

## Modified MrBayes 3.2.1 with compound Dirichlet priors

Chi Zhang, 9/20/2012

### Introduction

This is a modified version of MrBayes 3.2.1, implementing the compound Dirichlet priors for branch lengths described by Rannala et al. (2012) and Zhang et al. (2012). Compared with the exponential and uniform priors for branch lengths in the standard versions of MrBayes, the Dirichlet priors appear more reasonable and may avoid the problem of extremely long trees, as discussed by Brown et al. (2010) and Marshall (2010). The two-exponential prior on internal and external branch lengths described by Yang & Rannala (2005) and Yang (2007) are also implemented in this modified version.

If you use this version of mrbayes, you should cite the original MrBayes papers besides Rannala et al. (2012) and Zhang et al. (2012).

Windows and MAC osx executables are included in this archive (mb3.2.Dir.osx and mb3.2.Dir.exe). To compile for linux, follow the instructions in the original package (CompileInstructions.txt).

I have made changes to 5 source files: bayes.c, command.c, mb.h, model.c, and mcmc.c. You can search for “//chi” in the source code to find my modifications.

I suggest that before analyzing your own data, you run the program without data (mcmc data=no) to confirm that the posterior estimates are consistent with the priors. Also if you analyze the lizards data set (Sceloporus.nex) of Leaché and Mulcahy (2007), you should be able to duplicate the results of table 3 in Zhang et al. (2012). Under GTR+ $\Gamma_4$ , the posterior mean and 95% CI of the tree length is 1.824 (1.689, 1.958) for gammadir(1,1,1,1), and is 1.821 (1.688, 1.954) for invgamdir(3,2,1,1).

If you have any problem with the modified version, please contact Chi Zhang <zhangchicool@gmail.com>.

### Commands for compound Dirichlet priors

#### *Gamma prior on the tree length*

```
prset brlenspr=unconstrained:gammadir( $\alpha_T, \beta_T, \alpha, c$ );
```

$\alpha_T$  and  $\beta_T$  are the shape and rate parameters of the gamma distribution on the tree length. We use  $\alpha_T=1$  by default, while the prior mean of the tree length is  $\alpha_T / \beta_T$ .  $\alpha$  and  $c$  are the parameters of the Dirichlet prior on the branch lengths.  $\alpha = c = 1$  specifies the uniform Dirichlet distribution.

#### *Inverse gamma prior on the tree length*

```
prset brlenspr=unconstrained:invgamdir( $\alpha_T, \beta_T, \alpha, c$ );
```

This prior is heavier-tailed than gamma. We use  $\alpha_T = 3$  by default, while the prior mean of the tree length is  $\beta_T / (\alpha_T - 1)$ .

### Commands for two-exponential priors

```
prset brlenspr=unconstrained:twoexp( $r_I, r_E$ );
```

$r_I$  and  $r_E$  are parameters for internal branches and external branches respectively. We use  $r_I = 100$ ,  $r_E = 10$  by default, while  $r_I = r_E = 10$  is equivalent to the default prior exponential(10).

The three commands above are newly introduced in this version. They all can be used together with other commands in prset. For example,

```
prset statefreqpr=fixed(equal) brlenpr=uncon:gamma(1,.5,.8,1);
```

## References

- Brown, J. M., S. M. Hedtke, A. R. Lemmon, and E. M. Lemmon. 2010. When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. *Syst. Biol.* 59:145-161.
- Leaché, A. D., and D. G. Mulcahy. 2007. Phylogeny, divergence times and species limits of spiny lizards (*Sceloporus magister* species group) in western North American deserts and Baja California. *Mol. Ecol.* 16:5216-5233.
- Marshall, D. C. 2010. Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. *Syst. Biol.* 59:108-117.
- Rannala, B., T. Zhu, and Z. Yang. 2012. Tail paradox, partial identifiability and influential priors in Bayesian branch length inference. *Mol. Biol. Evol.* 29:325-335.
- Yang, Z. 2007. Fair-balance paradox, star-tree paradox and Bayesian phylogenetics. *Mol. Biol. Evol.* 24:1639-1655.
- Yang, Z., and B. Rannala. 2005. Branch-length prior influences Bayesian posterior probability of phylogeny. *Syst. Biol.* 54:455-470.
- Zhang, C., B. Rannala, and Z. Yang. 2012. Robustness of compound Dirichlet priors for Bayesian inference of branch lengths. *Syst. Biol.* 61:779-784.