Molecular Clock Dating using MrBayes

Chi Zhang[1,2,*]

May 28, 2018

[1]*Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden;*
[2]*Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule Zürich, 4058 Basel, Switzerland;*
[*]*E-mail: zhangchicool@gmail.com*

MrBayes is a software for Bayesian phylogenetic inference (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Many new features have been implemented since version 3.2 (Ronquist et al., 2012b), including species tree inference under the multi-species coalescent model (BEST algorithm) (Liu and Pearl, 2007); compound Dirichlet priors for branch lengths (Rannala et al., 2012; Zhang et al., 2012); divergence time estimation using node dating (Hedges and Kumar, 2004; Yang and Rannala, 2006; Ho and Phillips, 2009) or total-evidence dating (Ronquist et al., 2012a; Zhang et al., 2016) methods under (relaxed) molecular clock (Huelsenbeck et al., 2000; Thorne and Kishino, 2002; Drummond et al., 2006; Lepage et al., 2007); marginal model likelihood estimation using stepping-stone sampling (Xie et al., 2011); topology convergence diagnostics using the average standard deviation of split frequencies (ASDSF) (Lakner et al., 2008); BEAGLE library (Ayres et al., 2012) support and parallel computing using MPI (Altekar et al., 2004); among others.

There are two modern approaches on dating species divergence using molecular data: node dating (e.g., Yang and Rannala, 2006; Drummond et al., 2006) and total-evidence dating (e.g., Ronquist et al., 2012a; Zhang et al., 2016). In a Bayesian framework, node dating calibrates one or several internal nodes of the tree, each with a prior distribution derived from the fossil record. While total-evidence dating uses the morphological data from the fossil record and morphological and sequence data from extant taxa together to infer the tree and divergence times. The age of each fossil is assigned a prior distribution directly.

Several steps involve in Bayesian dating analysis, importantly including data partitioning, specifying evolutionary model, calibrating internal nodes or fossils, and setting priors for the tree and the molecular clock model. In this tutorial, I demonstrate the molecular clock dating functionalities in MrBayes 3.2 step by step, while focusing on total-evidence dating, using an example dataset truncated from the Hymenoptera data analyzed in Ronquist et al. (2012a); Zhang et al. (2016).

1   Getting Started

The program MrBayes is available from http://mrbayes.net (latest version 3.2.7). After downloading and installing MrBayes by following the manual, we execute the program from terminal (or command prompt) by typing

`mb` (assuming the executable is in the user path and is named `mb` on Mac OSX/Linux or `mb.exe` on Windows).

```
                  MrBayes v3.2
             (Bayesian Analysis of Phylogeny)
       Distributed under the GNU General Public License
MrBayes >
```

The prompt `MrBayes >` at the bottom means that MrBayes is running and ready for your commands. In the following tutorial, the commands for MrBayes are colored RED, the commands typing in terminal (or command prompt) are colored BLUE.
(!!) Direct copy and paste the commands here may result in improper spacing or line breaking in the program. Please copy the plain-text commands included in the example file.

2 Run the analyses

2.1 Partition the data

The full data includes 60 extant and 45 fossil hymenopteran taxa and 8 outgroup taxa. The alignment is divided into 8 partitions (Ronquist et al., 2012a). To make it computationally tractable for this tutorial, the data is truncated to 10 extant taxa (including 1 outgroup and 9 hymenopteran taxa) and 10 fossils. Only the morphology (200 characters), 16S (100 sites) and EF1α (210 sites) partitions are included, all partially.

Use the `execute` command (`exe` for short) to read in the data named `hym.nex`.

```
execute hym.nex
```

The following commands partition the data into four partitions: the morphology, 16S, 1st and 2nd codon positions of Ef1α, and 3rd codon positions of Ef1α, and exclude the incompatible (constant) characters (these commands are included in the example data and have been executed automatically when reading it in).

```
charset MV = 1-200
charset 16S = 201-300
charset Ef1a = 301-510
```

```
charset Ef1a12 = 301-510\3 302-510\3
charset Ef1a3 = 303-510\3
partition four = 4:  MV, 16S, Ef1a12, Ef1a3
set partition = four
exclude 7 31 61 83 107 121 122 133 182 183 198
```

It is much worth mentioning that the build-in help in MrBayes is very informative and explanatory. We can type `help` followed by the keyword to retrieve the corresponding help message. For example,

```
help charset
help partition
help lset
```

2.2   Evolutionary model

For the morphology partition, we use the Mk Model (Lewis, 2001) with variable ascertainment bias (only variable characters scored), equal state frequencies and gamma rate variation across characters.

```
lset applyto = (1) coding = variable rates = gamma
```

If instant change is only allowed between adjacent states (e.g., only $0 \leftrightarrow 1$ and $1 \leftrightarrow 2$), these characters are specified using `ctype ordered`. The other characters are thus allowed to change instantly from one state to another.

```
ctype ordered:  20 23 27 30 36 41 42 44 46 48 59 65 75 78 79 89
   99 112 117 134 146 157 159 171 185 191 193 196
```

For the molecular partitions, we use the general time-reversible model with gamma rate variation across sites (GTR+Γ) (Yang, 1994a,b). The widely used invariable sites and gamma (+I+Γ) model is pathological due to strong correlation between the proportion of invariable sites (p_0) and the gamma shape parameter (α) (e.g., Zhang et al., 2012), and is avoided.

```
lset applyto = (2,3,4) nst = 6 rates = gamma
```

The numbers after `applyto` should match the order of partitions defined above. The default prior for the shape parameter α of gamma(α, β) ($\alpha = \beta$, Figure 1) is exponential(1.0), which can be changed using `prset shapepr`. We keep the default here.

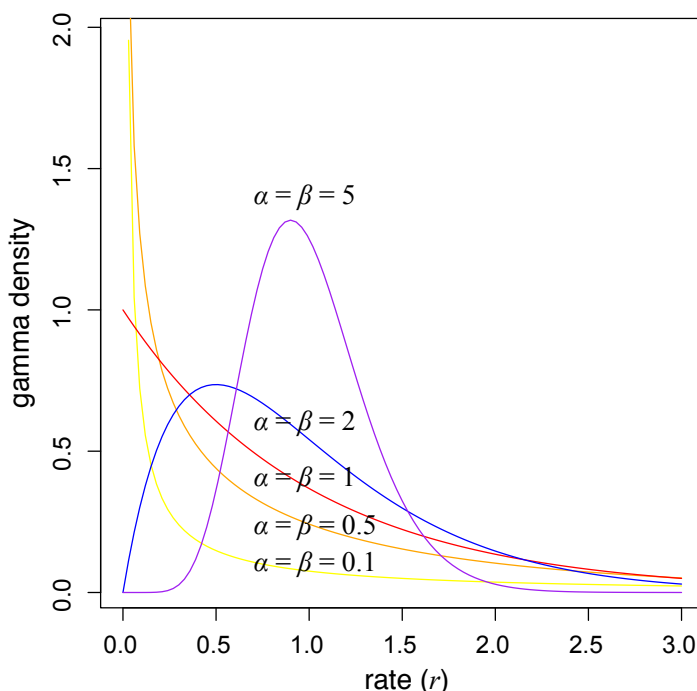Different partitions are assumed to have independent substitution parameters, thus we unlink them.

Figure 1: Probability density function of the gamma distribution. The shape α and rate β are fixed equal so that the mean is 1.0. The exponential distribution is a special case of gamma when $\alpha = 1$.

```
unlink statefreq = (all) revmat = (all) shape = (all)
```

It is reasonable to account for evolutionary rate variation across partitions, so that different partitions can have different substitution rates.

```
prset applyto = (all) ratepr = variable
```

The partition-specific rate-multipliers, $\{m_i\}$, have mean equal to 1.0. Specifically, $\sum_i p_i m_i = 1$, where p_i is the proportion of sites in partition i to the total number of sites. By default, `ratepr = variable` specifies a uniform Dirichlet prior for $\{p_i m_i\}$ (A special case of Dirichlet is uniform on 2 partitions). Note the difference between partition-specific rate and site-specific rate inside a partition. The sites in partition i have the same partition-specific rate m_i, while each site again has a site-specific rate r_i from gamma(α_i, α_i) distribution in the $+\Gamma$ model.

Molecular data only provide evolutionary distances in units of evolution-

ary change, such as substitutions per site. Branch lengths thus are the product of the geological time duration (e.g. in Myr) and the evolutionary rate (e.g. in substitutions per site per Myr). To estimate times and rates separately, it is necessary to introduce additional model assumptions.

Early studies assumed the evolutionary rate is constant over time (called global clock or strict clock) (Zuckerkandl and Pauling, 1965). To set a strict clock model with mean rate at 0.001 per site per Myr, we can use a lognormal(-7,0.6) prior for the clock rate, c. The branch length v_j is the geological time duration t_j multiplied by c.

```
prset clockratepr = lognorm(-7,0.6)
prset clockvarpr = strict
```
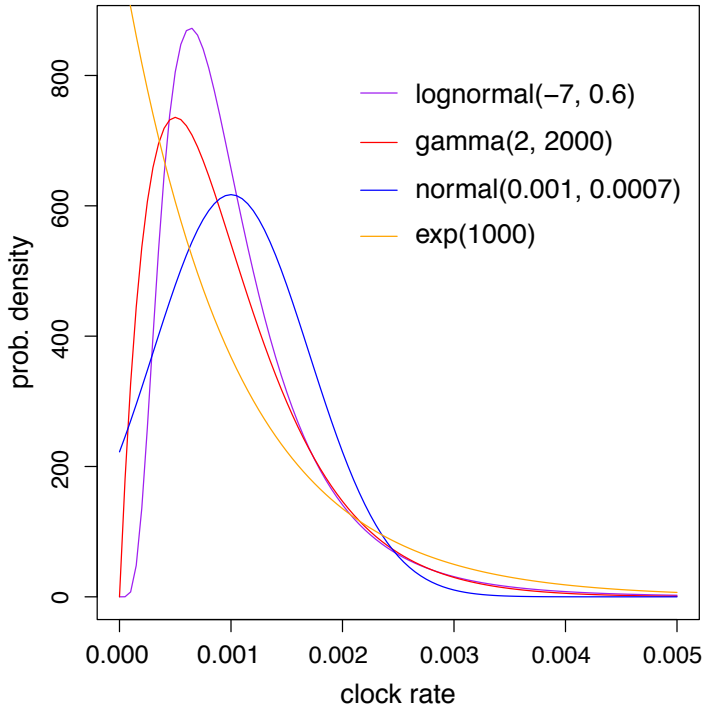

Figure 2: Probability density functions of normal, lognormal and gamma distributions, all with mean 0.001 and standard deviation 0.0007. The exponential distribution has mean and standard deviation both equal to 0.001.

There are several options for the clock rate prior, including `fixed`, `normal`

6

(truncated at 0), `lognormal`, and `gamma` (use `help prset` for more details). The probability density functions of the distributions (all with mean 0.001) are shown in Figure 2. Here I take some space to explain the lognormal distribution, as I think it is confusing. If $x \sim$ lognormal(μ, σ) then $log(x) \sim$ normal(μ, σ). The mean of x is $e^{(\mu+\sigma^2/2)}$ and the median of x is e^μ. The variance of x is $(e^{\sigma^2} - 1)e^{(2\mu+\sigma^2)}$. Thus the mean for lognormal(-7, 0.6) is $e^{(-7+0.6^2/2)} = 0.001$ and the standard deviation is $\sqrt{(e^{0.6^2} - 1)e^{(-2\times 7+0.6^2)}} = 0.0007$.

It is more realistic to assume variable evolutionary rate over time. There are three relaxed clock models implemented in MrBayes: compound Poisson process (CPP, Huelsenbeck et al., 2000), autocorrelated lognormal (TK02, Thorne and Kishino, 2002) and independent gamma rate (IGR, Lepage et al., 2007). We just focus on the IGR and TK02 model, as the CPP model is not working currently for total-evidence dating (see below).
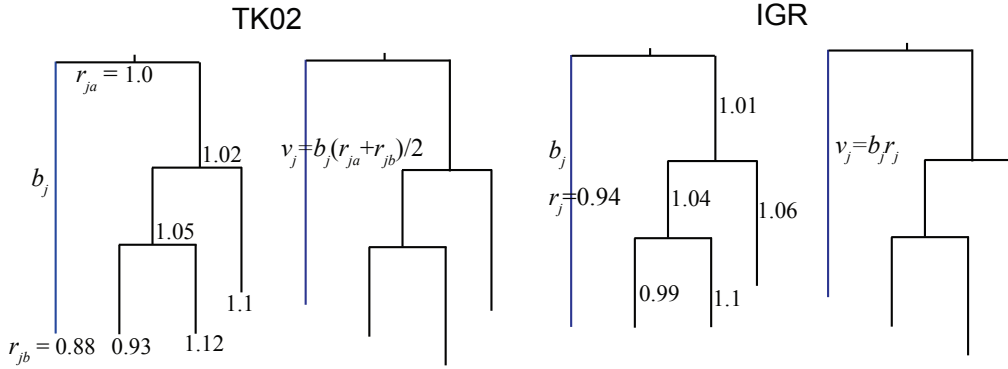

Figure 3: Illustration of the TK02 and IGR relaxed clock models. The relaxed clock tree on the right is resulted from the strict clock tree by multiplying its branch lengths with the relaxed clock rates.

In the TK02 model (Figure 3), the (relative) evolutionary rate changes along the branches as Brownian motion on the log scale, starting from 1.0 (0.0 on the log scale) at the root. The rate at the end of a branch j on the tree is lognormal distributed with mean (not the log of the mean) equal to the rate at the beginning of the branch and variance equal to $b_j\sigma^2_{TK}$, where b_j is the product of geological time duration t_j and the base clock rate c. The branch length v_j is then calculated as b_j multiplied by the arithmetic mean of the two rates at both ends. The TK02 model is specified using

```
prset clockvarpr = tk02
```

The prior for c was already set above using `clockratepr`. The prior for σ^2_{TK} is specified by

```
prset tk02varpr = exp(1)
```

The IGR model assumes that the (relative) rate for branch j is gamma distributed with mean 1.0 and variance σ^2_{IG}/b_j (Figure 3). The rates for different branches are independent but not identical. The branch length v_j is then calculated as b_j multiplied by the branch rate. In equivalent, v_j is gamma distributed with mean b_j and variance $b_j\sigma^2_{IG}$ (original definition in Lepage et al., 2007). The IGR model is specified using

```
prset clockvarpr = igr
```

and the prior for σ^2_{IG} is specified by

```
prset igrvarpr = exp(10)
```

Note that when the branch rates are all fixed to 1.0 for TK02 or IGR, it becomes the strict clock model. For the likelihood calculation (Felsenstein, 1981), the evolutionary distance (number of substitutions per site) for branch j, site k in partition i is $d_{jik} = v_j m_i r_{ik}$ (where m_i is the partition-specific rate, r_{ik} is the site-specific rate).

Before doing total-evidence dating and node dating, we first define the outgroup and fossil group,

```
outgroup Raphidioptera
taxset fossils = Asioxyela Nigrimonticola Xyelotoma
   Undatoma Dahuratoma Cleistogaster Ghilarella
   Mesorussus Prosyntexis Pseudoxyelocerus
```

and some constraints for later use. Note these constraints are *not* enforced until we set `topologypr` explicitly (see below).

```
constraint root = 1-.
constraint HymenFossil = 2-.
constraint Hymenoptera = 2-10
constraint Holometabola = 1-10
constraint Tenthredinidae = 3-5
constraint CepSirOruApo = 7-10
```

2.3 Total-evidence dating

In the following, we incorporate fossil information and assign priors for the geological times. This is a typical step in total-evidence dating, where we calibrate the fossils instead of the internal nodes.

```
calibrate
   Asioxyela = unif(228,242)
   Nigrimonticola = unif(152,163)
   Xyelotoma = unif(152,163)
   Undatoma = unif(145,152)
   Dahuratoma = fixed(134)
   Cleistogaster = unif(168,191)
   Ghilarella = unif(113,125)
   Mesorussus = unif(94,100)
   Prosyntexis = unif(80,86)
   Pseudoxyelocerus = fixed(182)
prset nodeagepr = calibrated
```

The last command `nodeagepr` is *necessary* to enable the calibrations. There is more information in `help calibrate`.
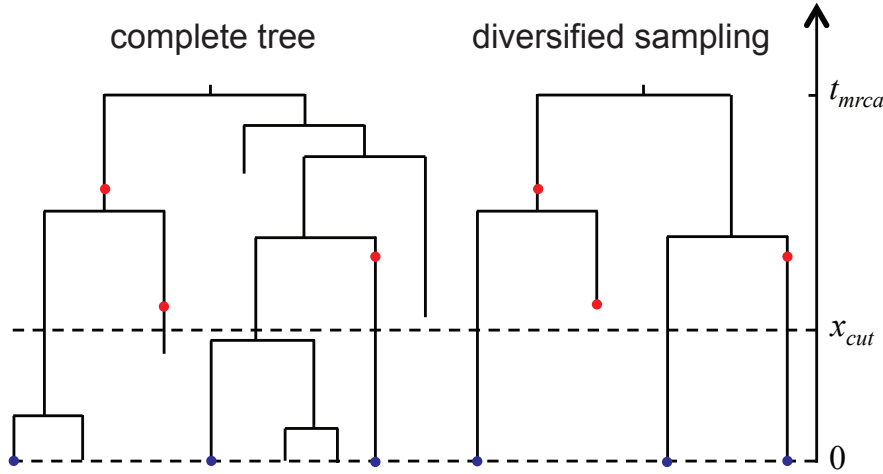

Figure 4: The fossilized birth-death (FBD) process under diversified sampling.

The speciation, extinction and fossilization process is explicitly modeled

9

using the fossilized birth-death (FBD) process (Stadler, 2010; Heath et al., 2014; Gavryushkina et al., 2014; Zhang et al., 2016). The complete tree is generated from the birth-death process, with speciation rate λ and extinction rate μ (Figure 4). Two sampling strategies are assumed. For random sampling, the extant taxa are sampled uniformly at random with probability ρ, while the fossils are sampled with a constant rate ψ through time. For diversified sampling (Zhang et al., 2016), exactly one representative taxa per clade descending from time x_{cut} is sampled, resulting a proportion of ρ extant taxa sampled. The fossils are sampled with a constant rate ψ before x_{cut} and zero after. The observed FBD tree is resulted when all lineages without a fossil or sampled extant descendant have been pruned. A fossil can be either tip or ancestor of other taxa (Figure 4).

The FBD prior is enabled using

```
prset brlenspr = clock:fossilization
```

The sampling proportion ρ is fixed to 0.0001, based on the living number of hymenopteran species at about $10/0.0001 = 100,000$.

```
prset sampleprob = 0.0001
```

The sampling strategy is set to be diversified, as it is arguably suitable for this species-level dataset.

```
prset samplestrat = diversity
```

For inference, rather then operating on λ, μ and ψ, which may range from 0 to infinity, we re-parametrize them as $d = \lambda - \mu$ (net diversification), $e = \mu/\lambda$ (turnover), and $s = \psi/(\mu + \psi)$ (fossil sampling proportion), so that the later two parameters range from 0 to 1. The default priors for d, e, and s are:

```
prset speciationpr = exp(10)
prset extinctionpr = beta(1,1)
prset fossilizationpr = beta(1,1)
```

In order to root the tree properly, we enable the constraint `HymenFossil` defined above. This forces the Hymenoptera with fossils form a monophyletic group.

```
prset topologypr = constraint(HymenFossil)
```

The FBD prior is conditioned on the root age of the tree (t_{mrca}). It is important to set it properly. Here we use an offset exponential distribution with minimal age 300 Ma and mean age 390 Ma.

```
prset treeagepr = offsetexp(300,390)
```

You may have noticed that various prior distributions for tree age (and for calibration) are available, and the way specifying the prior parameters as minimal age and mean age is *different* from the parameterization used elsewhere in the program. This setting is aiming to ease the user and to assure a proper prior is specified. Several probability densities all with mean 390 and minimal 300 are shown in Figure 5, including the offsetexp(300, 390) and other candidates (offsetlognormal(300, 390, 118), offsetgamma(300, 390, 64), truncatednormal(300, 390, 60)).

Here I mention that there is another tree prior, the uniform prior (`clock:uniform`) (Ronquist et al., 2012a), that fits in the total-evidence dating framework. The model assumes that the internal nodes are draw from uniform distributions and the fossils are only tips of the tree (so-called tip dating). It is also conditioned on the root age and requires setting `treeagepr`, but there is no speciation-extinction-fossilization-sampling parameter.

All models and priors are set :). To see the current settings, use

```
showmodel
```

These messages will also be printed at the beginning of the run.

Now it is time to run the analysis. We first set the Markov chain Monte Carlo (MCMC) (Metropolis et al., 1953; Hastings, 1970) without running it.

```
mcmcp nrun = 2 nchain = 4 ngen = 500000 samplefr = 100
mcmcp filename = hym.te printfr = 1000 diagnfr = 5000
```

This setting uses 2 independent runs and 4 chains (1 cold and 3 heated) per run for 500,000 iterations, and samples every 100 iterations. The output file names will be hym.te.∗. The chain states will be printed to screen every 1000 iterations. The convergence diagnostics (acceptance ratios and average standard deviation of split frequencies (ASDSF)) will be printed every 5000 iterations. See `help mcmc` for other default settings that you may want to change as well.

To run the MCMC, just type

```
mcmc
```

then you will see the log likelihoods printed to the screen, with [ ] for the cold chain and ( ) for the hot, and the time left to finish.

```
0 -- [-3530.775] (-3407.296) (-3461.278) (-3476.937) *
   [-3537.572] (-3494.511) (-3460.711) (-3407.261)
```
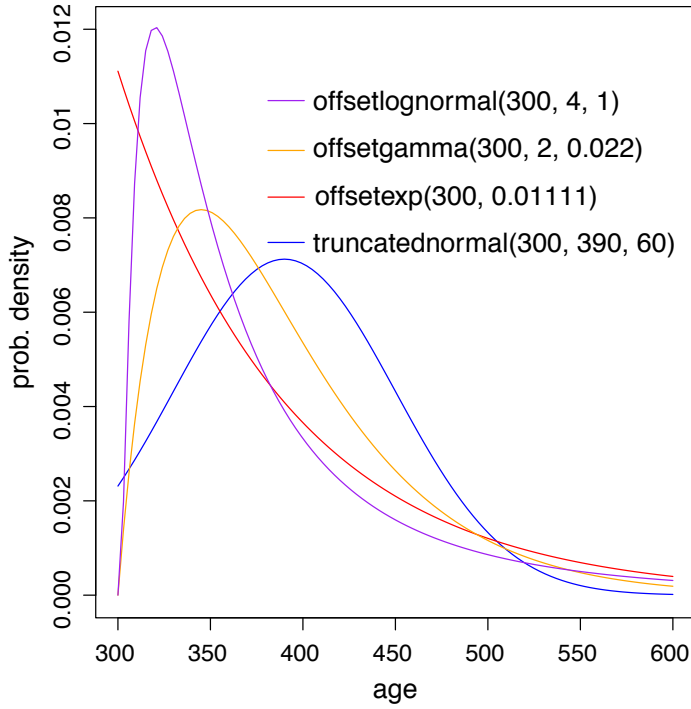
Figure 5: Probability density functions of offsetlognormal(m, μ, σ), offsetgamma(m, α, β), offsetexp(m, λ), and truncatednormal(m, μ, σ). For offsetlognormal the mean is $m+e^{(\mu+\sigma^2/2)}$ (the median is $m+e^\mu$) and the standard deviation is $\sqrt{(e^{\sigma^2}-1)e^{(2\mu+\sigma^2)}}$; for offsetgamma the mean is $m+\alpha/\beta$ and the standard deviation is $\sqrt{\alpha}/\beta$; for offsetexp the mean is $m+1/\lambda$ and the standard deviation is $1/\lambda$; for truncatednormal the mean is slightly larger then μ and the standard deviation is slightly smaller then σ, depending on the left truncation point.

```
1000 -- (-3281.546) [-3286.212] (-3193.371) (-3198.067) *
   (-3249.083) (-3208.938) [-3214.627] (-3142.797) -- 0:08:19
...
   Average standard deviation of split frequencies:  0.101638
101000 -- (-2884.632) (-2873.325) (-2869.337) [-2874.607] *
   [-2876.804] (-2871.278) (-2872.583) (-2867.121) -- 0:07:22
...
500000 -- (-2871.038) (-2876.999) (-2866.260) [-2866.105] *
   (-2853.373) [-2868.731] (-2881.175) (-2867.806) -- 0:00:00
   Average standard deviation of split frequencies:  0.085192
Continue with analysis?  (yes/no):  no
...
```

The ASDSF is decreasing slowly toward 0, indicating the tree topologies sampled from different runs are getting similar and converging to the same (stationary) distribution. We stop the run by typing **no** when prompted. Then at the end, it will print the chain swap information.

To summarize the results, we use

sump
sumt

First we look at the outputs from **sump**. The likelihood traces for the two runs are mixed together, this is also a good indication of convergence. It also helps us to determine the number of burnin. By default, 25% of the samples are discarded. This can be changed by **sump burninfrac = 0.4**, say. The traces are followed by a table of the parameter estimates. Since MCMC is sampling correlated samples, the effective sample size (ESS) is smaller than the actual number of samples (500000/100=5000 in this example). Ideally ESS should larger than 200 for all parameters to make good estimates. We do not run longer for this tutorial, but note that we can increase the number of iterations and append to the current samples.

mcmc ngen=1000000 append=yes

Then go the the outputs from **sumt**. It first lists the taxon bipartitions and the corresponding IDs. The root ID is 0; the extant taxa IDs are 1 to 10; the fossil IDs are 11 to 20. The following summaries are matched to the bipartition IDs. The tree is printed at the end.

The consensus tree including all fossils is highly unresolved due to uncertainty in the placement of the fossils. In order to display the ages clearly, we

13

remove the fossils and redraw an extant taxa tree. The output filename is changed to avoid overwriting the existing ones.

```
delete fossils
sumt output = hym.rf
```

2.4   Node dating

In node dating, we calibrate internal nodes instead of fossils. The calibration priors are derived from second interpretation of the fossil record. Thus we remove the fossils, the morphological characters of fossils are not used.

```
delete fossils
exclude 24 130 168
```

Then we calibrate the root, and another two internal nodes. These three probability densities are shown in Figure 6.

```
calibrate
   root = offsetexp(300,390)
   Tenthredinidae = offsetgamma(100,150,25)
   CepSirOruApo = truncatednormal(140,175,25)
prset nodeagepr = calibrated
```

Note that the root age calibration here is equivalent to `prset treeagepr = offsetexp(300,390)` (see above), thus either of them is sufficient.

We can still use the FBD prior but fix the fossilization rate to 0 (no fossil sampling).

```
prset fossilizationpr = fixed(0)
```

Alternatively, we can use the birth-death prior.

```
prset brlenspr = clock:birthdeath
```

Comparing to the FBD prior, the birth-death prior does not assume fossil sampling, thus `fossilizationpr` is irrelevant. The priors for the root age, speciation rate, extinction rate, sampling strategy (diversified) and sampling proportion (0.0001) are not changed (see above).

The uniform tree prior (`clock:uniform`) is also applicable, then the speciation-extinction-fossilization-sampling priors are irrelevant.

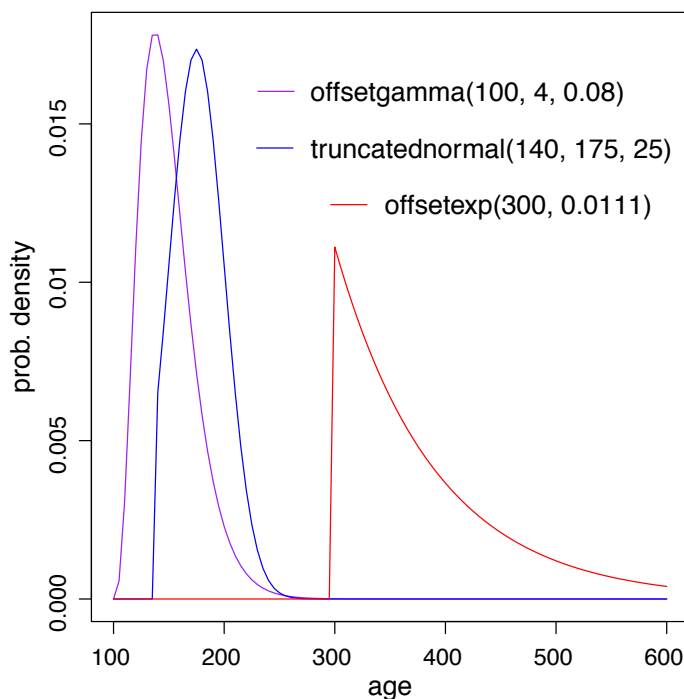It is important to force the calibrated node to be monophyletic, and enable the constraints using

Figure 6: Probability density functions of offsetexp(300, 0.011) (offsetexp(300, 390) in MrBayes), offsetgamma(100, 4, 0.08) (offsetgamma(100, 150, 25) in MrBayes), and truncatednormal(140, 175, 25).

```
prset topologypr = constraint(Hymenoptera,
                        Tenthredinidae,CepSirOruApo)
```

The Hymenoptera constraint helps to root the tree properly.

We change the output filename to avoid overwriting existing ones. We also need to reset the starting values, as this node dating run is continued after the total-evidence dating run in the same session.

```
mcmcp filename = hym.nd startp = reset startt = random
```

The other settings are kept the same as in total-evidence dating above.

```
mcmc
sump
sumt
```

15

2.5 Batch mode

In practice, instead of typing commands line by line in MrBayes, we usually write the commands in file, either after the data block as in this example, or in separate files. Each command should end with a semicolon ;. Texts in a pair of square brackets [ ] is a comment, and is ignored by the program. Below is an example of commands in a separate file (named mbcmd.nex), in the same folder as hym.nex.

```
#NEXUS
Begin mrbayes;
    [read in data]
    exe hym.nex;
    [partition data]
    ...
    [model and prior]
    ...
    prset brlenspr = uncons:gammadir(1,1,1,1);
    ...
    mcmc;
    sump;
    sumt;
end;
```

To run the analysis, just type

execute mbcmd.nex

after the MrBayes > prompt (MrBayes is running), or

mb mbcmd.nex

in terminal (or command prompt. MrBayes is not running).

2.6 A non-clock analysis

Before looking at the results, we do a non-clock analysis without assuming molecular clock, to compare the tree with the relaxed clock tree. The branch lengths are measured in expected substitutions per site. This is a typical analysis most people do using MrBayes.

We do not use fossils either, and do not constrain the topology so that they are uniformly distributed. The evolutionary model is kept the same but the setting for clock rate is ignored.

```
delete fossils
prset brlenspr = uncons:gammadir(1,1,1,1)
prset topologypr = uniform
mcmcp filename = hym.un
mcmc
sump
sumt
```

Note the prior for branch lengths is a gamma-Dirichlet(α_T, β_T, α, c) prior (Rannala et al., 2012; Zhang et al., 2012). The prior assigns a gamma(α_T, β_T) distribution for the tree length (sum of branch lengths), and a Dirichlet(α, c) prior for the proportion of branch lengths to the tree length. In the Dirichlet, the parameter for external branches is α and for internal branches is αc, so that the prior ratio between internal and external branch is c. In this case, we assign gamma(1, 1) (i.e. exponential(1), cf. Figure 1) for the tree length and uniform Dirichlet for the proportions.

The gamma-Dirichlet (compound Dirichlet) prior was shown to help avoiding overestimation of tree length (Zhang et al., 2012) and is now the default prior in MrBayes (after v3.2.3). It was independent and identically distributed (i.i.d.) exponential(10) for each branch length before. For this 10 extant taxa example, the prior distributions for the tree length and for each branch length are shown in Figure 7. The gamma-Dirichlet prior appears less informative and more flexiable than the i.i.d. exp(10).
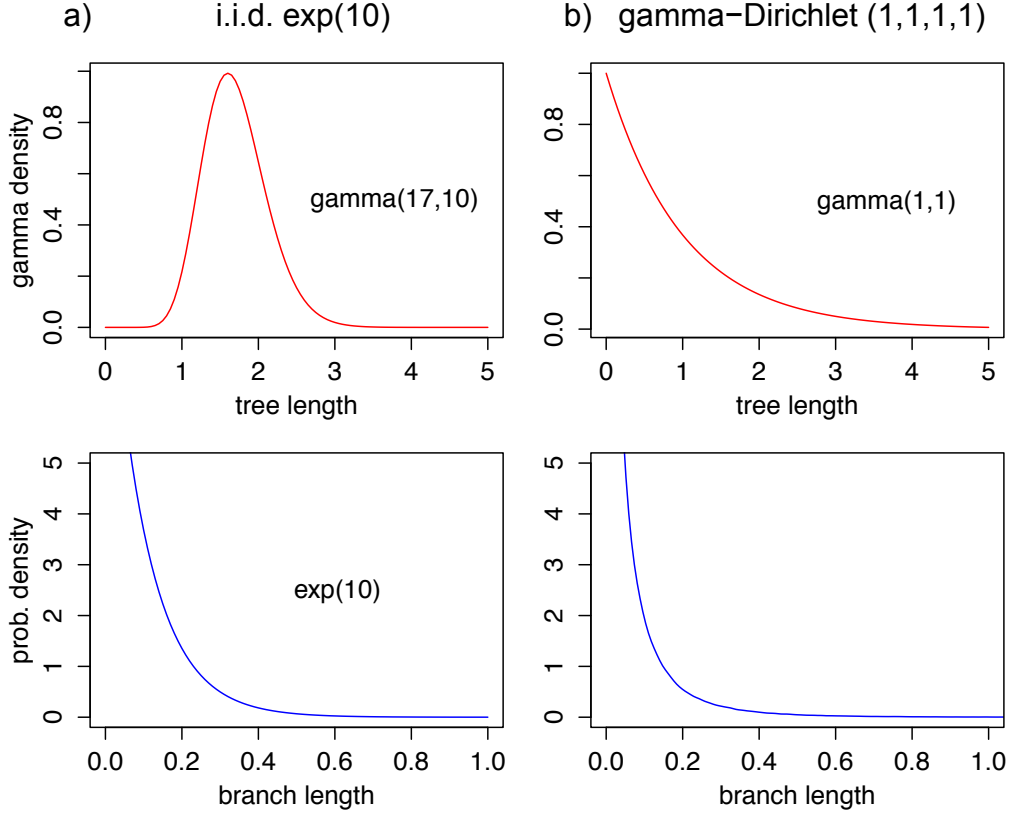
Figure 7: Probability densities for the tree length and for each branch length under a) i.i.d. exponential(10) prior and b) gamma-Dirichlet(1, 1, 1, 1) prior. The sum of n i.i.d. gamma(α, β) is gamma($n\alpha$, β), and the proportion to the sum is Dirichlet(α). Thus the sum of $n = 2 \times 10 - 3 = 17$ i.i.d. exp(10) (gamma(1, 10)) is gamma-Dirichlet(17, 10, 1, 1). Under gamma-Dirichlet(1, 1, 1, 1), the priors for branch lengths are i.i.d. but not exactly gamma.

3 Summarize the results

The posterior estimates on the screen are also output to files, with names specified using the commands above, each with a particular extension. These files can be opened using a plain-text editor.

The partition rate multipliers m_i are in file `hym.*.pstat`. The first and second codon positions (`m{3}`) evolve much slower than the third codon position (`m{4}`). The morphology (`m{1}`) and 16S (`m{2}`) partitions evolve at similar rate.

There are many tools to visualize the trees in file `hym.*.con.tre`. Here I use FigTree http://tree.bio.ed.ac.uk/software/figtree/. It is very compatible with the consensus tree format output from MrBayes. There are various options on the left panel of FigTree to adjust the display.

The clock trees from total-evidence dating and node dating under diversified FBD prior and IGR model are shown in Figure 8. The ages of root and hymenopteran crown are shown in Table 1. These estimates match the corresponding node bars in Figure 8. The HPD intervals are in file `hym.*.vstat`. The bipartition ID of `age{all}[.]` is in `hym.*.parts`. The ID for root is 0 which includes all taxa (all `*`), while the ID for Hymenoptera is that excludes the outgroup and fossils (`.*********`).

Comparing the non-clock tree (Figure 9) with the clock trees (Figure 8), it is obvious that the evolutionary rate is not constant over time. The Xyela and Onycholyda lineages evolve much slower than the Orussus and Vespidae clade, and there are indeed dramatic rate changes between adjacent branches. The IGR model is thus presumably more suitable than the autocorrelated TK02 model.

The ages inferred from the total-evidence dating are slightly younger than those from the node dating under the IGR model (Figure 8), with relatively narrower credibility intervals. The FBD model in the total-evidence dating approach models the fossilization (and sampling) process explicitly and incorporates all available information from fossil record. In comparison, the node dating approach discards the fossil morphologies and stratigraphic times, but uses second interpretation of the fossil record as node calibration distributions. Thus the total-evidence dating approach appears more objective and rigorous, and provides an ideal platform for exploring and further improving the models used for Bayesian divergence-time estimation.

Nevertheless, the results from this truncated small dataset are mainly for demonstration of MrBayes' functionalities. For more results and discussions
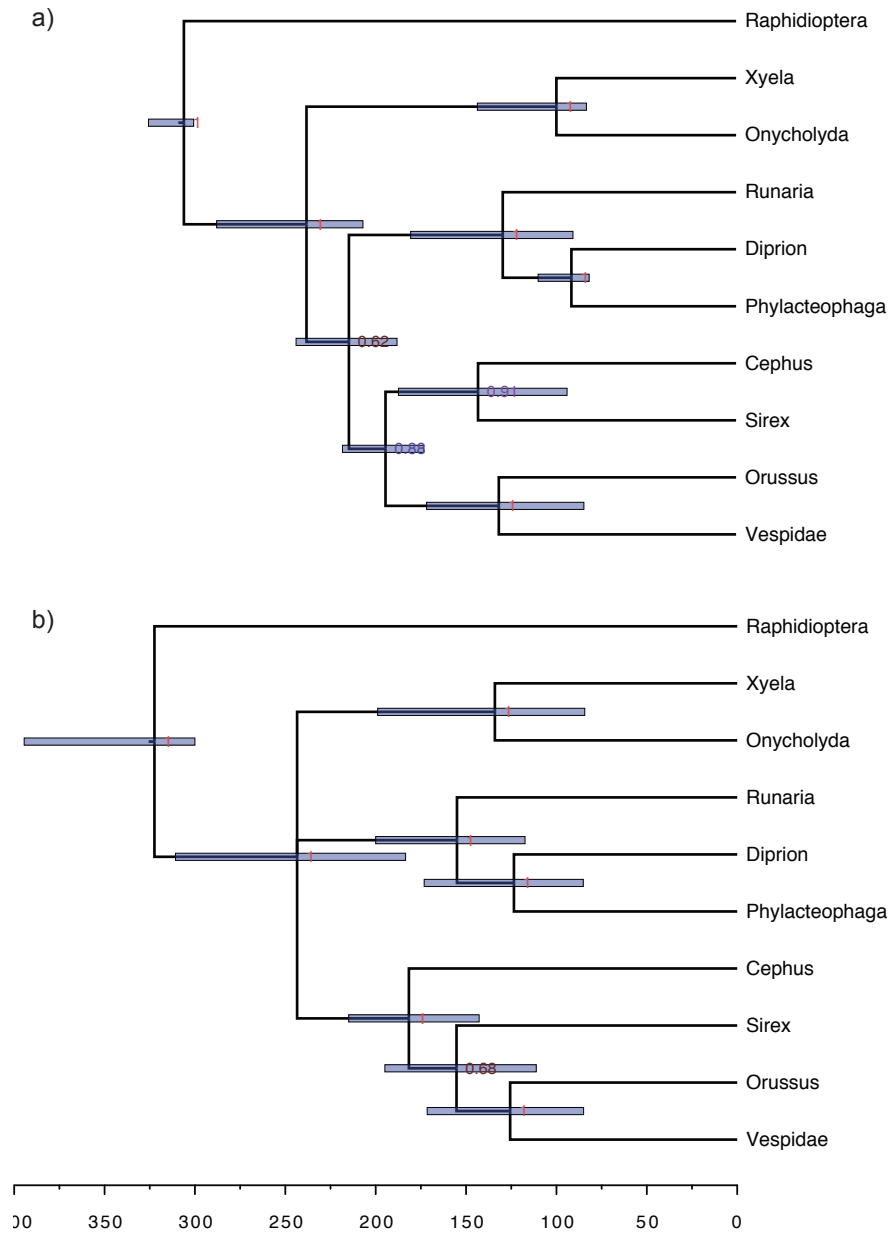
Figure 8: Majority-rule consensus trees from a) total-evidence dating and b) node dating, under diversified FBD and IGR model.

Table 1: Ages (Ma) of root and Hymenoptera (median and 95% HPD) from total-evidence dating and node dating under diversified FBD prior.

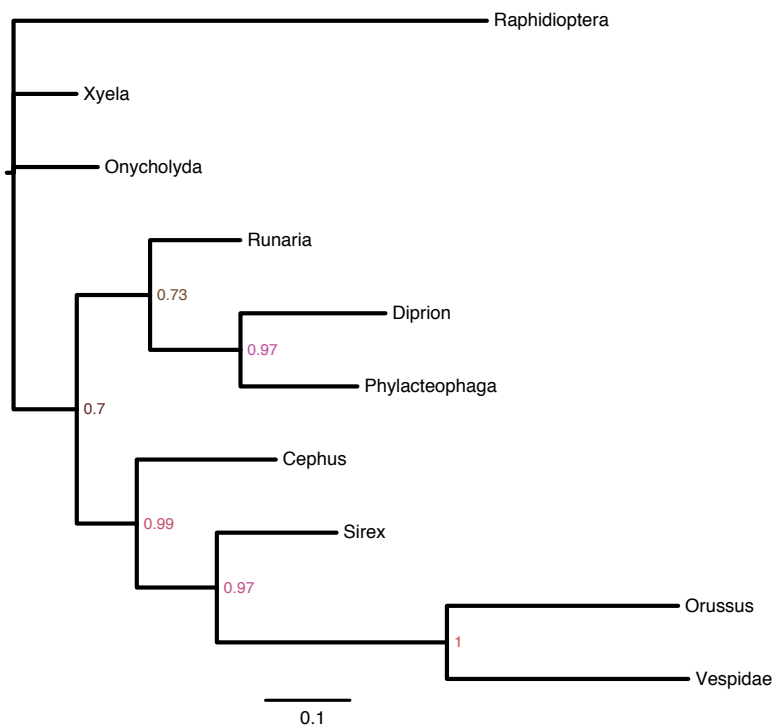|	Root	Hymenoptera
IGR		
TE dating	305.4 (300.0, 325.0)	237.6 (206.4, 287.3)
Node dating	322.4 (300.0, 394.5)	243.5 (183.4, 310.7)
TK02		
TE dating	306.2 (300.0, 327.7)	289.0 (257.6, 314.0)
Node dating	325.0 (300.0, 404.0)	236.6 (185.9, 298.6)

Figure 9: Majority-rule consensus tree from non-clock analysis under the gamma-Dirichlet prior.

using the whole data, please see Ronquist et al. (2012a); Zhang et al. (2016).

4 Acknowledgements

I thank Johan Nylander for valuable discussions and for organizing a workshop of MrBayes using this tutorial.

References

Altekar, G., S. Dwarkadas, J. P. Huelsenbeck, and F. Ronquist. 2004. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics 20:407–415.

Ayres, D. L., A. Darling, D. J. Zwickl, P. Beerli, M. T. Holder, P. O. Lewis, J. P. Huelsenbeck, F. Ronquist, D. L. Swofford, M. P. Cummings, A. Rambaut, and M. A. Suchard. 2012. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. Systematic Biology 61:170–173.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. PLoS biology 4:e88.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17:368–376.

Gavryushkina, A., D. Welch, T. Stadler, and A. J. Drummond. 2014. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. PLoS Computational Biology 10:e1003919.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

Heath, T. A., J. P. Huelsenbeck, and T. Stadler. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. Proceedings of the National Academy of Sciences of the United States of America 111:E2957–66.

Hedges, S. B. and S. Kumar. 2004. Precision of molecular time estimates. Trends in Genetics 20:242–247.

Ho, S. Y. W. and M. J. Phillips. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58:367–380.

Huelsenbeck, J. P., B. Larget, and D. L. Swofford. 2000. A compound poisson process for relaxing the molecular clock. Genetics 154:1879–1892.

Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Lakner, C., P. van der Mark, J. P. Huelsenbeck, B. Larget, and F. Ronquist. 2008. Efficiency of Markov chain Monte Carlo tree proposals in Bayesian phylogenetics. Systematic Biology 57:86–103.

Lepage, T., D. Bryant, H. Philippe, and N. Lartillot. 2007. A general comparison of relaxed molecular clock models. Molecular Biology and Evolution 24:2669–2680.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50:913–925.

Liu, L. and D. K. Pearl. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Systematic Biology 56:504–514.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. Journal of Chemical Physics 21:1087–1092.

Rannala, B., T. Zhu, and Z. Yang. 2012. Tail paradox, partial identifiability, and influential priors in Bayesian branch length inference. Molecular Biology and Evolution 29:325–335.

Ronquist, F. and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Ronquist, F., S. Klopfstein, L. Vilhelmsen, S. Schulmeister, D. L. Murray, and A. P. Rasnitsyn. 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Systematic Biology 61:973–999.

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61:539–542.

Stadler, T. 2010. Sampling-through-time in birth-death trees. Journal of theoretical biology 267:396–404.

Thorne, J. L. and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. Systematic Biology 51:689–702.

Xie, W., P. O. Lewis, Y. Fan, L. Kuo, and M.-H. Chen. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Systematic Biology 60:150–160.

Yang, Z. 1994a. Estimating the pattern of nucleotide substitution. Journal of Molecular Evolution 39:105–111.

Yang, Z. 1994b. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. Journal of Molecular Evolution 39:306–314.

Yang, Z. and B. Rannala. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Molecular Biology and Evolution 23:212–226.

Zhang, C., B. Rannala, and Z. Yang. 2012. Robustness of compound Dirichlet priors for Bayesian inference of branch lengths. Systematic Biology 61:779–784.

Zhang, C., T. Stadler, S. Klopfstein, T. A. Heath, and F. Ronquist. 2016. Total-evidence dating under the fossilized birth-death process. Systematic Biology 65:228–249.

Zuckerkandl, E. and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. Evolving genes and proteins 97:97–166.