimensional integrals.

↓

1990s:
Markov chain Monte Carlo (MCMC)
Development of computer.

↓

MCMC has made it possible to implement sophisticated parameter-rich models for which the likelihood analysis would not be feasible
=> Divergence time estimation.

Yang (2006) P146

Prior and Posterior Distribution

Distribution of the parameter – **Prior**: without sequence data.
Posterior: with sequence data.

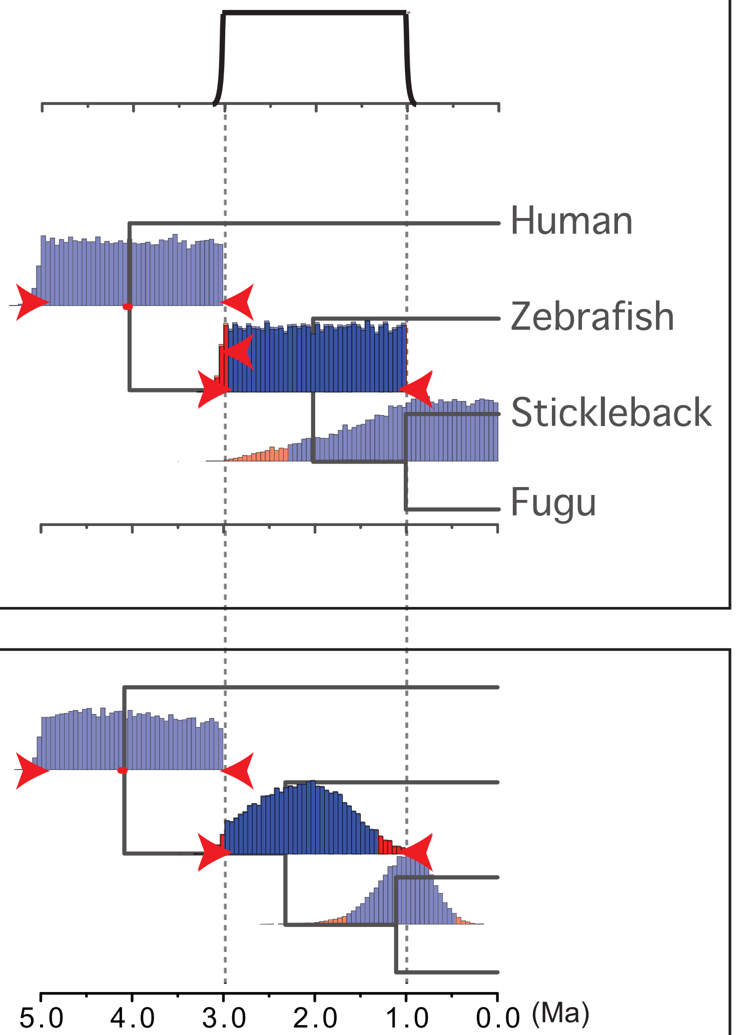
Prior

Uniform distribution assumed before analysis

Estimated distribution without sequence data

Posterior

Estimated distribution with sequence data



Bayesian analysis aims to estimate the posterior distribution.

Posterior Distribution is Given as:

Likelihood — Prior distribution

$$f(\mathbf{t}, \mathbf{r}, \theta | X) = \frac{f(\mathbf{X}|\mathbf{t}, \mathbf{r}, \theta)\ f(\mathbf{r}|\mathbf{t}, \theta)\, f(\mathbf{t}|\theta)\, f(\theta)}{f(X)}$$

Posteriors distribution

X: Sequence data.

$\mathbf{t}$: Internal node times ($T_0, T_1, \ldots. T_k$).

$\mathbf{r}$: Rates. ($R_0, R_1, \ldots. R_k$).

θ : Other parameters.

Marginal probability of the data:
A high-dimensional integral over t, r, θ.
=> Difficult to calculate.

↓

Metropolis-Hastings Algorithms is used to approximate $f(\mathbf{t}, \mathbf{r}, \theta | X)$

Yang (2006) P246

Metropolis-Hastings Algorithm

Yang (2006) P163

Likelihood — Prior distribution

$$f(\mathbf{t}, \mathbf{r}, \theta|X) = \frac{f(X|\mathbf{t}, \mathbf{r}, \theta)\ f(\mathbf{r}|\mathbf{t}, \theta)\, f(\mathbf{t}|\theta)\, f(\theta)}{f(X)}$$

Posterior distribution

Marginal distribution of the data

Metropolis-Hastings algorithm

Compare only numerator

Next state

$$R = \min\left\{1, \frac{f(X|\mathbf{t}^*, \mathbf{r}^*)\, f(\mathbf{r}^*|\mathbf{t}^*, \theta^*)\, f(\mathbf{t}^*|\theta^*)\, f(\theta^*)}{f(X|\mathbf{t}, \mathbf{r})\, f(\mathbf{r}|\mathbf{t}, \theta)\, f(\mathbf{t}|\theta)\, f(\theta)} \times \frac{q(\theta, \mathbf{t}, \mathbf{r}|\theta^*, \mathbf{t}^*, \mathbf{r}^*)}{q(\theta^*, \mathbf{t}^*, \mathbf{r}^*|\theta, \mathbf{t}, \mathbf{r})}\right\}$$

Acceptance ratio

Current state

Proposal ration :
correct the asymmetry of the proposal

Repeating many cycles of this procedure followed by acceptance/rejection
=> a Markov chain with a stationary distribution
that is the desired posterior distribution $f(\mathbf{t}, \mathbf{r}, \theta|X)$

Approximate Likelihood Calculation in MCMCTREE

Likelihood — Prior distribution

$$f(\mathbf{t}, \mathbf{r}, \theta | X) = \frac{f(X | \mathbf{t}, \mathbf{r}, \theta)\, f(\mathbf{r} | \mathbf{t}, \theta)\, f(\mathbf{t} | \theta)\, f(\theta)}{f(X)}$$

Posterior distribution — Marginal probability of the data

Calculation of the likelihood function is computationally expensive.
- To achieve computational efficiency, a normal approximation to the maximum likelihood estimates (MLEs) of branch lengths is used.

=> 2 steps:

1. Estimation of the gradient and Hessian of the branch lengths at the MLEs of the branch lengths.
2. MCMC analysis using multivariate normal distribution of MLEs of branch lengths

Thorne et al. (1998)
Yang (2006)
PAML manual

Topics

I. Actual Studies

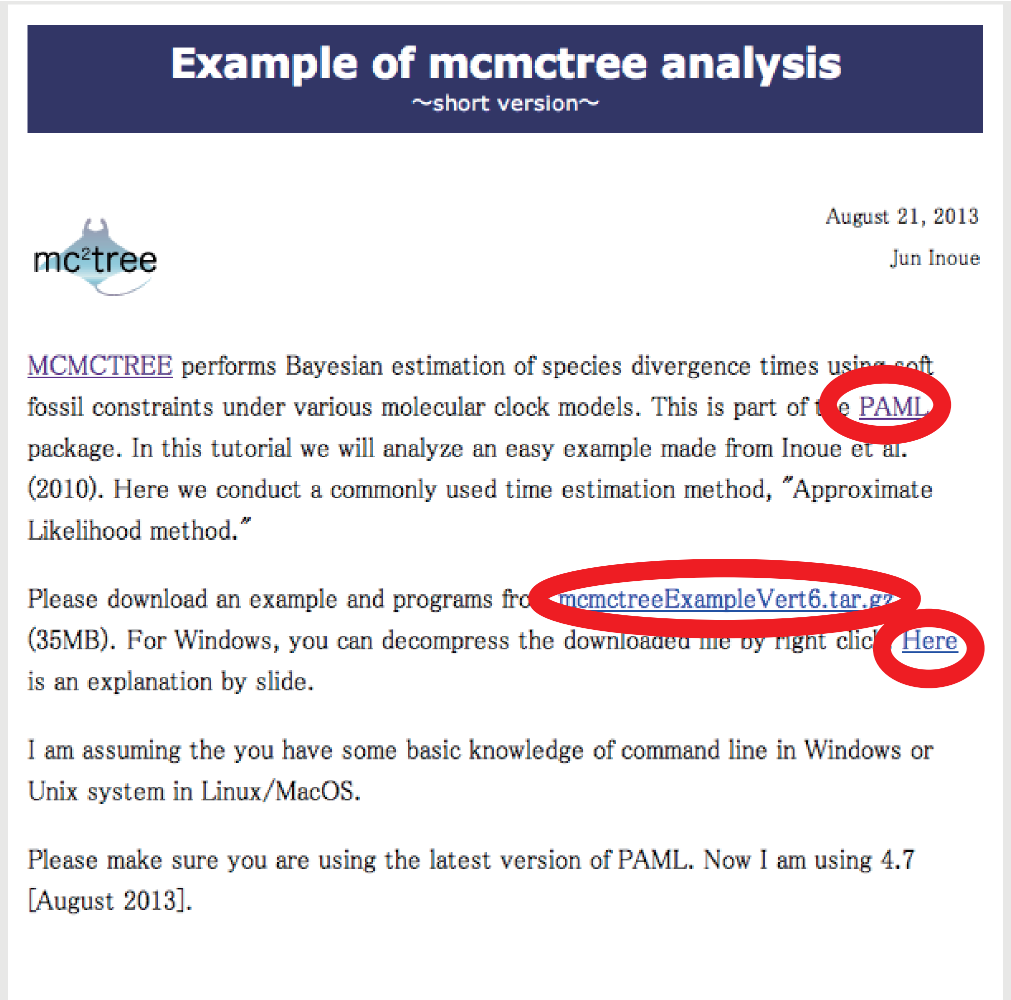
- Conventional Time Estimation
- Bayesian Time Estimation

II. MCMCTREE Analysis

- Theory
- Tutorial

MCMCTREE Tutorial

English version



Example of mcmctree analysis
~short version~

mc²tree

August 21, 2013

Jun Inoue

MCMCTREE performs Bayesian estimation of species divergence times using soft fossil constraints under various molecular clock models. This is part of the PAML package. In this tutorial we will analyze an easy example made from Inoue et al. (2010). Here we conduct a commonly used time estimation method, "Approximate Likelihood method."

Please download an example and programs from mcmctreeExampleVert6.tar.gz (35MB). For Windows, you can decompress the downloaded file by right click. Here is an explanation by slide.

I am assuming the you have some basic knowledge of command line in Windows or Unix system in Linux/MacOS.

Please make sure you are using the latest version of PAML. Now I am using 4.7 [August 2013].

Download 3 files:

Program: paml4.8.tgz

Ex. data: mcmctreeExampleVert6.tar

Slide: 000_InoueMCTenshuEng.pdf

Find a directory:

mcmctreeExampleVert6/examples

http://www.geocities.jp/ancientfishtree/mcmctreeExampleVert6/text1Eng.html

Search for "MCMCTREE-Easy Example"

Set up of example directory (Windows)

1. Download and install unzipping program, e.g. Lhaplus (free).

2. Unzip. Drag and drop mcmctreeExampleVert6.tar to the alias of Lhaplus.

3. Show extensions. Uncheck the following:
   整理 > フォルダと検索のオプション > 表示 > 登録されている拡張子は表示しない

4. Cancel the read only. Right clock mcmctreeExampleVert6/examples directory and un-check the following:
   Property > Read only

5. Open it by your editor as a trial. Right click baseml.ctl and select your editor.

6. Move in to the mcmctreeExampleVert6/examples directory from command prompt (Program >Accessory >Command prompt). Type the following commands:

```
cd Desktop
cd mcmctreeExampleVert6/examples
dir
```

Preparation of Programs

Textile manipulation

Win: ?

Mac: TextWrangler etc.

Tree drawing

FigTree

Time estimation

MCMCTREE (PAML package: paml4.7.tgz)

Win: paml4.7/bin/

```
baseml.exe,
mcmctree.exe
```

=> Copy to mcmctreeExampleVert6/examples

Mac:

from terminal,

```
cd into paml4.7/src
```

and type

```
make -f Makefile
```

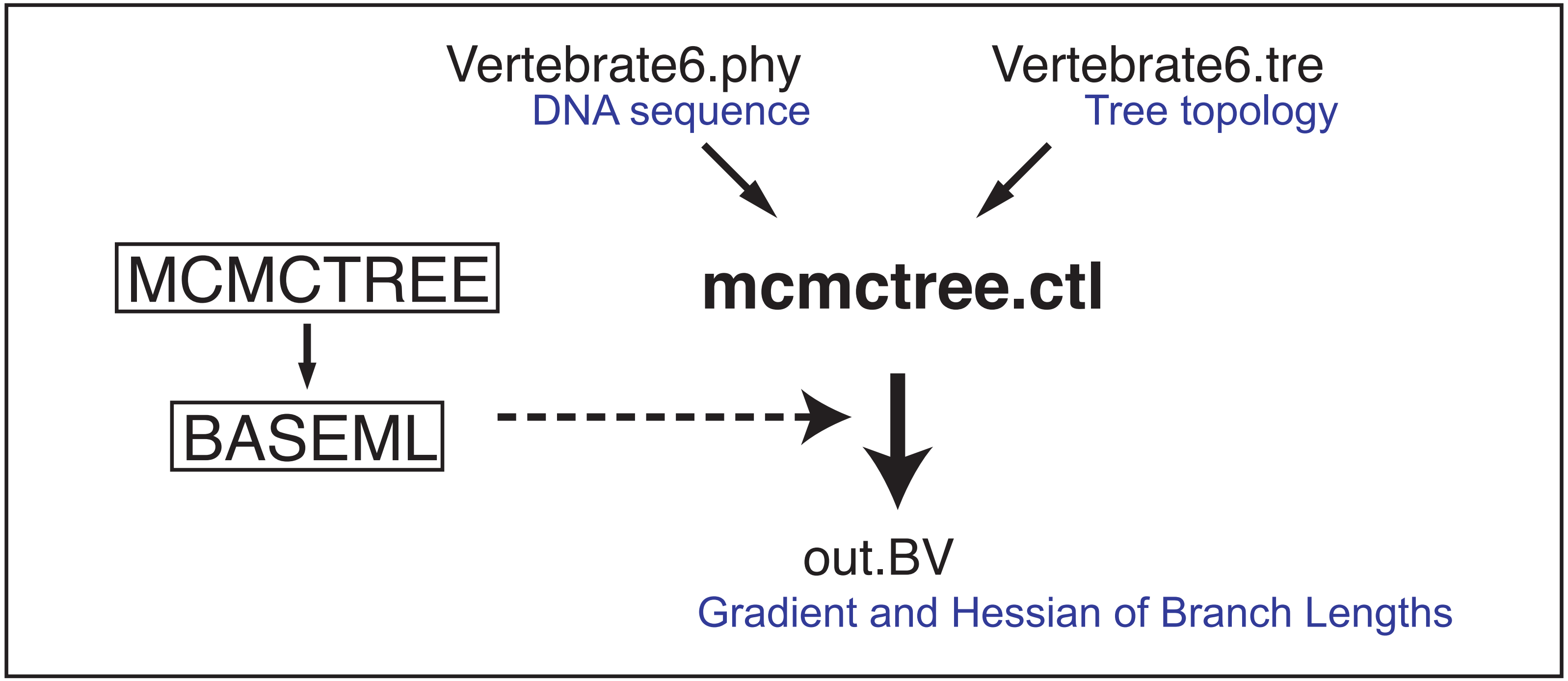
Then,

```
baseml
mcmctree
```

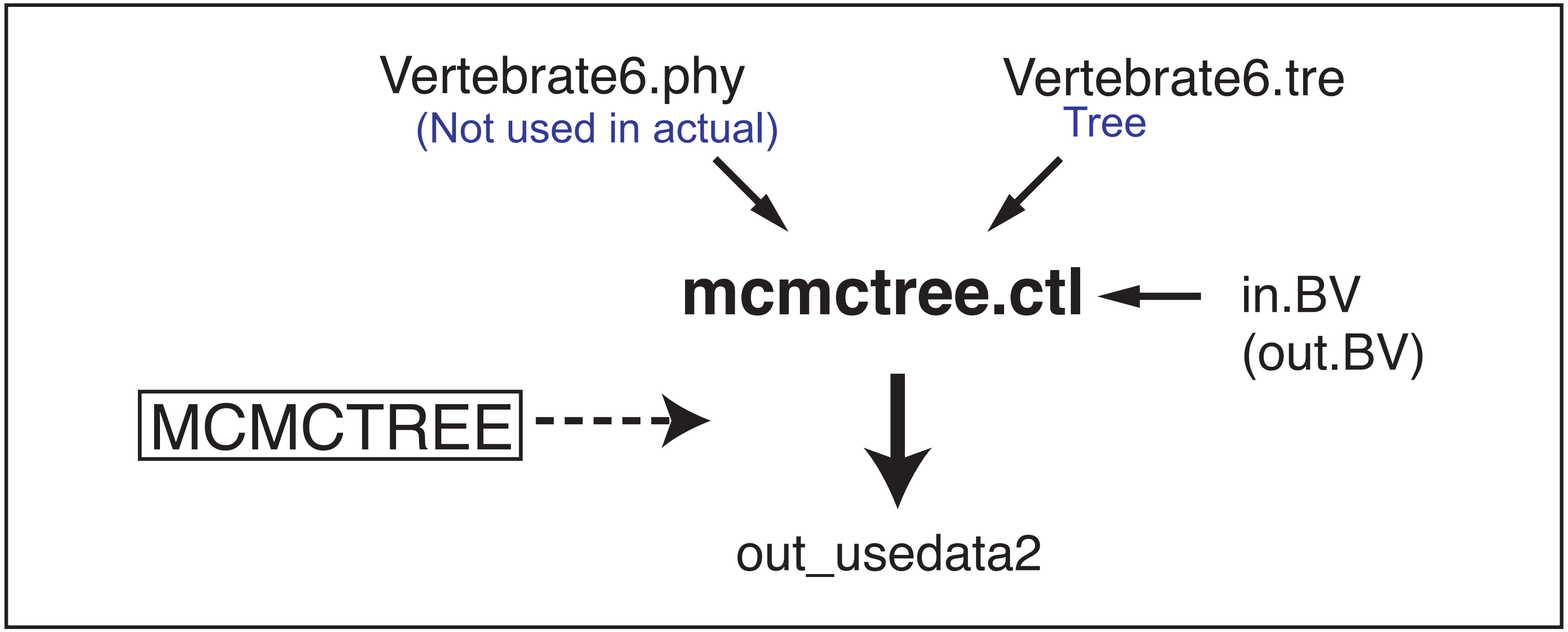
=> Copy to mcmctreeExampleVert6/examples

Workflow of MCMCTREE Analysis

1.
Estimation of Gradient and Hessian



2.
MCMC Analysis



Time Estimation using MCMCTREE

0. Preparation => Rough Estimation of Rate

1. Estimation of Gradient and Hessian
   (usedata=3)

2. MCMC Analysis
   (usedata=2)

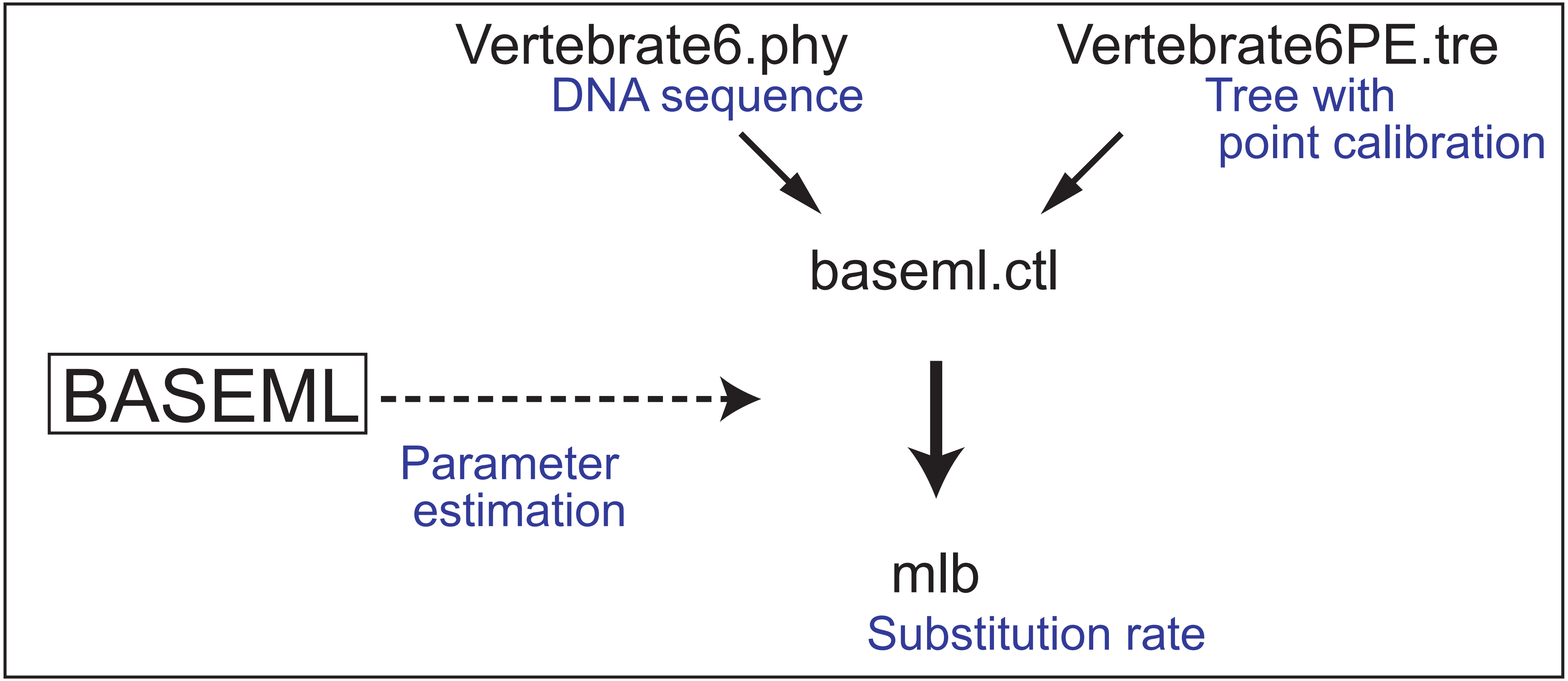
0: mcmctree.ctl File

```
          seed = -1
       seqfile = Vertebrate6.phy
      treefile = Vertebrate6.tre
       outfile = out_usedata3

         ndata = 1
       usedata = 3      * 0: no data; 1:seq like; 2:normal a
         clock = 2      * 1: global clock; 2: independent ra
       RootAge = '<10'  * constraint on root age, used if no

         model = 7    * 0:JC69, 1:K80, 2:F81, 3:F84, 4:HKY85
         alpha = 0.5  * alpha for gamma rates at sites
         ncatG = 5    * No. categories in discrete gamma

     cleandata = 0    * remove sites with ambiguity data (1

       BDparas = 1 1 0    * birth, death, sampling
   kappa_gamma = 6 2      * gamma prior for kappa
   alpha_gamma = 1 1      * gamma prior for alpha

   rgene_gamma = 1 12.5   * gamma prior for rate for genes
  sigma2_gamma = 1 4.5    * gamma prior for sigma^2     (fo

      finetune = 0: 0.06  0.5  0.008  0.12 0.4  * times, ra
*      finetune = 0.06  0.5  0.006  0.12 0.4  * times, rate

         print = 1
        burnin = 50000
      sampfreq = 50
       nsample = 10000

*** Note: Make your window wider (100 columns) when running
```

-Substitution rate (overall gene [partition]) prior probability → **rgene_gamma = 1 12.5**

0: Estimation of Substitution Rate (rgene_gamma)

Estimate overall rate using BASEML.



0: baseml.ctl File

- Sequence file
- Output file
- Tree file
- Substitution Model (7 is GTR)
- Strict clock model

```
      seqfile = Vertebrate6.phy * sequence data file name
      outfile = mlb          * main result file
     treefile = Vertebrate6PE.tre  * tree structure file 

        noisy = 3   * 0,1,2,3: how much rubbish on the sc
      verbose = 1   * 1: detailed output, 0: concise outp
      runmode = 0   * 0: user tree;  1: semi-automatic;  
                    * 3: StepwiseAddition; (4,5):Perturba

        model = 7   * 0:JC69, 1:K80, 2:F81, 3:F84, 4:HKY8
        Mgene = 0   * 0:rates, 1:separate; 2:diff pi, 3:d

    fix_kappa = 0
        kappa = 2   * initial or given kappa

    fix_alpha = 0
        alpha = 0.5  * initial or given alpha, 0:infinity
       Malpha = 0   * 1: different alpha's for genes, 0: 
        ncatG = 5   * # of categories in the dG, AdG, or 

      fix_rho = 1
          rho = 0.  * initial or given rho,   0:no correl
        nparK = 0   * rate-class models. 1:rK, 2:rK&fK, 3

        clock = 1   * 0: no clock, unrooted tree, 1: cloc
        nhomo = 1   * 0 & 1: homogeneous, 2: kappa's, 3: 
        getSE = 1   * 0: don't want them, 1: want S.E.s o
 RateAncestor = 0   * (1/0): rates (alpha>0) or ancestral
    cleandata = 0  * remove sites with ambiguity data (1:
```

0: Sequence File

Vertebrate6.phy Sequence data

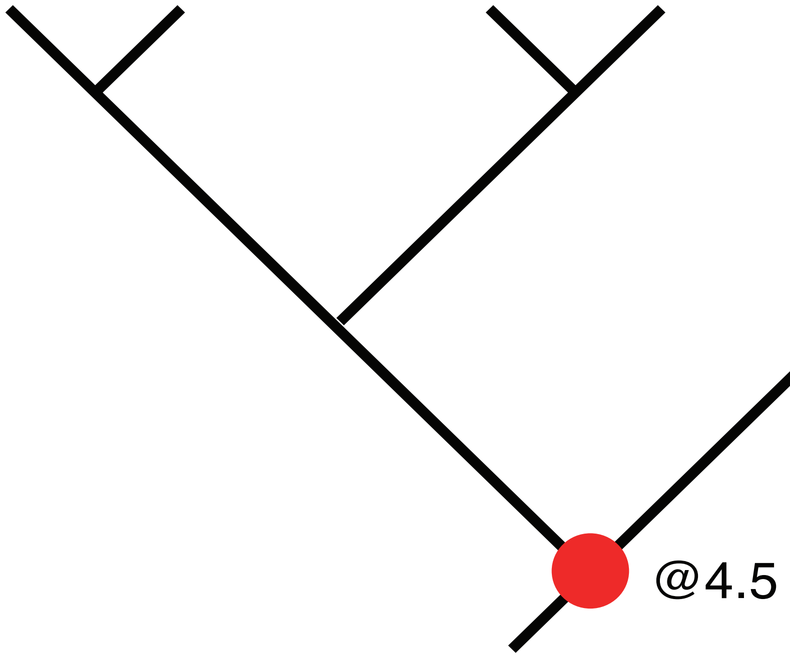
```
6 1146
Carp       ---ATGGCA---AGCCTACGAAAAACACACCCTCTCATTAAAATCGCTAACGACGCACTAGTTGA
Fugu       ---ATGGCC---AGCCTACGCAAAACCCACCCCCTACTAAAAATCGTAAACGACATAGTAATTGA
Frog       ---ATGGCACCCAACATCCGTAAATCTCATCCATTAATTAAAATTATTAATAATTCTTTCATTGA
Human      ------ATGACCCCAATACGCAAAACTAACCCCCTAATAAAATTAATTAACCACTCATTCATCGA
Shark      ---ATGGCCACAAACATCCGAAAAACTCACCCTCTACTAAAAATTGTAAATCATGCCCTAATTGA
Stingray   ATGACTACTACAAACACCCGCAAAACCCATCCCCTATTCAAAATTATCAATAACTCCGTAATTGA
```

Phylip format

0: Tree File

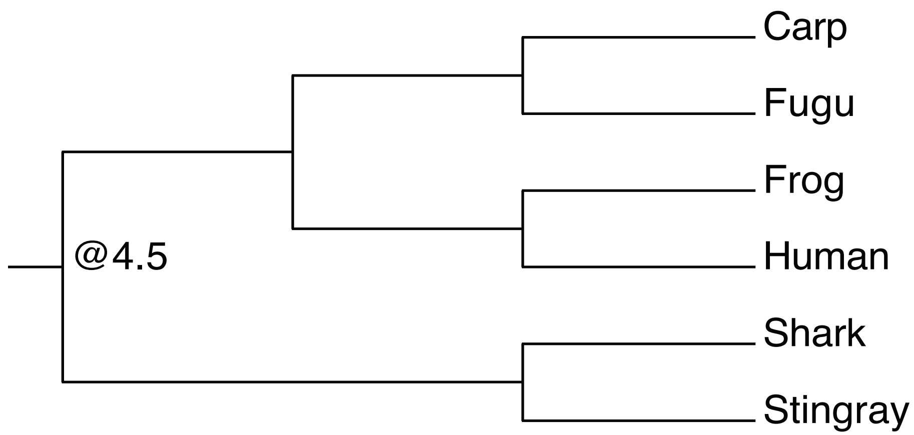
Vertebrate6PE.tre   Tree with point calibration

```
6  1
(((Carp,Fugu),(Frog,Human)),(Shark,Stingray))'@4.5';
```



Newick format

Open with FigTree



@4.5 means point constraint with 450 million years ago.
It is based on fossil literature.

Rough estimation of rate using point calibration 4.5 at root.
This will be used as a prior of MCMC analysis.

0: Run baseml

From command prompt (or terminal), type

./baseml

```
[JunInoue:12tr]$ baseml
BASEML in paml version 4.4d, March 2011

ns = 6     ls = 1146
Reading sequences, sequential format..
Reading seq # 1: Carp

  .............


 29 h-m-p  1.6000 8.0000   0.0000 C      5551.140599  0 1.4859   734 | 0/14
 30 h-m-p  1.6000 8.0000   0.0000 --------------C  5551.140599  0 0.0000   779
Out...
lnL  = -5551.140599
Calculating SE's

Time used:  0:00
[JunInoue:12tr]$
```

Outfile (mlb).

```
Substitution rate is per time unit
    0.084825 +- 0.008316
```

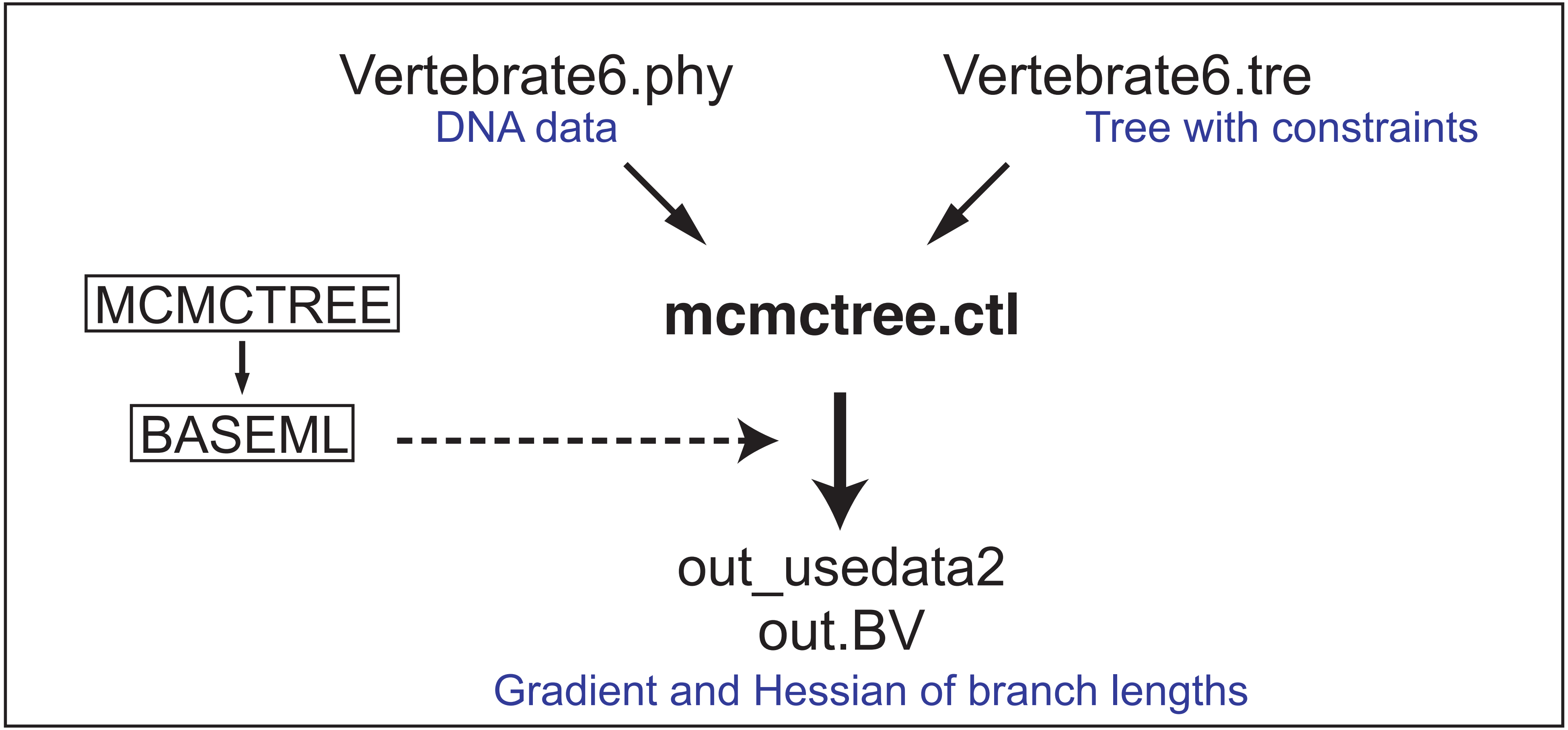
Time Estimation using MCMCTREE

0. Preparation => Rough Estimation of Rate

1. Estimation of Gradient and Hessian
   (usedata=3)

2. MCMC Analysis
   (usedata=2)

1. Estimation of Gradient and Hessian



Obtain maximum likelihood estimates (MLEs) of branch lengths together with the gradient and Hessian.

1: mcmctree.ctl File (usedata=3)

- Sequence file
- Tree file
- Output file
- Gradients and Hessian estimation mode
- Substitution model (7 is GTR)

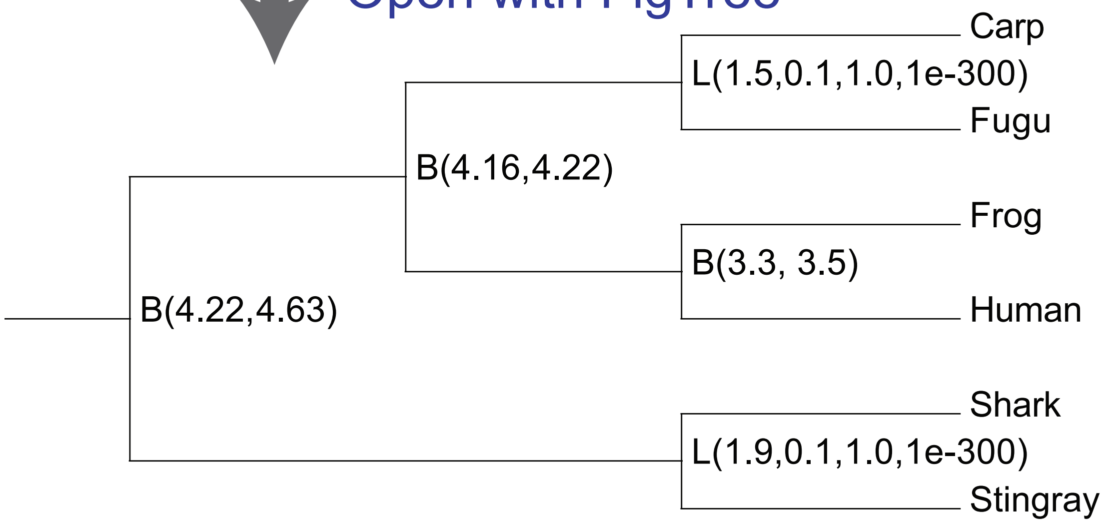
```
         seed = -1
      seqfile = Vertebrate6.phy
     treefile = Vertebrate6.tre
      outfile = out_usedata3

        ndata = 1
      usedata = 3      * 0: no data; 1:seq like; 2:normal a
        clock = 2      * 1: global clock; 2: independent ra
      RootAge = '<10'  * constraint on root age, used if no

        model = 7    * 0:JC69, 1:K80, 2:F81, 3:F84, 4:HKY8
        alpha = 0.5  * alpha for gamma rates at sites
        ncatG = 5    * No. categories in discrete gamma

    cleandata = 0    * remove sites with ambiguity data (1

      BDparas = 1 1 0    * birth, death, sampling
  kappa_gamma = 6 2      * gamma prior for kappa
  alpha_gamma = 1 1      * gamma prior for alpha

  rgene_gamma = 1 12.5   * gamma prior for rate for genes
 sigma2_gamma = 1 4.5    * gamma prior for sigma^2      (fo

     finetune = 0: 0.06  0.5  0.008  0.12 0.4  * times, ra
*     finetune = 0.06  0.5  0.006  0.12 0.4  * times, rate

        print = 1
       burnin = 50000
     sampfreq = 50
      nsample = 10000

*** Note: Make your window wider (100 columns) when running
```

1: tree File

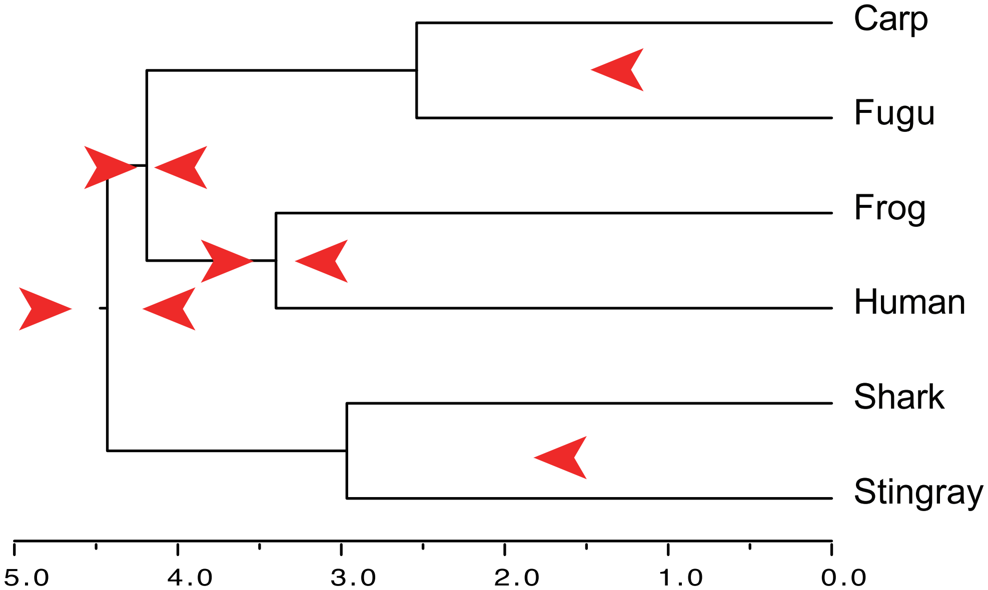
Vertebrate6.tre   Tree with constraints

```
6 1
(((Carp,Fugu)'L(1.5,0.1,1.0,1e-300)',(Frog,Human)'B(3.3, 3.5)')'B(4.16,
4.22)',(Shark,Stingray)'L(1.9,0.1,1.0,1e-300)')'B(4.22, 4.63)';
```

Open with FigTree



: Constraints based on literature



B(4.22, 4.63)

Paired constraint (Maximum 4.63, Minimum 4.22)

L(1.5, 0.1, 1.0, 1e-300)

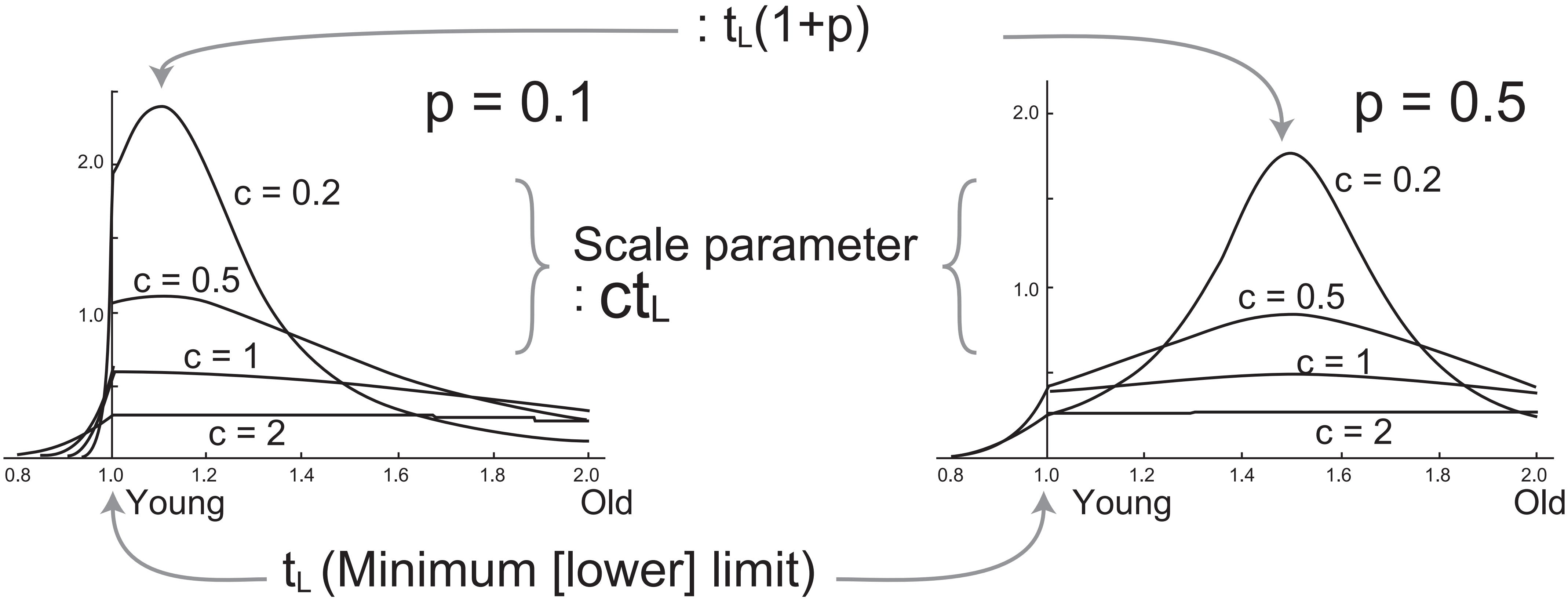
Minimum constraint (1.5, p=0.1, c=1.0, hard bound)

Caution：4.63 means 463 Mya. MCMCTREE uses time values 1-10 usually.

Prior Distribution for Minimum Constraint

MCMCTREE (PAML):

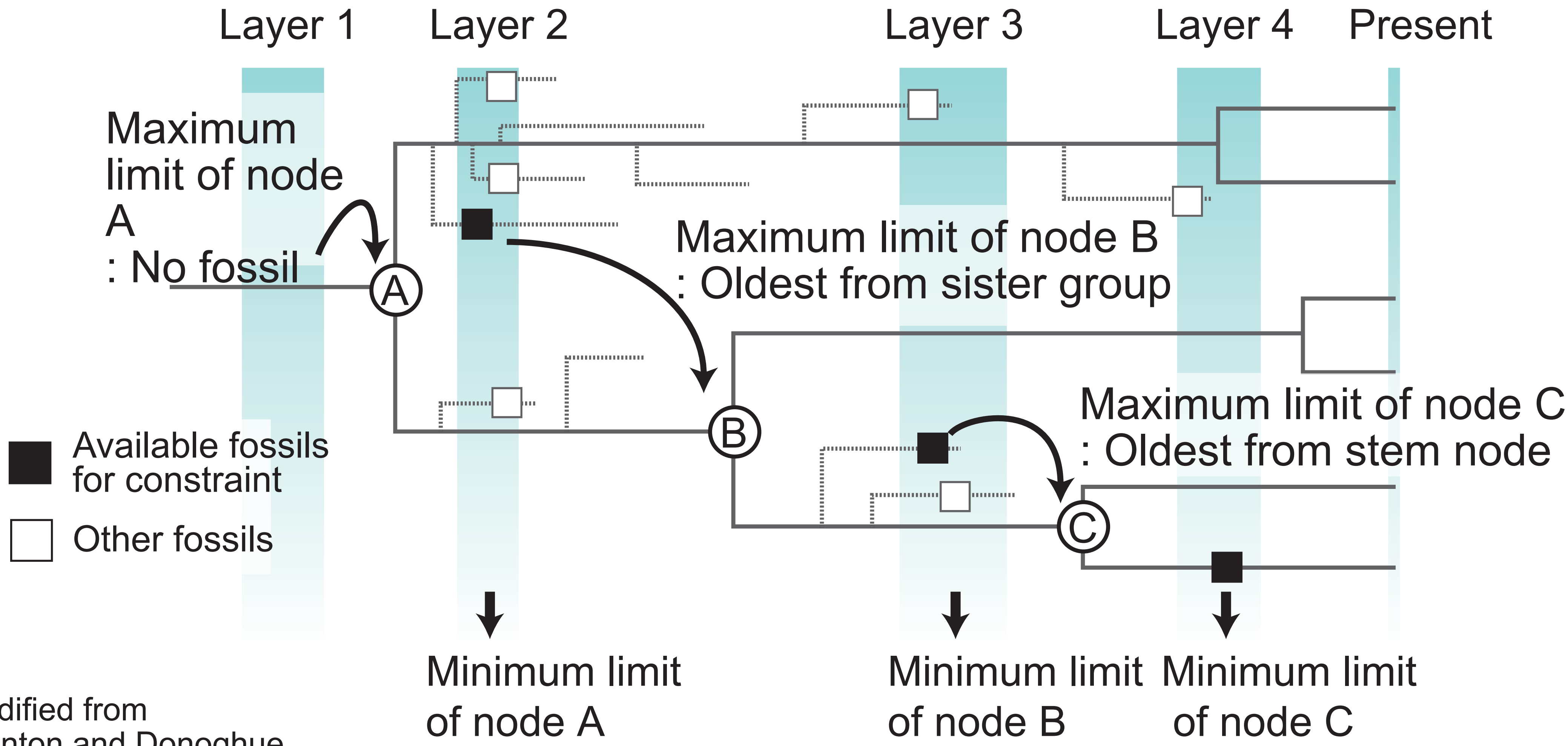
Cauchy distribution: ‘$L(t_L, p, c)$’



Inoue et al. (2010)

Fossil Selection for Constraints

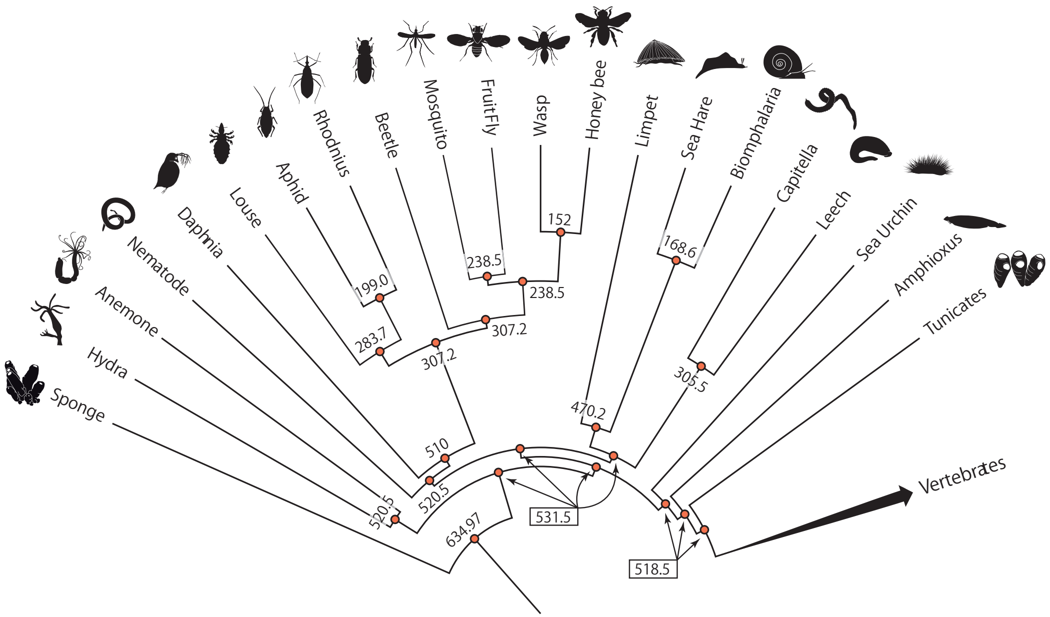
Requirement: Phylogenetic position of fossil

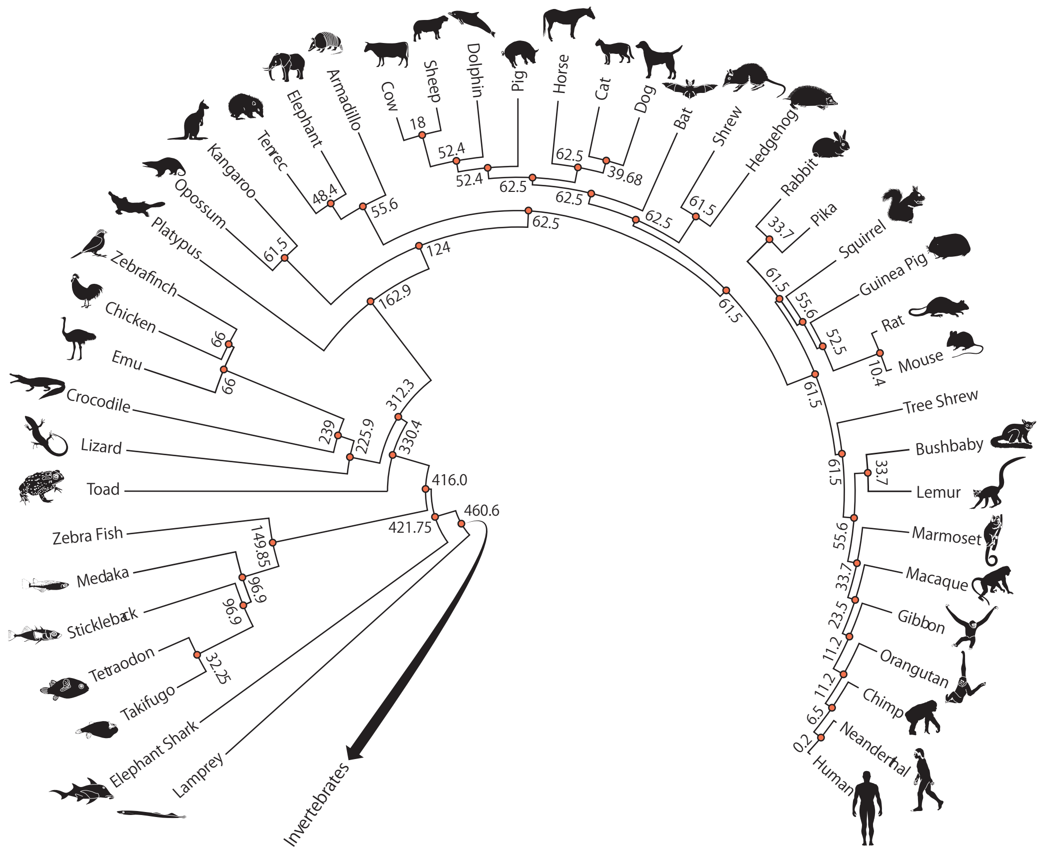


Modified from Benton and Donoghue (2007)

- Minimum is easy.
- Maximum is difficult. We need further study.

Fossils for Time Estimation





Benton et al. (2009)

Benton, M. J., P. C. J. Donoghue, and R. J. Asher. 2009. Calibrating and constraining molecular clocks. Pages 35-86 in Timetree of Life (S. B. Hedges, and S. Kumar, eds.). Oxford University Press.

- Some reliable constraints are suggested in Benton et al. (2009)
- For other nodes, see Fossil Records II (Benton 1993).

http://www.fossilrecord.net/fossilrecord/index.html

1: Run MCMCTREE (usedata=3)

From command prompt, type,

./mcmctree

```
[JunInoue:12tr]$ mcmctree
MCMCTREE in paml version 4.4c, August 2010

Reading options from mcmctree.ctl..
Reading master tree.
(((Carp, Fugu), (Frog, Human)), (Shark, Stingray));

  ..........

 25 h-m-p  1.6000 8.0000   0.0000 C      5540.322816  0 1.5426   623 | 0/15
 26 h-m-p  1.6000 8.0000   0.0000 C      5540.322816  0 1.6000   656 | 0/15
 27 h-m-p  1.6000 8.0000   0.0000 C      5540.322816  0 2.4792   689
Out...
lnL  = -5540.322816
Calculating SE's

Time used:  0:00
[JunInoue:12tr]$
```

Branch lengths, gradient, and Hessian will be saved in out.BV file.

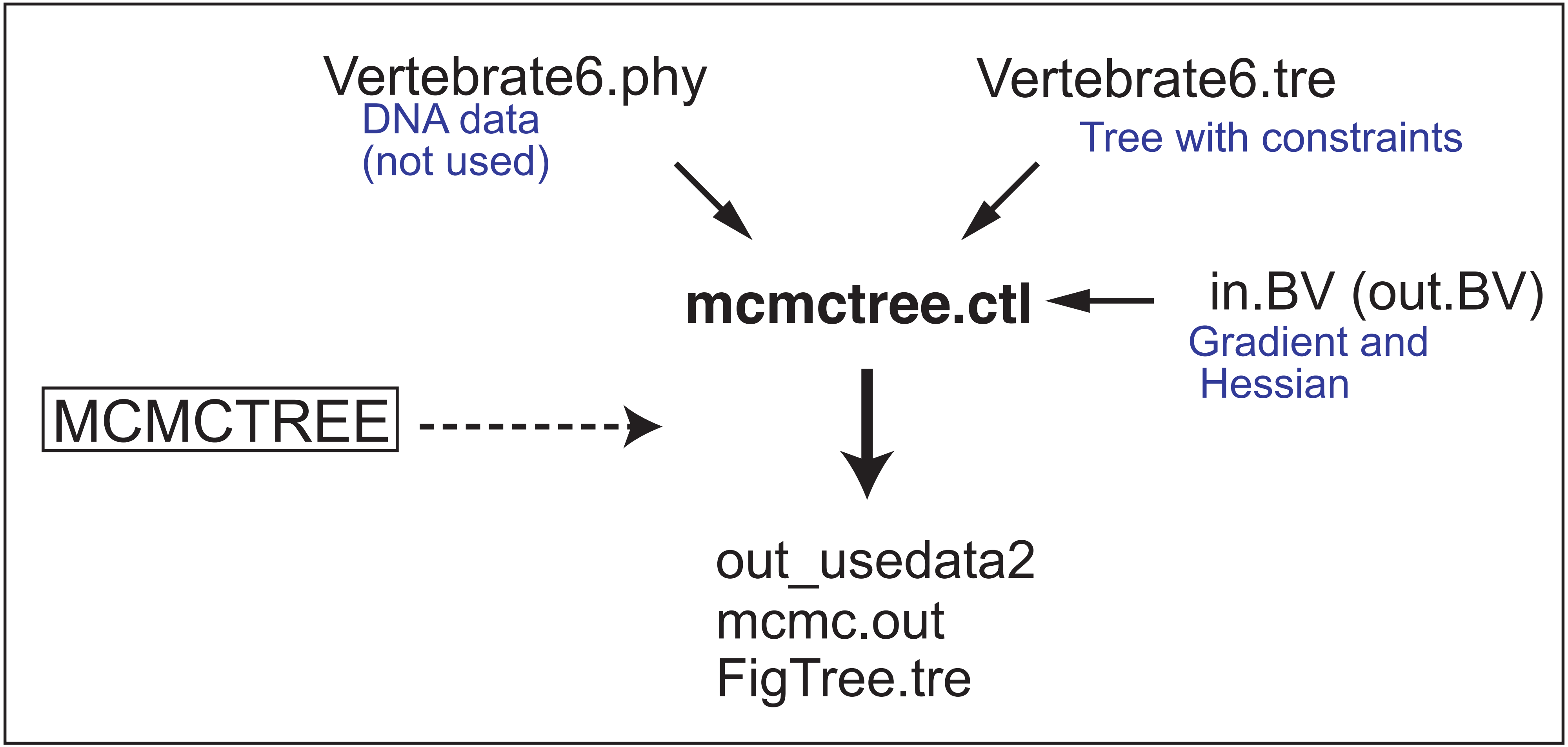
Time Estimation using MCMCTREE

0. Preparation => Rough Estimation of Rate

1. Estimation of Gradient and Hessian
   (usedata=3)

2. MCMC Analysis
   (usedata=2)

2. MCMC Analysis



Bayesian estimation of rate and time.

2: in.BV File (out.BV <= change name)

in.BV file (out.BV : obtained from usedata = 3)

Number of species
ML tree with branch lengths
Vector of branch lengths
Gradient

Vector of first derivative of the log-likelihood function

Hessian matrix

(Variance-covariance matrix of branch lengths)

Matrix of second derivative of the log-likelihood function

```
 6

((Carp: 0.178357, Fugu: 0.230865): 0.059097, (Frog: 0.300002, Human: 0.480660): 0.045538, (

  0.059097  0.178357  0.230865  0.045538  0.300002  0.480660  0.123633  0.341101  0.273660

  0.000000  0.000000  0.000000  0.000000  0.000000  0.000000  0.000000  0.000000  0.000000


Hessian

  -1.9e+03  -1.3e+02        22  -3.2e+02  -1.1e+02        13  -2.7e+02        55        45
  -1.3e+02  -1.4e+03  -1.8e+02        54        73   1.1e+02        32        24        43
        22  -1.8e+02  -1.1e+03   1.1e+02        99   1.1e+02        93        78       8.9
  -3.2e+02        54   1.1e+02  -1.8e+03  -3.3e+02       -51       -21        98        67
  -1.1e+02        73        99  -3.3e+02  -6.7e+02        72        53        48        81
        13   1.1e+02   1.1e+02       -51        72  -3.7e+02        75        70        59
  -2.7e+02        32        93       -21        53        75  -1.1e+03       -66  -1.2e+02
        55        24        78        98        48        70       -66  -5.2e+02        16
        45        43       8.9        67        81        59  -1.2e+02        16  -6.7e+02
```

The gradient and Hessian describe the shape of the log-likelihood surface around the MLEs.

↓

Approximate the log-likelihood curve for each MCMC generation.
Caution: Sequence data is not used.

Thorne, 1998, p1650
Yang 2006 p22, MCT manual.

2: mcmctree.ctl File (usedata=2)

- MCMC analysis
- Prior of Substitution rate
- Parameters specify how variable the rates are across branches
- finetune (1 is automatic) Step lengths for the proposals
- Iterations of MCMC – Conduct at least twice to ensure consistency.

```
         seed = -1
      seqfile = Vertebrate6.phy
     treefile = Vertebrate6.tre
      outfile = out_usedata2

        ndata = 1
     usedata = 2      * 0: no data; 1:seq like; 2:normal a
        clock = 2      * 1: global clock; 2: independent ra
      RootAge = '<10'  * constraint on root age, used if no

        model = 7    * 0:JC69, 1:K80, 2:F81, 3:F84, 4:HKY85
        alpha = 0.5  * alpha for gamma rates at sites
        ncatG = 5    * No. categories in discrete gamma

    cleandata = 0    * remove sites with ambiguity data (1

      BDparas = 1 1 0    * birth, death, sampling
  kappa_gamma = 6 2      * gamma prior for kappa
  alpha_gamma = 1 1      * gamma prior for alpha

  rgene_gamma = 1 12.5   * gamma prior for rate for genes
 sigma2_gamma = 1 4.5    * gamma prior for sigma^2     (for

     finetune = 1: 0.06  0.5  0.008  0.12 0.4  * times, ra
*     finetune = 0.06  0.5  0.006  0.12 0.4  * times, rate

        print = 1
       burnin = 50000
     sampfreq = 50
      nsample = 10000

*** Note: Make your window wider (100 columns) when running
```

rgene_gamma

-Specifies the distribution of overall rate parameter μ.

Mean (m), standard deviation (s),
shape parameter (α), scale parameter (β) :

$$\alpha = (m/s)^2$$
$$\beta = m/s^2.$$

Substitution rate is per time unit (from baseml output file)

```
0.084825 +- 0.008316
```

from mlb file, we will set

$$m = 0.08, s=0.08,$$

then $\alpha = (0.08/0.08)^2$
$= 1$
$\beta = 0.08/0.08^2$
$=12.5$

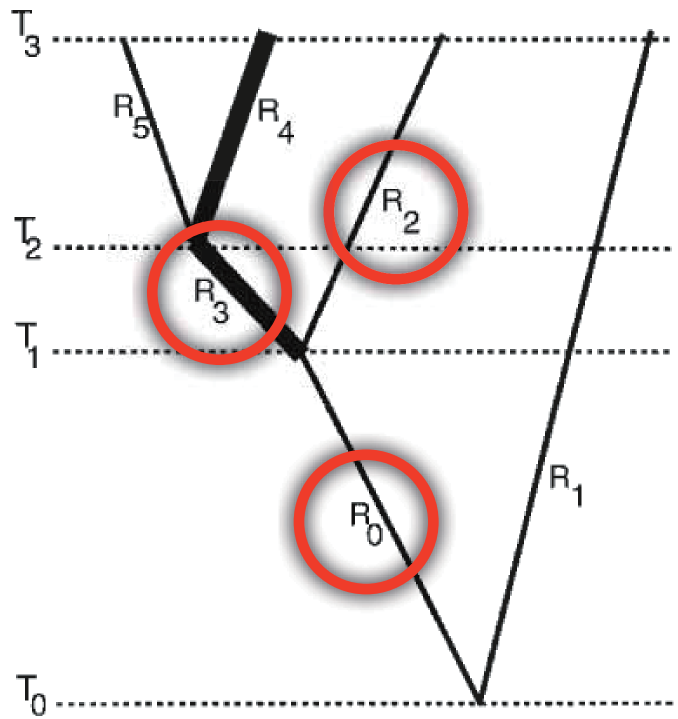
sigma2_gamma

- Specifies how variable the rates are across branches.

=> MULTIDIVTIME manual

```
rttm* m = 1
```

If we set rttm = 4.5,
$m = 1/4.5$
When $m = s = 1/4.5$,
$\alpha = (m/s)^2$
$= 1$
$\beta = m/s^2$
$= 4.5$



MCMCTREE manual

2: Run MCMCTREE (usedata=2)

```
[JunInoue:12tr]$ mcmctree
MCMCTREE in paml version 4.4c, August 2010

Reading options from mcmctree.ctl..
Reading master tree.
(((Carp, Fugu), (Frog, Human)), (Shark, Stingray));

  ...........

r_n11       0.0757  0.0802 (0.0392, 0.1469)
lnL        -4.3260 -4.6698 (-9.7050, -1.3740)

time prior: Birth-Death-Sampling
rate prior: Log-Normal

Time used:  0:14
[JunInoue:12tr]$
```

out_usedata2
mcmc.out
FigTree.tre

will be obtained.

2: finetune (finetune = 0, manual control)

Acceptance proportion (screen out) should be ranged 0.2– 0.4.

```
[JunInoue:12tr]$ mcmctree
MCMCTREE in paml version 4.4c, August 2010
  ......

Starting MCMC (np = 17) . . .
finetune steps (time rate mixing para RatePara?):  0.0600  0.5000  0
  paras: 5 times, 1 mu, 1 sigma2 (& rates, kappa, alpha)
-10% 0.46 0.28 0.46 0.00 0.33  4.589 4.254 2.346 3.359 2.866 0.094 -
 -5% 0.40 0.22 0.54 0.00 0.40  4.444 4.189 2.898 3.410 3.287 0.085 -
  0% 0.42 0.19 0.60 0.00 0.40  4.410 4.183 3.127 3.414 3.053 0.089 -
  5% 0.40 0.22 0.54 0.00 0.41  4.439 4.189 2.909 3.411 3.242 0.085 -
 10% 0.40 0.22 0.54 0.00 0.41  4.437 4.189 2.911 3.411 3.232 0.086 -
 15% 0.40 0.22 0.54 0.00 0.41  4.438 4.189 2.920 3.411 3.235 0.086 -
 20% 0.40 0.22 0.54 0.00 0.40  4.439 4.189 2.931 3.411 3.243 0.086 -
 25% 0.40 0.22 0.54 0.00 0.40  4.439 4.189 2.939 3.411 3.239 0.085 -
^C
[JunInoue:12tr]$
```

Too big,

try 0.01.

```
 finetune = 0.06  0.5  0.006  0.12 0.4  * times, rates, mixing, para
```

2: mcmc.out File (MCMC samplings)

mcmc.out file

t — $\sigma 2$ — Log likelihood

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Gen	t_n24	t_n25		mu1	mu2	sigma2_1	sigma2_2	r_g1_n1	r_g1_n2		r_g2_n1	r_g2_n2		lnL
2	1	4.6042825	4.1914538		0.0598536	0.0128381	0.4430608	0.524874	0.0378056	0.0794104		0.0108554	0.0153361		−107.919
3	50	4.5909781	4.175952		0.057184	0.0169996	0.5520485	0.5658893	0.0343797	0.0641597		0.0145256	0.0127715		−99.092
4	100	4.5861	4.1740957		0.0527724	0.0177151	0.2069208	0.4219372	0.034395	0.0641882		0.0074754	0.0116595		−89.257
5	150	4.6090388	4.1697957		0.0658565	0.0158464	0.3986258	0.3627929	0.0278656	0.0440668		0.0071804	0.0088083		−77.982
6	200	4.5593259	4.1960602		0.0583959	0.0176448	0.1631693	0.3153585	0.0341905	0.0440716		0.0112959	0.0116771		−93.709
7	250	4.5773135	4.1732455		0.0585138	0.0152161	0.2247225	0.2818855	0.0309738	0.0475023		0.0070665	0.0117115		−85.539

Iterations — Overall rate μ — rate of each branch r

- Start from random values of Time t, rate r, parameters θ.
- Propose a new state with, t^*, r^*, θ^*
- Sample the chain for every k iterations : save t, r, θ to mcmc.out.
- Calculate mean and sd in the end.

Here, sampling for every 50 iterations.

mcmctree.ctl file

```
    print = 1
   burnin = 50000
 sampfreq = 50
  nsample = 10000
```

$$R = \min\left\{1, \frac{f(X|\mathbf{t}^*, \mathbf{r}^*) f(\mathbf{r}^*|\mathbf{t}^*, \theta^*) f(\mathbf{t}^*|\theta^*) f(\theta^*)}{f(X|\mathbf{t}, \mathbf{r}) f(\mathbf{r}|\mathbf{t}, \theta) f(\mathbf{t}|\theta) f(\theta)} \times \frac{q(\theta, \mathbf{t}, \mathbf{r}|\theta^*, \mathbf{t}^*, \mathbf{r}^*)}{q(\theta^*, \mathbf{t}^*, \mathbf{r}^*|\theta, \mathbf{t}, \mathbf{r})}\right\}$$

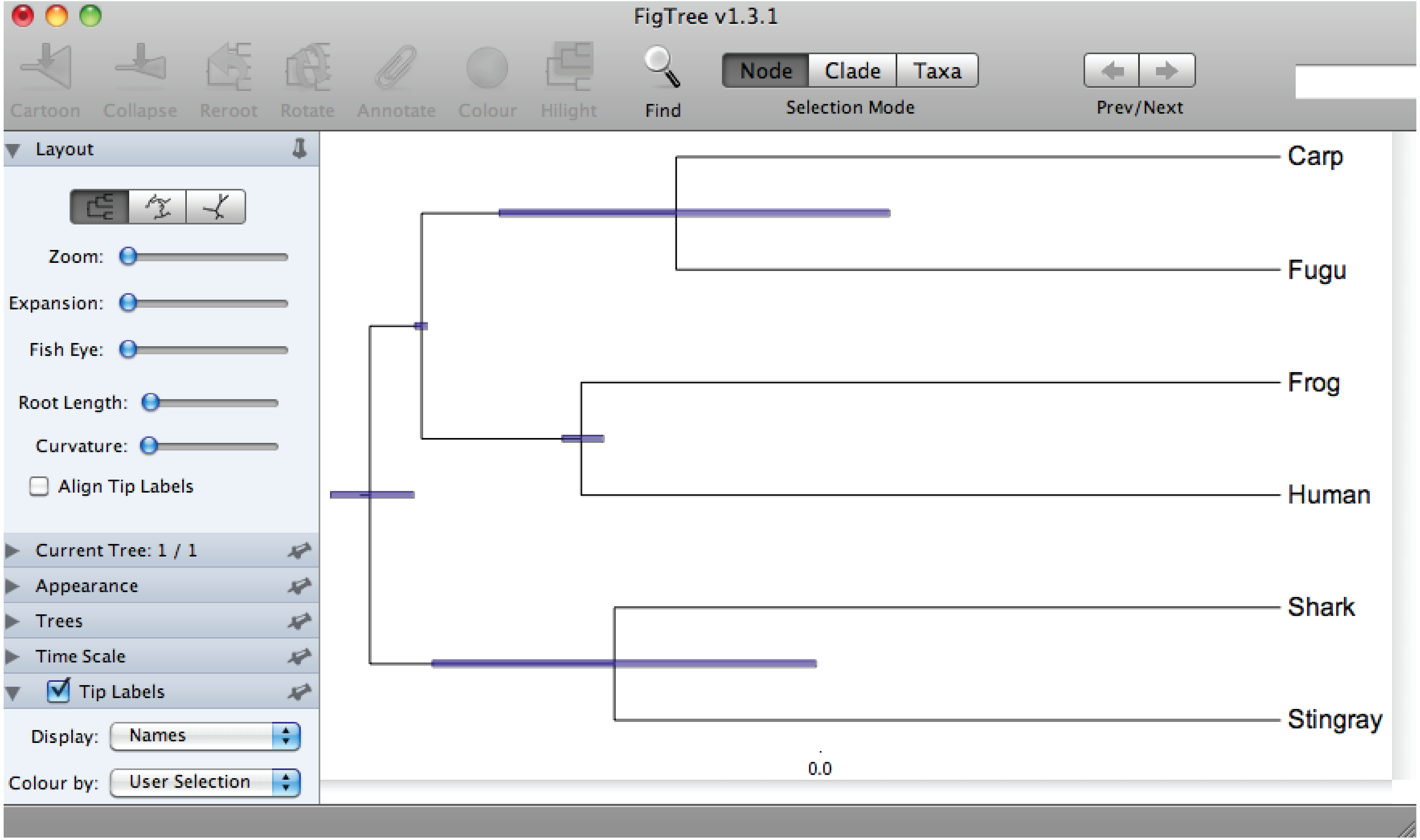
Yang (2006)

2: Output File (out_usedata2 file)

```
Species tree for TreeView.  Branch lengths = posterior mean times
(((1_Carp, 2_Fugu) 9 , (3_Frog, 4_Human) 10 ) 8 , (5_Shark, 6_Sti

(((Carp: 2.947415, Fugu: 2.947415): 1.241267, (Frog: 3.410426, Hu
0.252082, (Shark: 3.247640, Stingray: 3.247640): 1.193124);

(((Carp: 2.947415, Fugu: 2.947415) '1.906-3.811': 1.241267, (Frog
'3.302-3.501': 0.778256) '4.160-4.220': 0.252082, (Shark: 3.24764
'2.263-4.135': 1.193124) '4.226-4.632';

rategram locus 1:
(((Carp: 0.062332, Fugu: 0.076731): 0.062943, (Frog: 0.081017, Hu
0.084457, (Shark: 0.099224, Stingray: 0.082360): 0.080458);


Posterior median mean & 95% credibility interval

t_n7        4.4500  4.4419 (4.2264, 4.6320)  (age of jeffnode 10)
t_n8        4.1877  4.1885 (4.1598, 4.2198)  (age of jeffnode  9)
t_n9        2.9756  2.9477 (1.9057, 3.8108)  (age of jeffnode  8)
t_n10       3.4162  3.4107 (3.3021, 3.5008)  (age of jeffnode  7)
t_n11       3.2640  3.2474 (2.2630, 4.1346)  (age of jeffnode  6)
mu          0.0828  0.0843 (0.0631, 0.1141)
sigma2      0.1145  0.1434 (0.0269, 0.4243)
r_n1        0.0606  0.0623 (0.0416, 0.0940)
r_n2        0.0743  0.0767 (0.0513, 0.1179)
```

- Mean and 95 CI of time (t_n7 – t_n9)

2: FigTree.tre File



Tree drawing by FigTree
http://tree.bio.ed.ac.uk/software/figtree/

Reference

岸野洋久，Jeffrey L. Thorne．2002．分子進化速度のベイズ型階層モデル．統計数理．50: 17-31．http://www.ism.ac.jp/editsec/toukei/pdf/50-1-017.pdf

Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Molecular Biology and Evolution 15:1647–1657.

Yang, Z. http://abacus.gene.ucl.ac.uk/software/paml.html

Yang Z. 2006. Computational Molecular Evolution. Oxford Univ. Press.