

PHYLOBAYES 3. SUPPLEMENTARY INFORMATION

NICOLAS LARTILLOT, THOMAS LEPAGE AND SAMUEL BLANQUART

1. CONSENSUS TREES

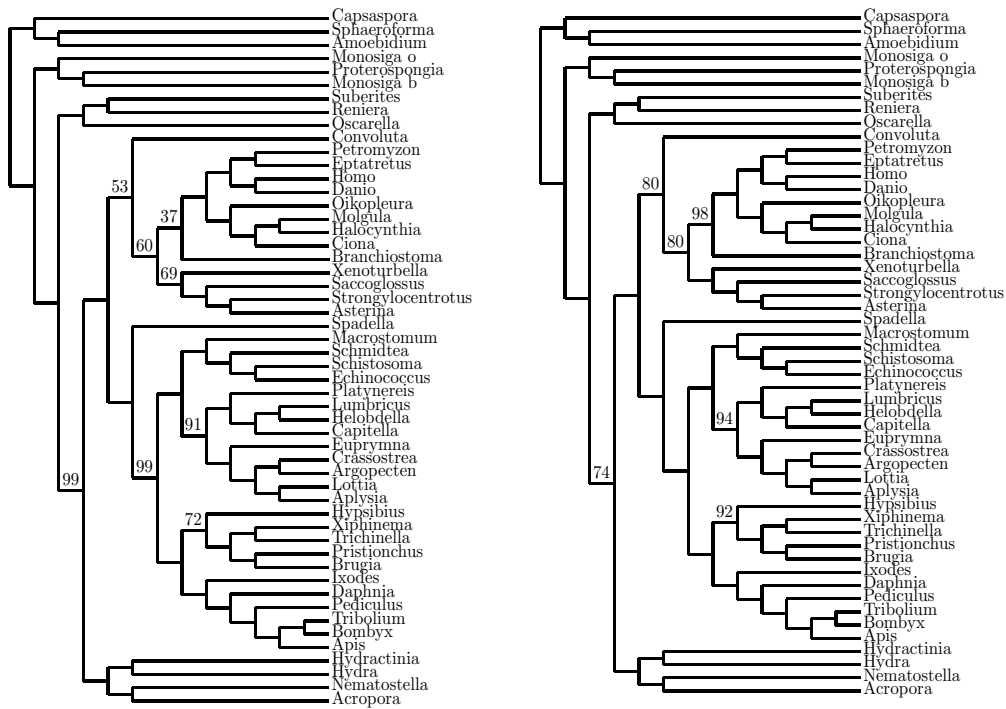


FIGURE 1. Consensus trees obtained under a Dirichlet process mixture of profiles combined with flat (CAT model, left) or free (CAT-GTR model) relative exchangeabilities, and gamma distributed rates across sites. Support values are $100 \times$ posterior probabilities obtained by merging 2 independent MCMC runs.

2. MOLECULAR DATING ANALYSES

Date: june, 8, 2009.

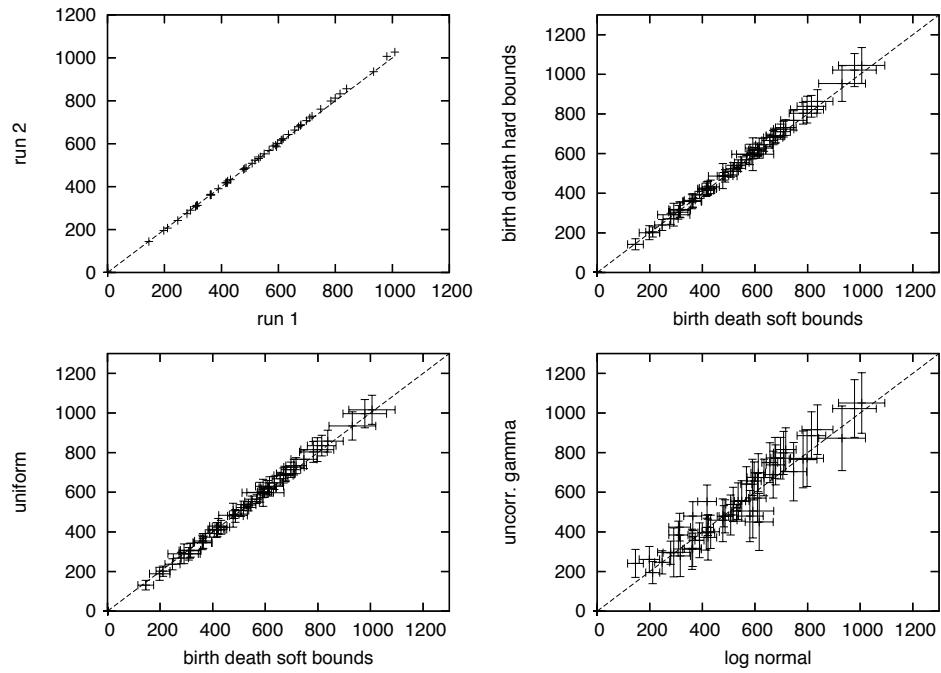


FIGURE 2. Comparisons of node age estimates obtained under two independent runs (top left), soft versus hard constraints (top right), birth death versus uniform prior (bottom left), and log-normal versus uncorrelated gamma models of clock relaxation (bottom right). All chains were run under the CAT model.

3. CHECKS OF THE IMPLEMENTATION

TABLE 1. Comparison of MULTIDIVTIME and PHYLOBAYES under the prior (see Table 1, Inoue, Donoghue and Yang, 2009, Systematic Biology, in press.)

cases	MULTIDIVTIME			PHYLOBAYES		
	t_1	t_2	t_3	t_1	t_2	t_3
1	1.185	0.894	0.449	1.186	0.897	0.450
2	1.240	0.954	0.627	1.242	0.954	0.628
3	0.758	0.393	0.150	0.757	0.394	0.150
4	0.711	0.313	0.156	0.712	0.312	0.156
5	0.782	0.429	0.364	0.781	0.428	0.364
6	0.782	0.429	0.364	0.782	0.429	0.364
7	0.761	0.402	0.201	0.764	0.402	0.201

TABLE 2. Comparison of MCMCTREE and PHYLOBAYES under the prior and using hard bounds (see Table 1, Inoue et al, 2009)

cases	MCMCTREE			PHYLOBAYES		
	t_1	t_2	t_3	t_1	t_2	t_3
1 improper prior	1.374	0.936	0.468	1.375	0.937	0.468
1 $L(t_l, 0.1, 2)$	1.309	0.856	0.427	1.310	0.858	0.429
1 $L(t_l, 0.1, 1)$	1.256	0.787	0.394	1.254	0.786	0.391
1 $L(t_l, 0.1, 0.5)$	1.202	0.711	0.356	1.201	0.712	0.355
1 $L(t_l, 0.1, 0.2)$	1.147	0.630	0.315	1.144	0.629	0.315
2 improper prior	1.548	1.153	0.726	1.548	1.152	0.727
2 $L(t_l, 0.1, 2)$	1.382	0.954	0.573	1.383	0.953	0.573
2 $L(t_l, 0.1, 1)$	1.290	0.836	0.496	1.290	0.835	0.495
2 $L(t_l, 0.1, 0.5)$	1.215	0.731	0.432	1.214	0.731	0.434
2 $L(t_l, 0.1, 0.2)$	1.150	0.633	0.378	1.147	0.634	0.378
3 improper prior	1.040	0.398	0.150	1.038	0.398	0.150
4 improper prior	1.017	0.246	0.123	1.016	0.246	0.123
5 improper prior	1.048	0.427	0.356	1.050	0.432	0.366
5 $L(t_l, 0.1, 1)$	1.050	0.431	0.364	1.051	0.431	0.364
6 improper prior	1.050	0.432	0.366	1.050	0.432	0.366
7 improper prior	1.040	0.398	0.199	1.041	0.398	0.199

TABLE 3. Comparison of MCMCTREE and PHYLOBAYES under the prior, using soft bounds (same conditions as in Table 1 of Inoue et al 2009, but run under soft constraints). Top: birth death with $\lambda = \mu = 1$ and $\rho = 0$. Bottom, $\lambda = 1$, $\mu = 0.5$ and $\rho = 0.5$.

cases	MCMCTREE			PHYLOBAYES		
	t_1	t_2	t_3	t_1	t_2	t_3
1	1.243	0.770	0.385	1.246	0.770	0.386
2	1.282	0.822	0.485	1.282	0.823	0.485
3	1.040	0.398	0.150	1.039	0.399	0.150
4	1.018	0.250	0.126	1.020	0.246	0.123
5	1.049	0.435	0.358	1.049	0.426	0.354
6	1.049	0.439	0.366	1.049	0.432	0.366
7	1.038	0.398	0.200	1.039	0.398	0.199
1	1.243	0.771	0.365	1.244	0.771	0.358
2	1.283	0.823	0.484	1.283	0.823	0.484
3	1.040	0.398	0.150	1.042	0.399	0.150

TABLE 4. Comparison of MCMCTREE and PHYLOBAYES under the posterior, using soft bounds. Primate example of MCMCTREE. Top, using one upper and one lower calibration on 2 separate nodes (human-chimp clade older than 0.6, ingroup younger than 1.6). Bottom: imposing an upper and a lower bound on 2 nodes (human-chimp clade between 0.6 and 0.8, and all ingroup between 1.2 and 1.6).

clade		MCMCTREE		PHYLOBAYES	
		mean	0.95 CI	mean	0.95 CI
gibbon	bonobo	3.237	[1.065 : 8.690]	3.190	[1.075 : 8.547]
sumatran	bonobo	1.218	[0.701 : 1.629]	1.197	[0.698 : 1.583]
gorilla	bonobo	1.034	[0.636 : 1.517]	1.022	[0.633 : 1.482]
human	bonobo	0.852	[0.575 : 1.371]	0.845	[0.570 : 1.335]
chimp	bonobo	0.422	[0.021 : 1.028]	0.425	[0.020 : 0.018]
sumatran	orang outang	0.737	[0.057 : 1.462]	0.724	[0.042 : 1.429]
gibbon	bonobo	1.627	[1.313 : 1,993]	1.717	[1.389 : 2.391]
sumatran	bonobo	1.497	[1.226 : 1.684]	1.499	[1.302 : 1.597]
gorilla	bonobo	0.917	[0.756 : 1.101]	0.928	[0.816 : 1.058]
human	bonobo	0.635	[0.539 : 0.768]	0.647	[0.602 : 0.749]
chimp	bonobo	0.292	[0.209 : 0.394]	0.299	[0.225 : 0.393]
sumatran	orang outang	0.527	[0.375 : 0.695]	0.528	[0.400 : 0.692]

TABLE 5. Comparison of MULTIDIVTIME and PHYLOBAYES under the posterior, using the Gene 1 example of multidivtime.

clade		MULTIDIVTIME		PHYLOBAYES	
		mean	0.95 CI	mean	0.95 CI
Taxon E	Taxon D	104.8	[100.2:109.7]	105.0	[100.3 : 109.7]
Taxon E	Taxon C	128.6	[106.7:168.9]	132.8	[107.9 : 173.3]
Taxon E	Taxon A	159.4	[121.7:235.6]	163.6	[123.7 : 234.8]

TABLE 6. Check of the CAT model, based on an evaluation of the hit probabilities. $K = 30$ simulations were performed, based on 30 parameter values drawn from the prior, and the fraction of bipartitions supported at $x \pm 5\%$ that are true (fraction expected to be approx. equal to x) is recorded, as a function of x .

x	obs. fraction.	abs. nb	total nb	expected tolerance
5	5	20	351	[12,23]
15	13	14	104	[11,21]
25	17	14	81	[15,26]
35	33	15	45	[12,20]
45	46	13	28	[9,16]
55	57	12	21	[9,14]
65	57	11	19	[10,15]
75	83	15	18	[11,15]
85	75	12	16	[12,15]
95	96	62	64	[59,62]