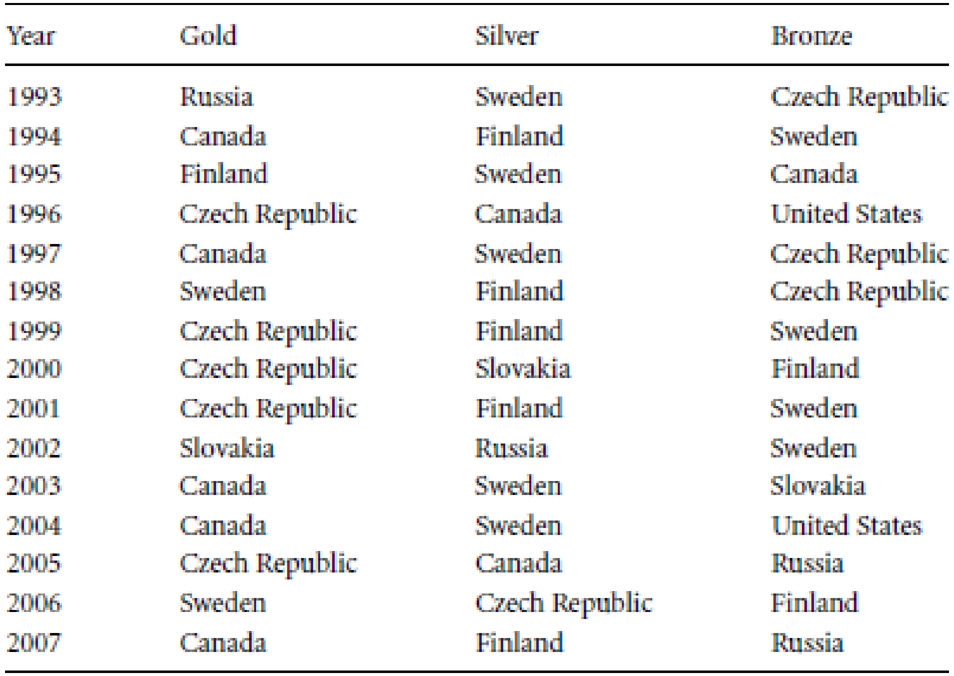
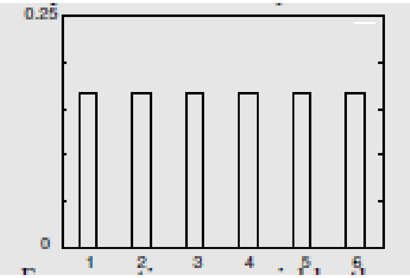
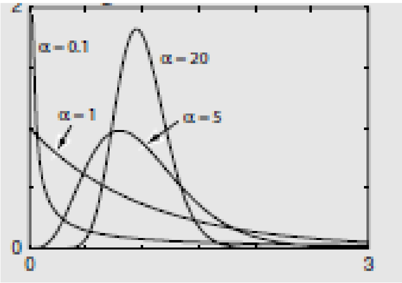
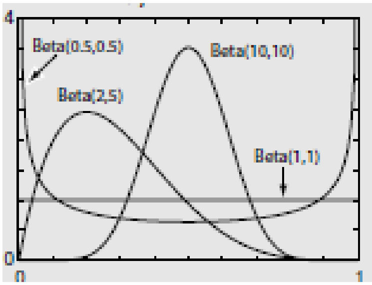
7. Bayesian filogenetik analiz

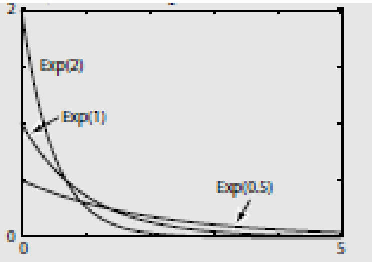


Year	Gold	Silver	Bronze
1993	Russia	Sweden	Czech Republic
1994	Canada	Finland	Sweden
1995	Finland	Sweden	Canada
1996	Czech Republic	Canada	United States
1997	Canada	Sweden	Czech Republic
1998	Sweden	Finland	Czech Republic
1999	Czech Republic	Finland	Sweden
2000	Czech Republic	Slovakia	Finland
2001	Czech Republic	Finland	Sweden
2002	Slovakia	Russia	Sweden
2003	Canada	Sweden	Slovakia
2004	Canada	Sweden	United States
2005	Czech Republic	Canada	Russia
2006	Sweden	Czech Republic	Finland
2007	Canada	Finland	Russia

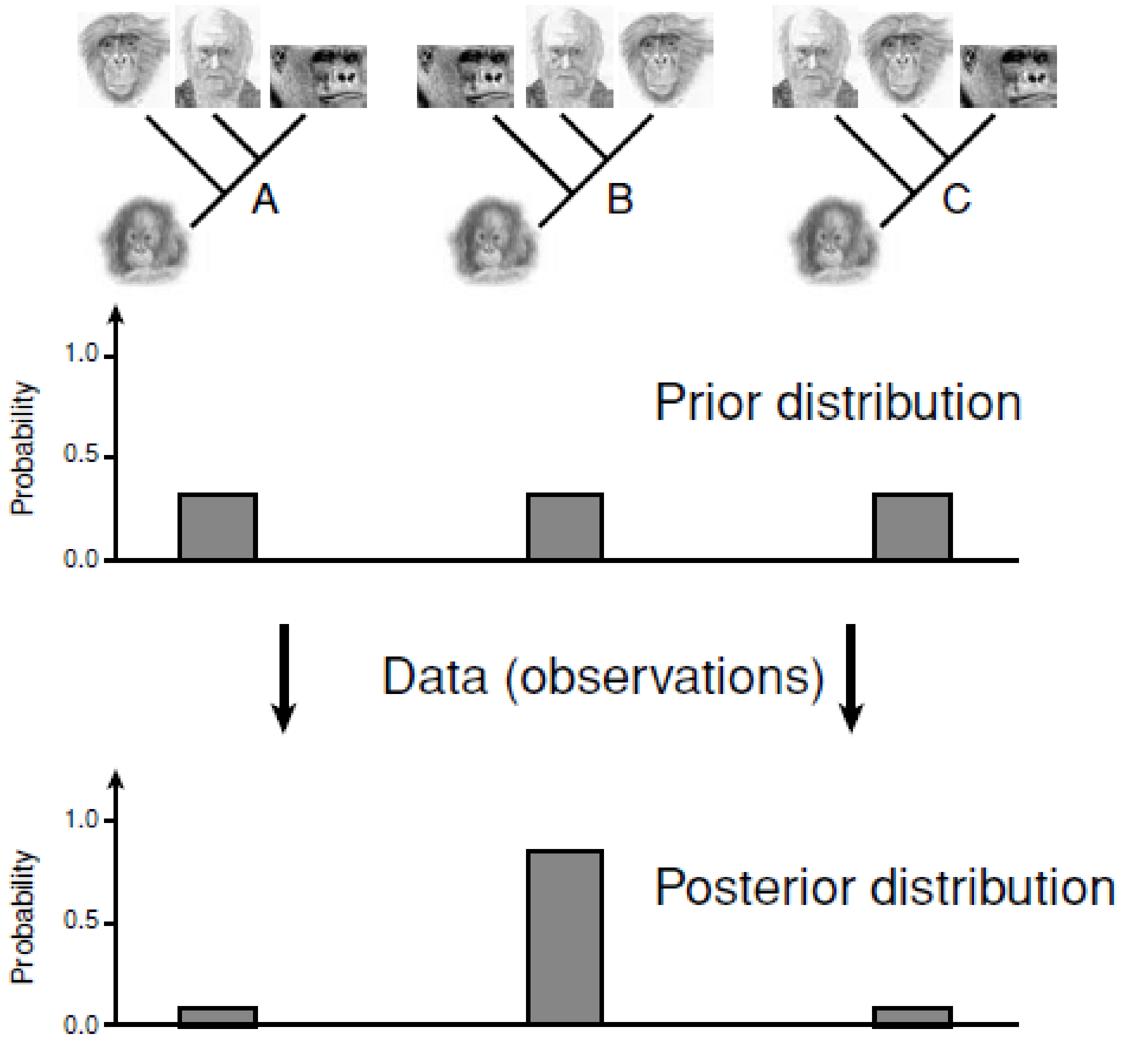


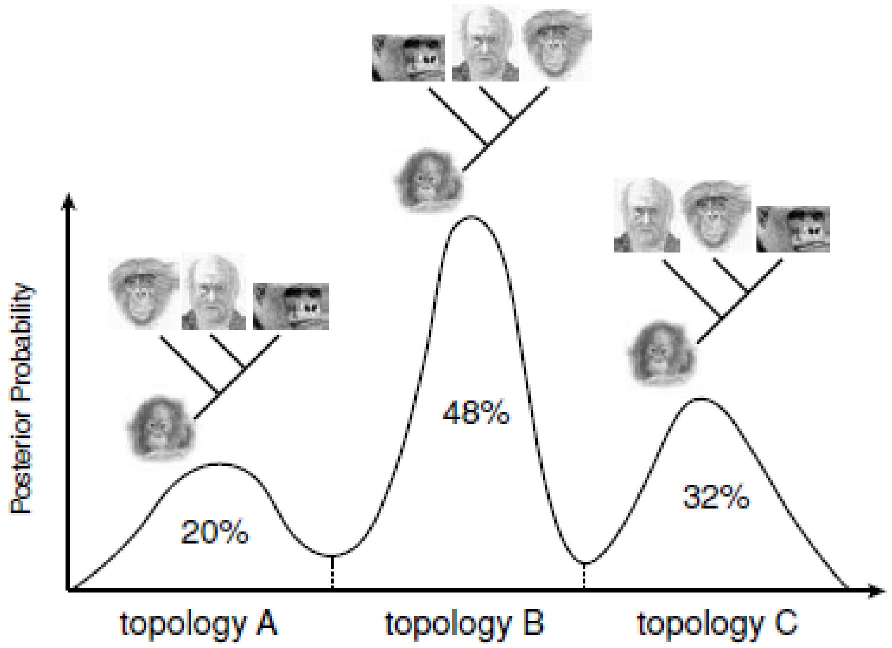




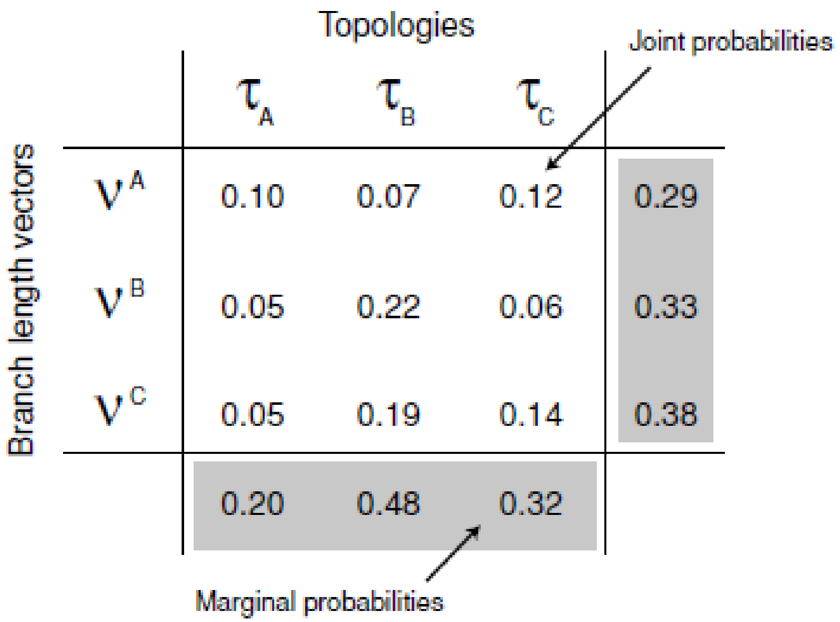


Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.





Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

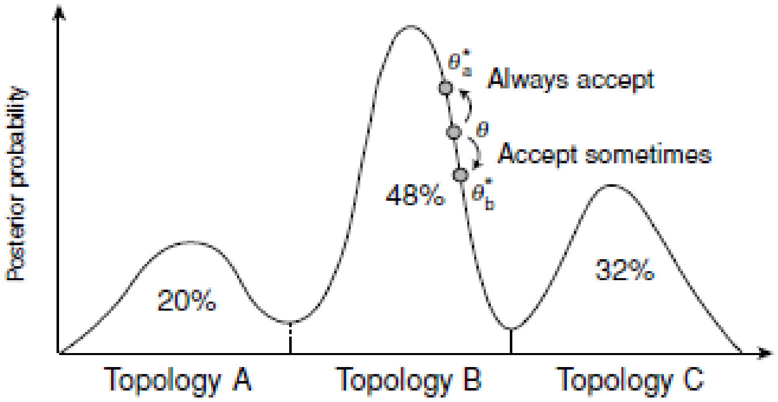


Topologies

Branch length vectors	τ_A	τ_B	τ_C	
ν^A	0.10	0.07	0.12	0.29
ν^B	0.05	0.22	0.06	0.33
ν^C	0.05	0.19	0.14	0.38
	0.20	0.48	0.32	

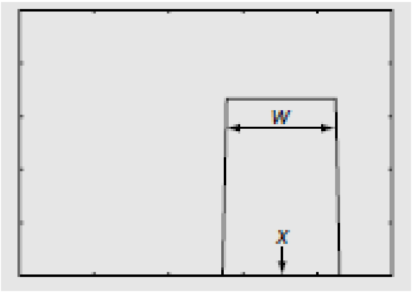
Joint probabilities

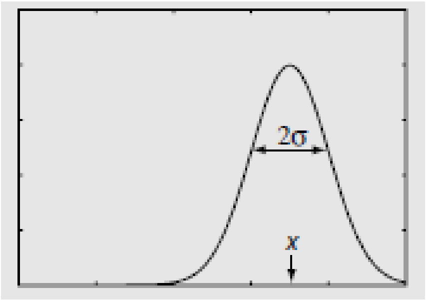
Marginal probabilities

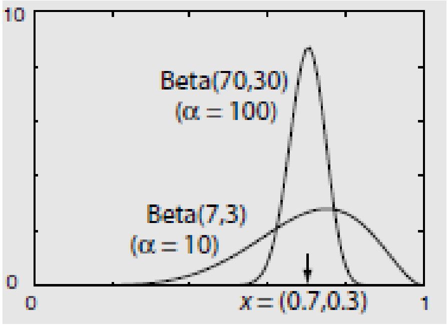


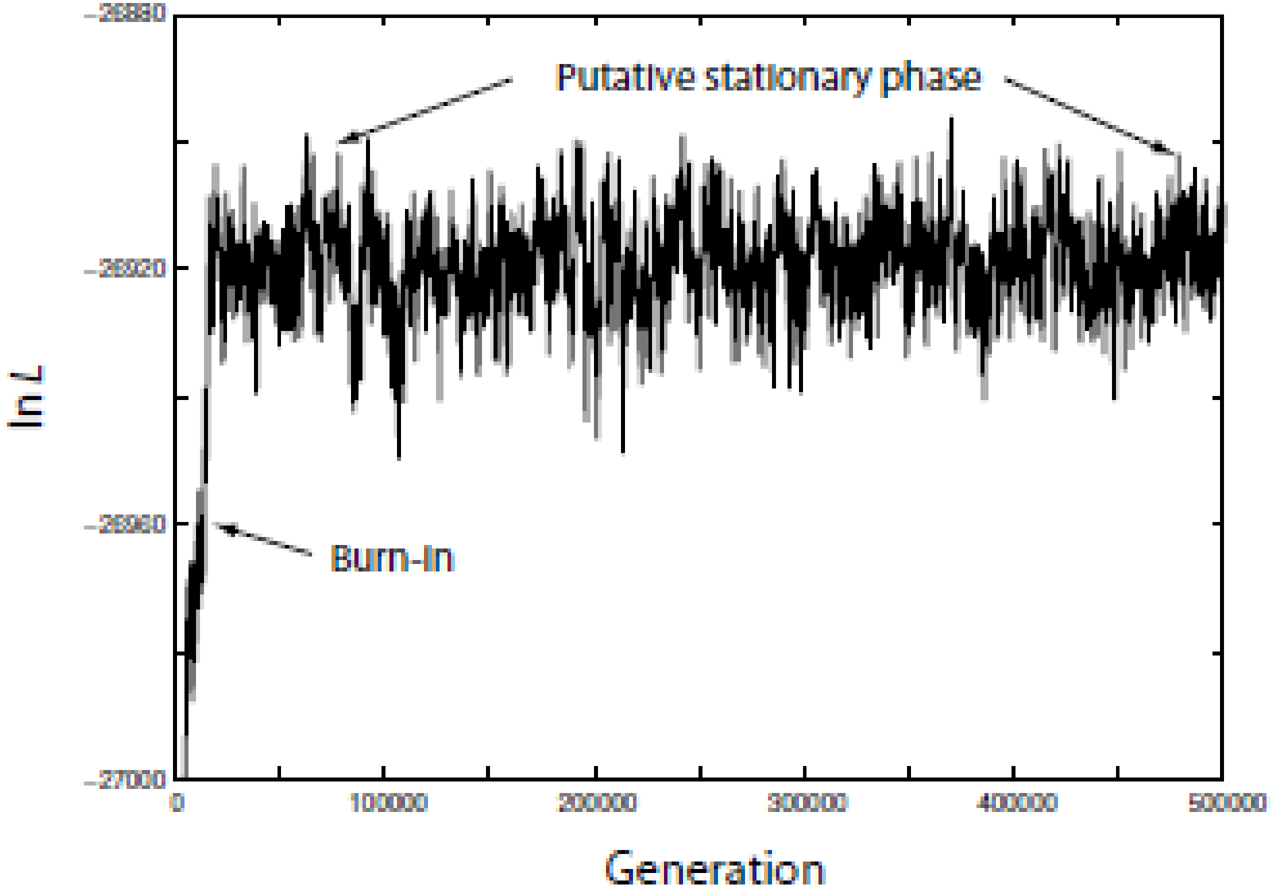
$$
\begin{aligned}
r &= \min\left(1, \frac{f(\theta^*|X)}{f(\theta|X)} \times \frac{f(\theta|\theta^*)}{f(\theta^*|\theta)}\right) \\
&= \min\left(1, \frac{f(\theta^*)\, f(X|\theta^*)/f(X)}{f(\theta)\, f(X|\theta)/f(X)} \times \frac{f(\theta|\theta^*)}{f(\theta^*|\theta)}\right) \\
&= \min\left(1, \frac{f(\theta^*}{f(\theta)} \times \frac{f(X|\theta^*)}{f(X|\theta)} \times \frac{f(\theta|\theta^*)}{f(\theta^*|\theta)}\right)
\end{aligned}
$$

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

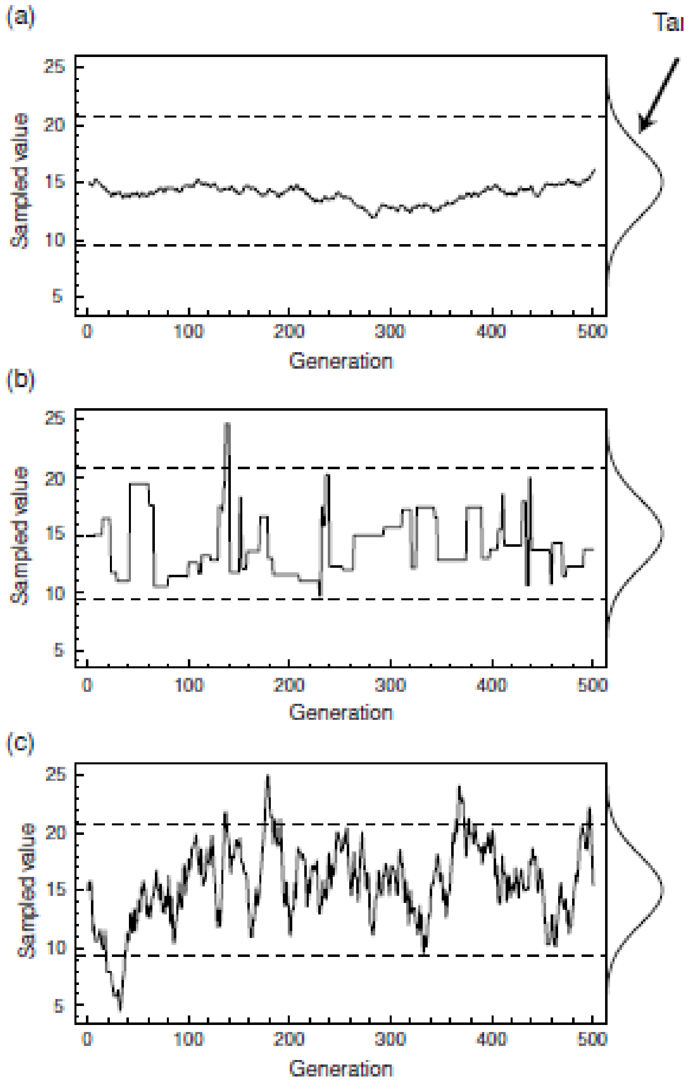


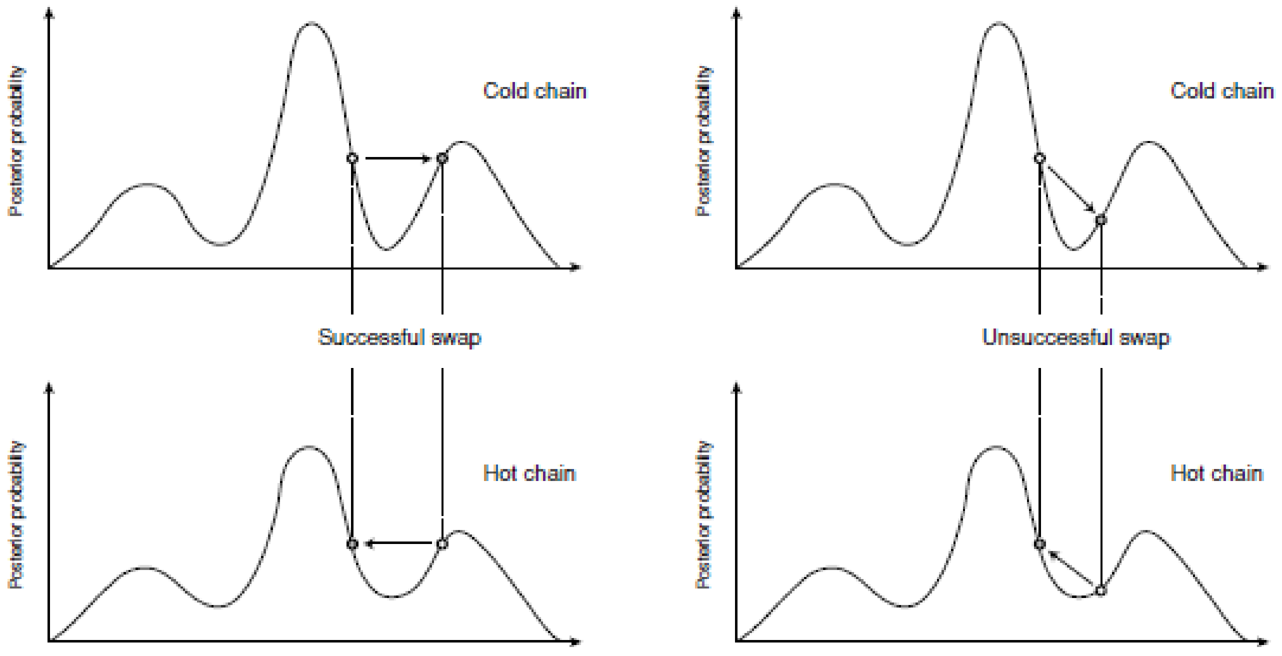






Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

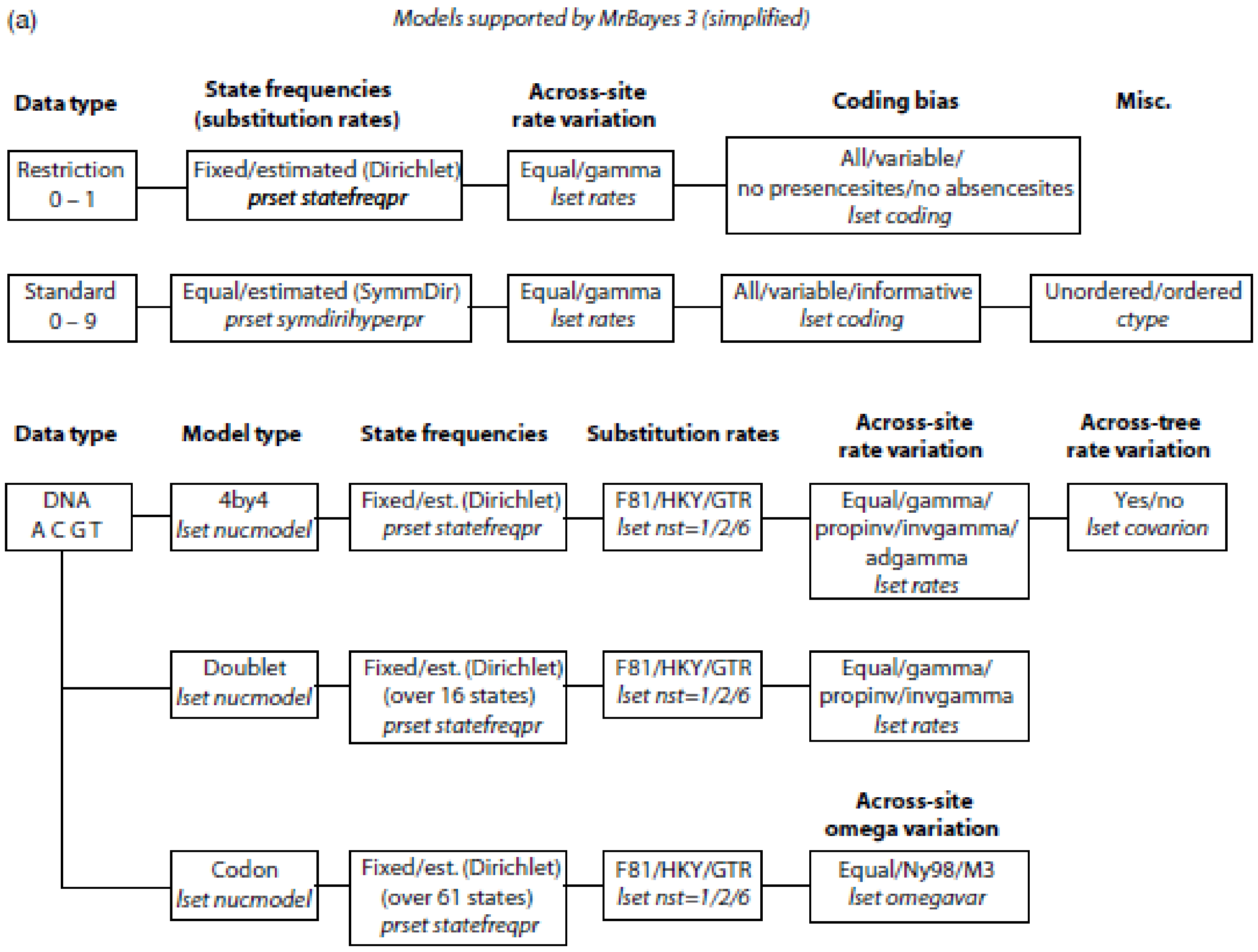




Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

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

$$\frac{f(M_0|X)}{f(M_1|X)} = \frac{f(M_0)\,f(X|M_0)}{f(M_1)\,f(X|M_1)} = \frac{f(M_0)}{f(M_1)} \times \frac{f(X|M_0)}{f(X|M_1)}$$



(a)

Models supported by MrBayes 3 (simplified)

Data type | **State frequencies (substitution rates)** | **Across-site rate variation** | **Coding bias** | **Misc.**

Restriction 0 – 1 — Fixed/estimated (Dirichlet) *prset statefreqpr* — Equal/gamma *lset rates* — All/variable/ no presencesites/no absencesites *lset coding*

Standard 0 – 9 — Equal/estimated (SymmDir) *prset symdirihyperpr* — Equal/gamma *lset rates* — All/variable/informative *lset coding* — Unordered/ordered *ctype*

Data type | **Model type** | **State frequencies** | **Substitution rates** | **Across-site rate variation** | **Across-tree rate variation**

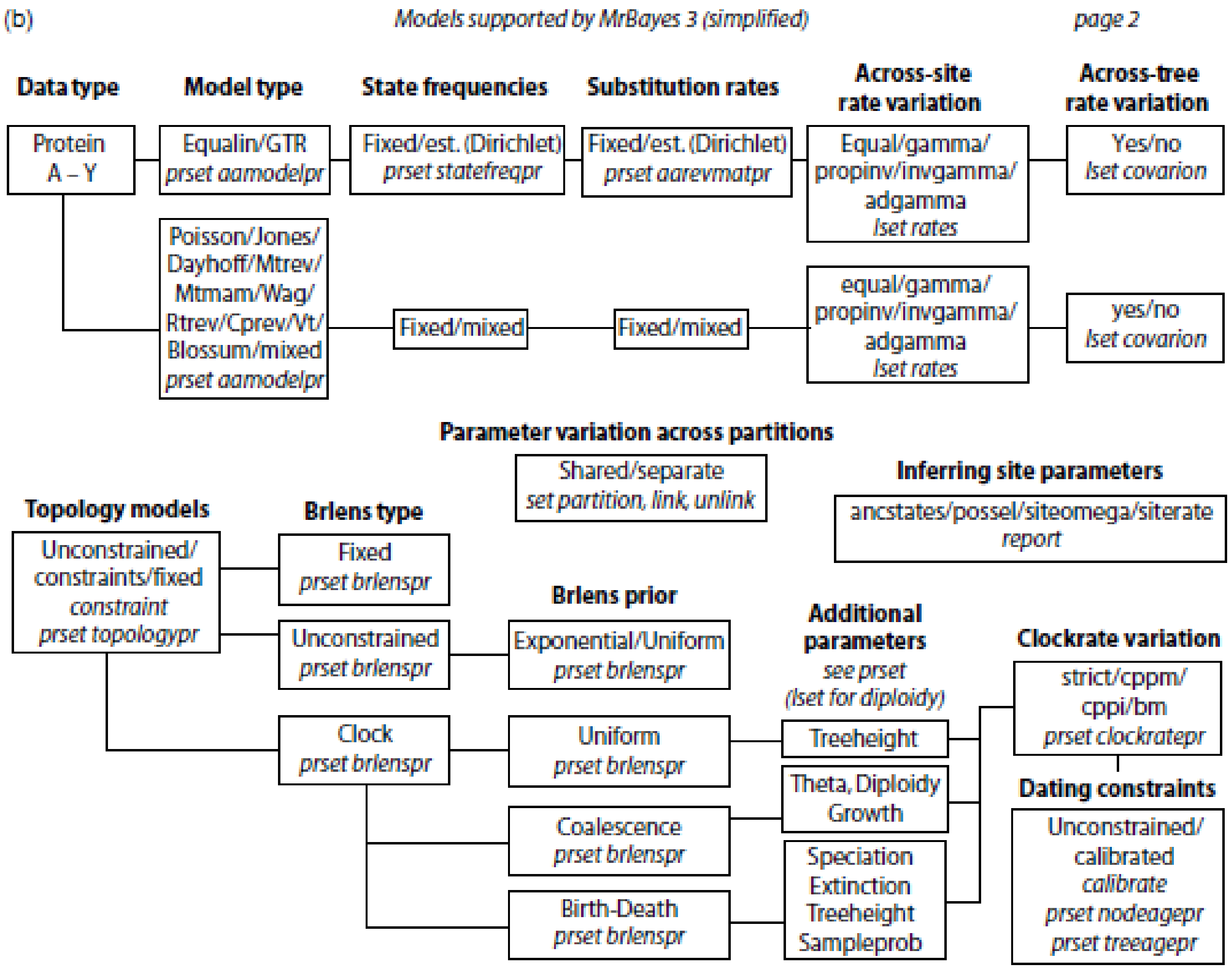
DNA A C G T — 4by4 *lset nucmodel* — Fixed/est. (Dirichlet) *prset statefreqpr* — F81/HKY/GTR *lset nst=1/2/6* — Equal/gamma/ propinv/invgamma/ adgamma *lset rates* — Yes/no *lset covarion*

Doublet *lset nucmodel* — Fixed/est. (Dirichlet) (over 16 states) *prset statefreqpr* — F81/HKY/GTR *lset nst=1/2/6* — Equal/gamma/ propinv/invgamma *lset rates*

Across-site omega variation

Codon *lset nucmodel* — Fixed/est. (Dirichlet) (over 61 states) *prset statefreqpr* — F81/HKY/GTR *lset nst=1/2/6* — Equal/Ny98/M3 *lset omegavar*

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.



(b) *Models supported by MrBayes 3 (simplified)* *page 2*

Data type

Protein A – Y

Model type

Equalin/GTR *prset aamodelpr*

Poisson/Jones/Dayhoff/Mtrev/Mtmam/Wag/Rtrev/Cprev/Vt/Blossum/mixed *prset aamodelpr*

State frequencies

Fixed/est. (Dirichlet) *prset statefreqpr*

Fixed/mixed

Substitution rates

Fixed/est. (Dirichlet) *prset aarevmatpr*

Fixed/mixed

Across-site rate variation

Equal/gamma/propinv/invgamma/adgamma *lset rates*

equal/gamma/propinv/invgamma/adgamma *lset rates*

Across-tree rate variation

Yes/no *lset covarion*

yes/no *lset covarion*

Parameter variation across partitions

Shared/separate *set partition, link, unlink*

Inferring site parameters

ancstates/possel/siteomega/siterate *report*

Topology models

Unconstrained/constraints/fixed *constraint* *prset topologypr*

Brlens type

Fixed *prset brlenspr*

Unconstrained *prset brlenspr*

Clock *prset brlenspr*

Brlens prior

Exponential/Uniform *prset brlenspr*

Uniform *prset brlenspr*

Coalescence *prset brlenspr*

Birth-Death *prset brlenspr*

Additional parameters

see prset (lset for diploidy)

Treeheight

Theta, Diploidy Growth

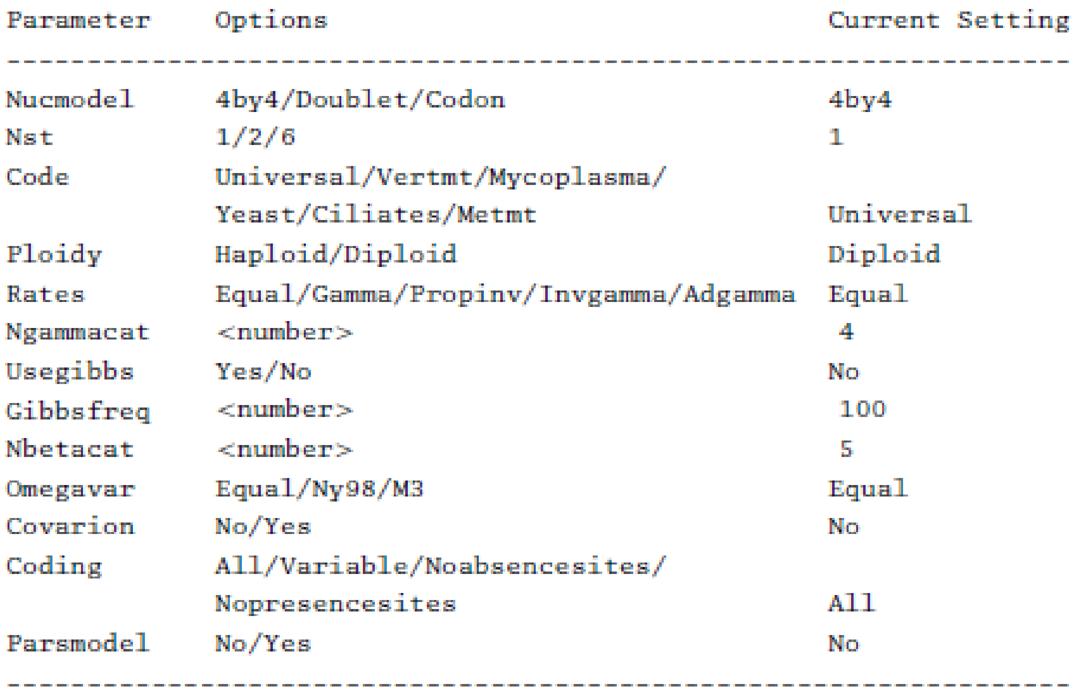
Speciation Extinction Treeheight Sampleprob

Clockrate variation

strict/cppm/cppi/bm *prset clockratepr*

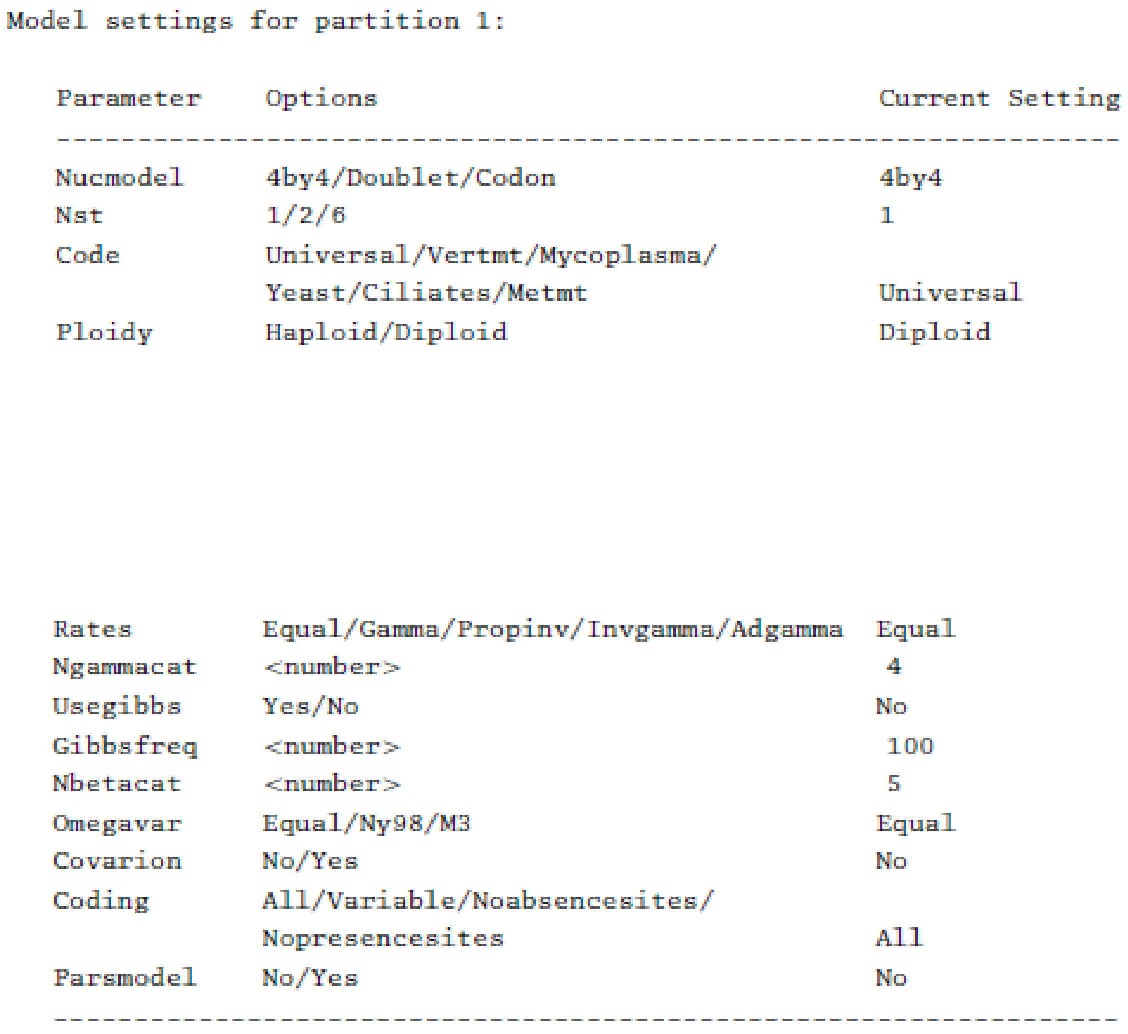
Dating constraints

Unconstrained/calibrated *calibrate* *prset nodeagepr* *prset treeagepr*

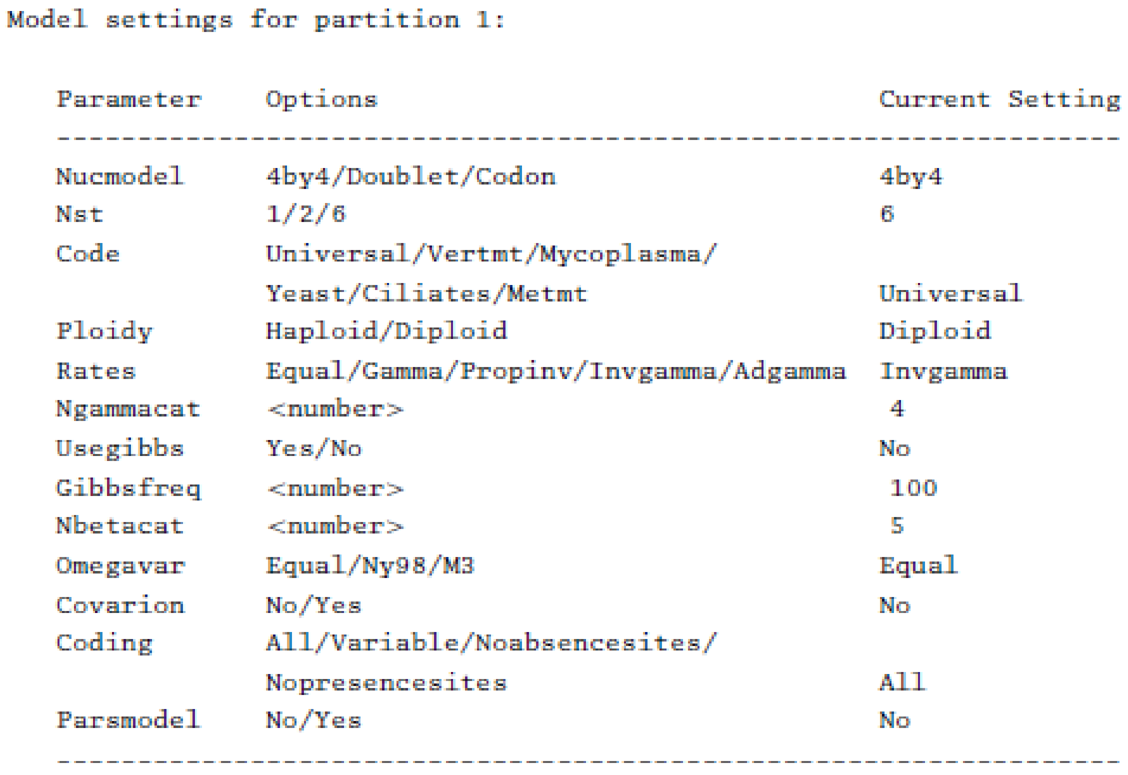


```
Parameter    Options                                  Current Setting
------------------------------------------------------------------------
Nucmodel     4by4/Doublet/Codon                       4by4
Nst          1/2/6                                    1
Code         Universal/Vertmt/Mycoplasma/
             Yeast/Ciliates/Metmt                     Universal
Ploidy       Haploid/Diploid                          Diploid
Rates        Equal/Gamma/Propinv/Invgamma/Adgamma     Equal
Ngammacat    <number>                                  4
Usegibbs     Yes/No                                   No
Gibbsfreq    <number>                                  100
Nbetacat     <number>                                  5
Omegavar     Equal/Ny98/M3                            Equal
Covarion     No/Yes                                   No
Coding       All/Variable/Noabsencesites/
             Nopresencesites                          All
Parsmodel    No/Yes                                   No
------------------------------------------------------------------------
```

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.



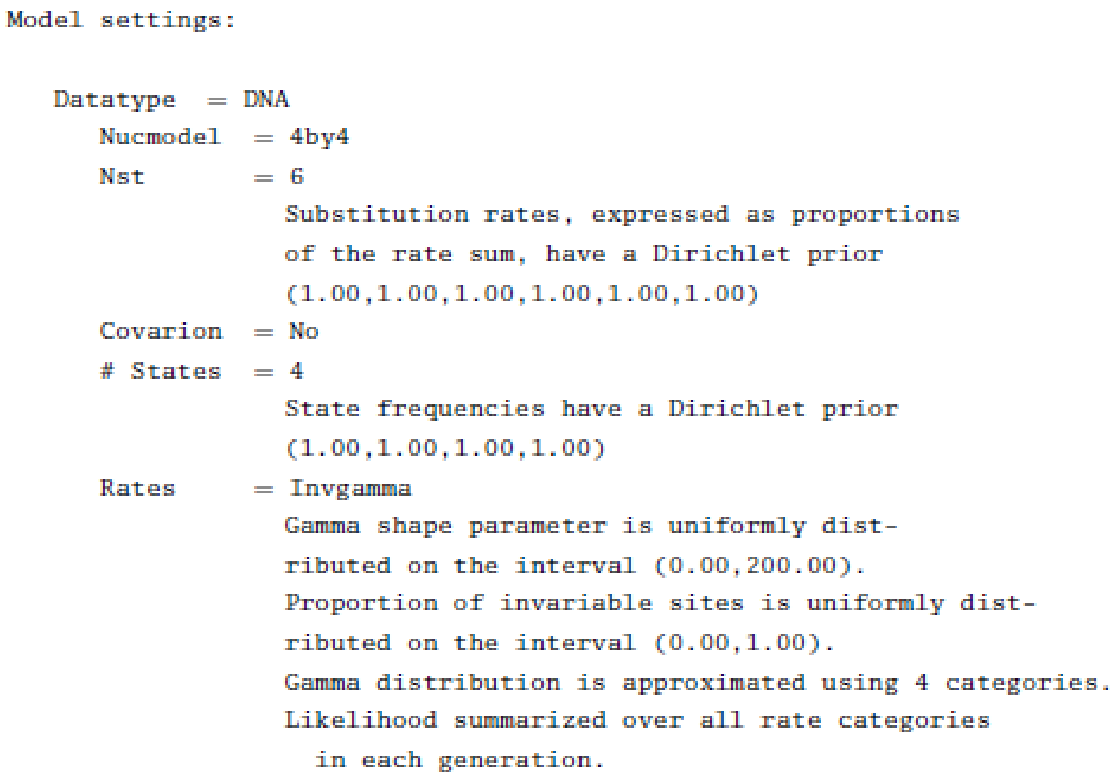
```
Model settings for partition 1:

   Parameter    Options                                Current Setting
   ------------------------------------------------------------------
   Nucmodel     4by4/Doublet/Codon                     4by4
   Nst          1/2/6                                  1
   Code         Universal/Vertmt/Mycoplasma/
                Yeast/Ciliates/Metmt                   Universal
   Ploidy       Haploid/Diploid                        Diploid



   Rates        Equal/Gamma/Propinv/Invgamma/Adgamma   Equal
   Ngammacat    <number>                                4
   Usegibbs     Yes/No                                 No
   Gibbsfreq    <number>                                100
   Nbetacat     <number>                                5
   Omegavar     Equal/Ny98/M3                          Equal
   Covarion     No/Yes                                 No
   Coding       All/Variable/Noabsencesites/
                Nopresencesites                        All
   Parsmodel    No/Yes                                 No
   ------------------------------------------------------------------
```



```
Model settings for partition 1:

   Parameter    Options                                Current Setting
   ------------------------------------------------------------------
   Nucmodel     4by4/Doublet/Codon                     4by4
   Nst          1/2/6                                  6
   Code         Universal/Vertmt/Mycoplasma/
                Yeast/Ciliates/Metmt                   Universal
   Ploidy       Haploid/Diploid                        Diploid
   Rates        Equal/Gamma/Propinv/Invgamma/Adgamma   Invgamma
   Ngammacat    <number>                                4
   Usegibbs     Yes/No                                 No
   Gibbsfreq    <number>                                100
   Nbetacat     <number>                                5
   Omegavar     Equal/Ny98/M3                          Equal
   Covarion     No/Yes                                 No
   Coding       All/Variable/Noabsencesites/
                Nopresencesites                        All
   Parsmodel    No/Yes                                 No
   ------------------------------------------------------------------
```

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

```
Model settings for partition 1:

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    Uniform(0.0,200.0)
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          Uniform
Brlenspr         Unconstrained/Clock          Unconstrained:Exp(10.0)
Treeheightpr     Exponential/Gamma            Exponential(1.0)
Speciationpr     Uniform/Exponential/Fixed    Uniform(0.0,10.0)
Extinctionpr     Uniform/Exponential/Fixed    Uniform(0.0,10.0)
Sampleprob       <number>                      1.00
Thetapr          Uniform/Exponential/Fixed    Uniform(0.0,10.0)
Nodeagepr        Unconstrained/Calibrated     Unconstrained
Treeagepr        Fixed/Uniform/
                 Offsetexponential            Fixed(1.00)
Clockratepr      Strict/Cpp/Bm                Strict
Cppratepr        Fixed/Exponential            Exponential(0.10)
Psigammapr       Fixed/Exponential/Uniform    Fixed(1.00)
Nupr             Fixed/Exponential/Uniform    Fixed(1.00)
Ratepr           Fixed/Variable=Dirichlet      Fixed
------------------------------------------------------------------------
```



```
Model settings:

   Datatype  = DNA
      Nucmodel  = 4by4
      Nst       = 6
                  Substitution rates, expressed as proportions
                  of the rate sum, have a Dirichlet prior
                  (1.00,1.00,1.00,1.00,1.00,1.00)
      Covarion  = No
      # States  = 4
                  State frequencies have a Dirichlet prior
                  (1.00,1.00,1.00,1.00)
      Rates     = Invgamma
                  Gamma shape parameter is uniformly dist-
                  ributed on the interval (0.00,200.00).
                  Proportion of invariable sites is uniformly dist-
                  ributed on the interval (0.00,1.00).
                  Gamma distribution is approximated using 4 categories.
                  Likelihood summarized over all rate categories
                    in each generation.
```

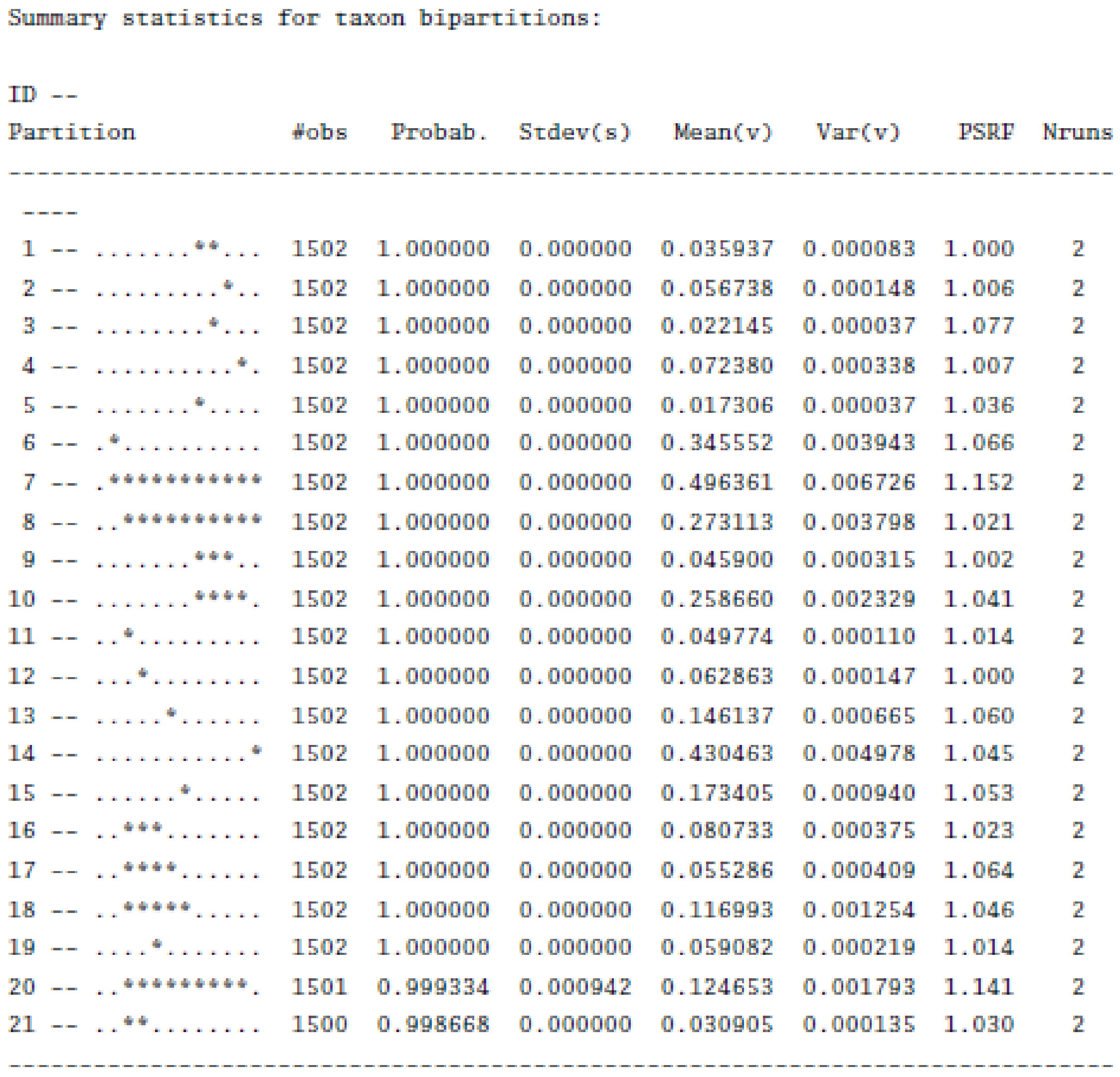
Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

```
The MCMC sampler will use the following moves:
     With prob.  Chain will change
        3.45 %   param. 1 (Revmat) with Dirichlet proposal
        3.45 %   param. 2 (Pi) with Dirichlet proposal
        3.45 %   param. 3 (Alpha) with Multiplier
        3.45 %   param. 4 (Pinvar) with Sliding window
       17.24 %   param. 5 (Tau) and 6 (V) with Extending subtree swapper
       34.48 %   param. 5 (Tau) and 6 (V) with Extending TBR
       17.24 %   param. 5 (Tau) and 6 (V) with Parsimony-based SPR
       17.24 %   param. 6 (V) with Random brlen hit with multiplier
```

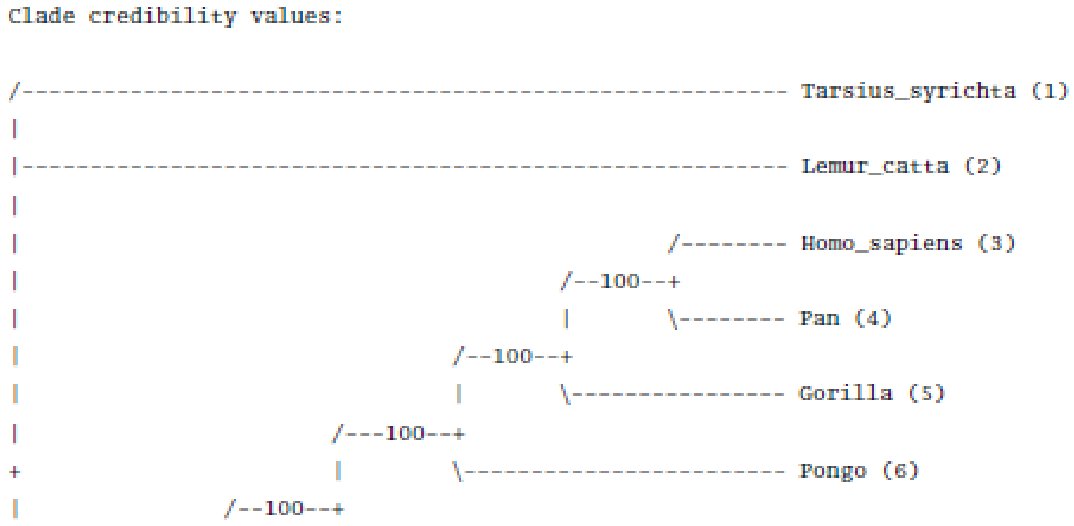
```
#NEXUS
  [ID: 9409050143]
  [Param: tree]
  begin trees;
translate
    1 Tarsius_syrichta,
    2 Lemur_catta,
    3 Homo_sapiens,
    4 Pan,
    5 Gorilla,
    6 Pongo,
    7 Hylobates,
    8 Macaca_fuscata,
    9 M_mulatta,
   10 M_fascicularis,
   11 M_sylvanus,
   12 Saimiri_sciureus;
tree rep.1 = ((12:0.486148,(((((3:0.042011,4:0.065025):0.034344,5:0.051939...
   ...
tree rep.10000 = (((((10:0.087647,(8:0.013447,9:0.021186):0.030524):0.0568...
 end;
```

```
+------------------------------------------------------------+ -5718.96
|                 2                           12             |
|            2                                               |
|              1 2         2                        2        |
|             1    22    1*1   2 22        2   1      2      |
|     2     2 2             1 2 2     2     2        1       |
|       11      1      2        1  2   2 2           2 2   2 |
|   1  2  1       1    12    1 1    *   2       2            |
|  1 11 2  *    2       1        1        2  2      1       *|
| *2      2      1   22              2  1        211    22   |
|        2   1                    1  1    11 1    22  1      |
|*     1       2            2      1        1 2          1*  |
|                            2        1          1      1    |
|           1       111  2    1        1 1                   |
|   22             1                                       1 |
|                                               1      1     |
+------+-----+-----+-----+-----+-----+-----+-----+-----+-----+ -5729.82
-                                                            -

2500                                                         10000
```

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.



```
Summary statistics for taxon bipartitions:

ID --
Partition           #obs   Probab.  Stdev(s)   Mean(v)    Var(v)    PSRF  Nruns
------------------------------------------------------------------------------
 ----
 1 -- .......**...  1502  1.000000  0.000000  0.035937  0.000083  1.000    2
 2 -- .........*..  1502  1.000000  0.000000  0.056738  0.000148  1.006    2
 3 -- ........*...  1502  1.000000  0.000000  0.022145  0.000037  1.077    2
 4 -- ..........*.  1502  1.000000  0.000000  0.072380  0.000338  1.007    2
 5 -- .......*....  1502  1.000000  0.000000  0.017306  0.000037  1.036    2
 6 -- .*..........  1502  1.000000  0.000000  0.345552  0.003943  1.066    2
 7 -- .***********  1502  1.000000  0.000000  0.496361  0.006726  1.152    2
 8 -- ..**********  1502  1.000000  0.000000  0.273113  0.003798  1.021    2
 9 -- .......***..  1502  1.000000  0.000000  0.045900  0.000315  1.002    2
10 -- .......****.  1502  1.000000  0.000000  0.258660  0.002329  1.041    2
11 -- ..*.........  1502  1.000000  0.000000  0.049774  0.000110  1.014    2
12 -- ...*........  1502  1.000000  0.000000  0.062863  0.000147  1.000    2
13 -- .....*......  1502  1.000000  0.000000  0.146137  0.000665  1.060    2
14 -- ...........*  1502  1.000000  0.000000  0.430463  0.004978  1.045    2
15 -- ......*.....  1502  1.000000  0.000000  0.173405  0.000940  1.053    2
16 -- ..***.......  1502  1.000000  0.000000  0.080733  0.000375  1.023    2
17 -- ..****......  1502  1.000000  0.000000  0.055286  0.000409  1.064    2
18 -- ..*****.....  1502  1.000000  0.000000  0.116993  0.001254  1.046    2
19 -- ....*.......  1502  1.000000  0.000000  0.059082  0.000219  1.014    2
20 -- ..*********.  1501  0.999334  0.000942  0.124653  0.001793  1.141    2
21 -- ..**........  1500  0.998668  0.000000  0.030905  0.000135  1.030    2
------------------------------------------------------------------------------
```



```
Clade credibility values:

/------------------------------------------------------ Tarsius_syrichta (1)
|
|------------------------------------------------------ Lemur_catta (2)
|
|                                              /------- Homo_sapiens (3)
|                                      /--100--+
|                                      |       \------- Pan (4)
|                              /--100--+
|                              |       \--------------- Gorilla (5)
|                     /---100--+
+                     |        \----------------------- Pongo (6)
|             /--100--+
```

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.

```
|               |       \--------------------------------- Hylobates (7)
|               |
|               |                                /-------- Macaca_fuscata (8)
|       /--100--+                        /--100--+
|       |       |                        |       \-------- M_mulatta (9)
|       |       |                /--100--+
|       |       |                |       \---------------- M_fascicularis (10)
\--100--+       \-------100------+
     |                        \------------------------ M_sylvanus (11)
        |
     \------------------------------------------------- Saimiri_sciureus (12)
```

```
/--------------------------------------- Tarsius_syrichta (1)
|
|--------------------------- Lemur_catta (2)
|
|                                                    /---- Homo_sapiens (3)
|                                                  /-+
|                                                  | \----- Pan (4)
|                                           /------+
|                                           |      \---- Gorilla (5)
|                                       /---+
+                                       |   \------------ Pongo (6)
|                              /--------+
|                              |        \------------- Hylobates (7)
|                              |
|                              |                         /-- Macaca_fuscata (8)
|                    /---------+                       /-+
|                    |         |                       | \-- M_mulatta (9)
|                    |         |                   /---+
|                    |         |                   |   \---- M_fascicularis (10)
\-------------------+         \-------------------+
                  |                             \------ M_sylvanus (11)
                     |
                  \---------------------------------- Saimiri_sciureus (12)

|--------------| 0.200 expected changes per site
```

Figures and tables are taken from: Lemey, P., Salemi, M., & Vandamme, A. M. (Eds.). (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press. For educational use only.