Supplementary figures and tables for StarBEAST2

Huw A. Ogilvie[*,1,2] and Alexei J. Drummond[2,3]

[1]Division of Evolution, Ecology and Genetics, Research School of Biology, Australian National University,

Canberra, Australia

[2]Centre for Computational Evolution, University of Auckland, Auckland, New Zealand

[3]Department of Computer Science, University of Auckland, Auckland, New Zealand

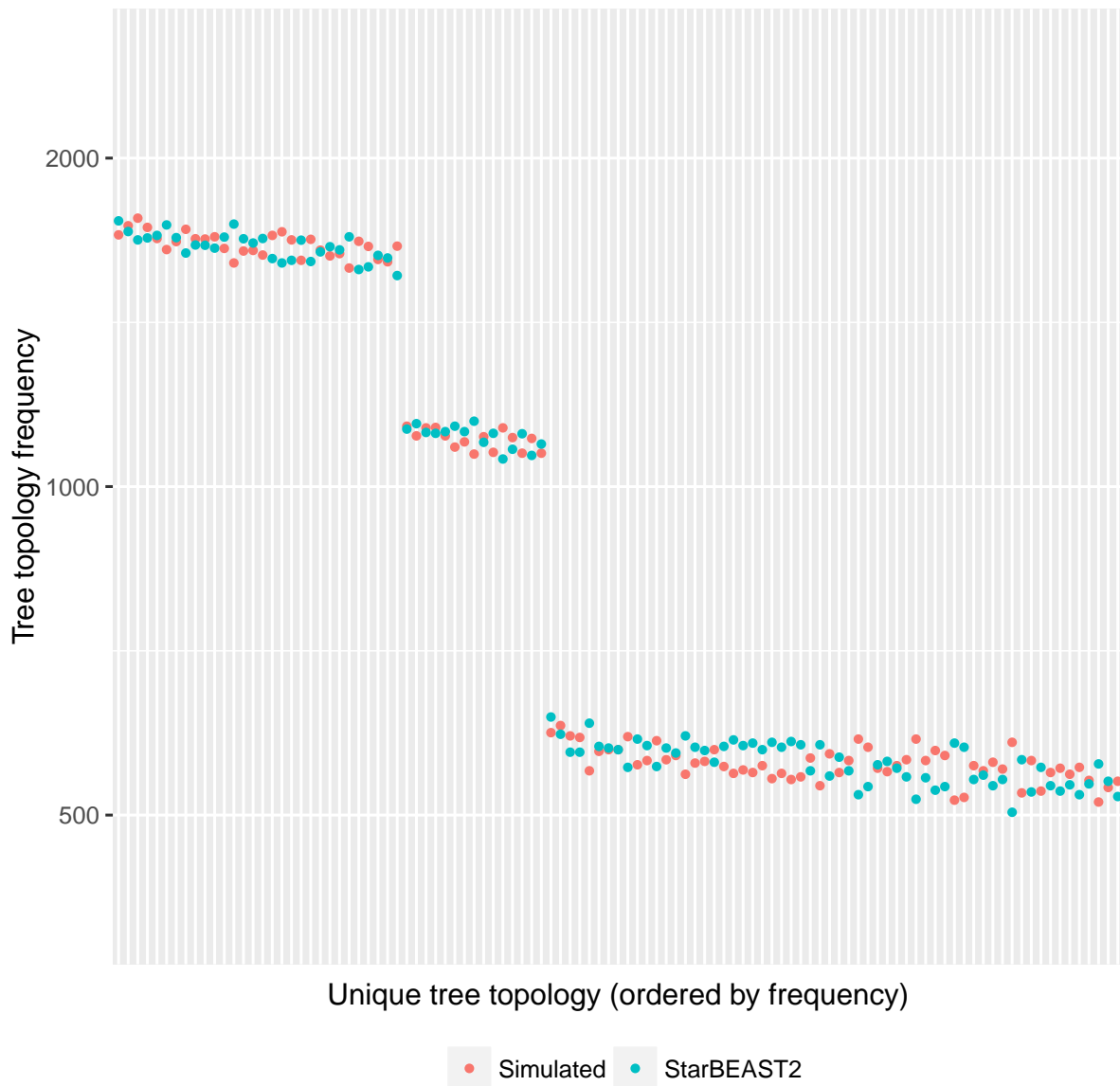June 23, 2016

[*]huw.ogilvie@anu.edu.au

Figure 1: Frequency of five-taxon species tree topologies sampled from a birth-death prior distribution. Topologies were sampled by simulating trees using biopy (red), or by using a StarBEAST2 MCMC chain (blue). Frequencies are identical apart from noise, indicating that StarBEAST2 is mathematically correct. Three levels of probability are evident from left to right — high probability balanced topologies e.g. ((a,b),((c,d),e)), middle probability intermediate topologies e.g. (((a,b),(c,d)),e), and low probability unbalanced topologies e.g. ((((a,b),c),d),e).
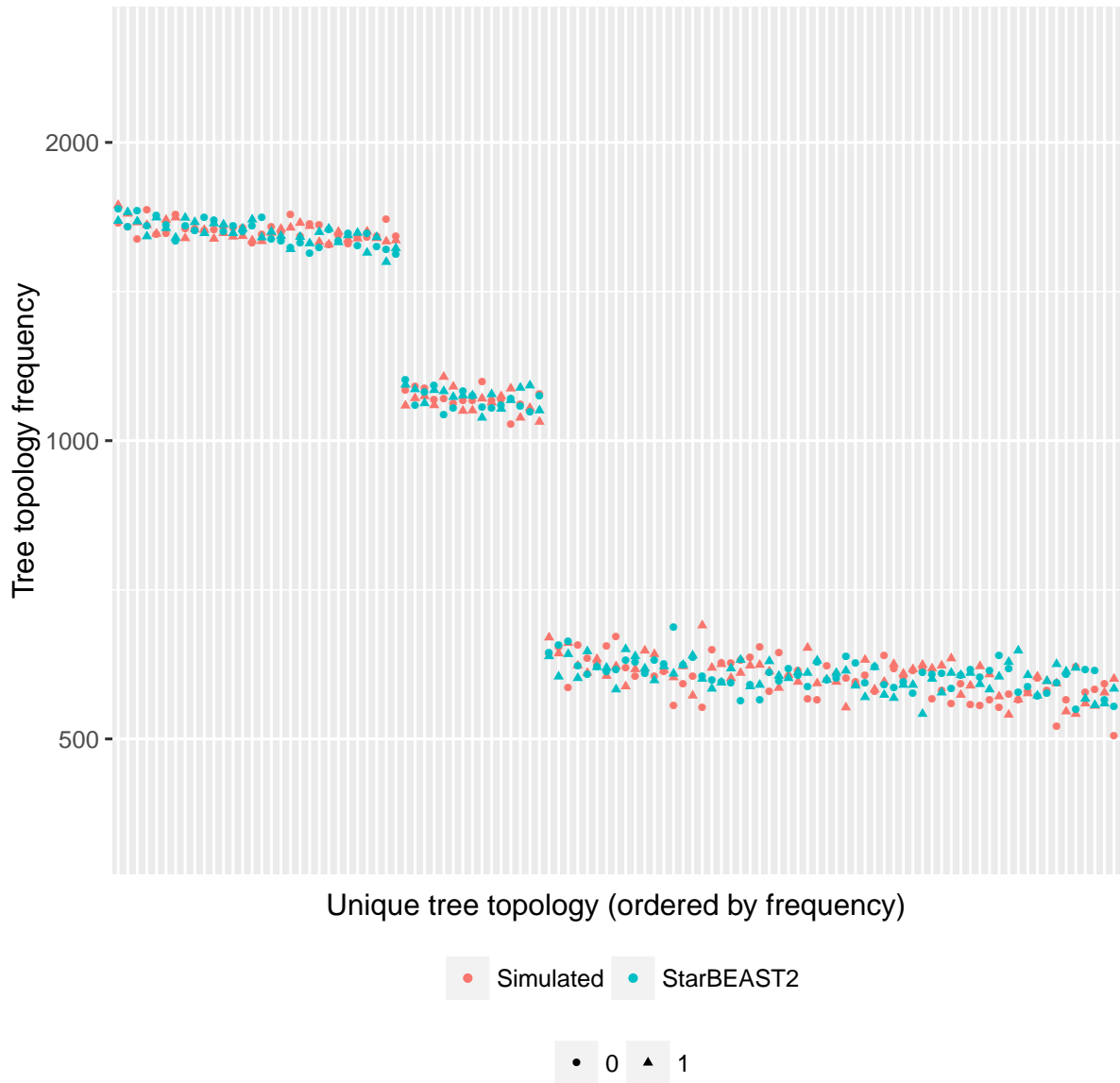
Figure 2: Frequency of five-taxon gene tree topologies sampled from a multispecies coalescent prior distribution. Topologies were sampled by simulating gene trees within species trees using biopy (red), or by using a StarBEAST2 MCMC chain (blue). Gene trees were sampled with two clock rates, 0.5 (circles) and 2.0 (triangles), although clock rate should not affect topology or node heights in units of time. Frequencies are identical apart from noise, indicating that StarBEAST2 is mathematically correct. Three levels of probability are evident from left to right — high probability balanced topologies e.g. ((a,b),((c,d),e)), middle probability intermediate topologies e.g. (((a,b),(c,d)),e), and low probability unbalanced topologies e.g. ((((a,b),c),d),e).

Figure 3: Probability densities of five-taxon species tree node heights sampled from a birth-death prior distribution. Node heights were sampled by simulating trees using biopy (red), or by using a StarBEAST2 MCMC chain (blue). Probability densities are plotted separately for each ranked node giving four peaks, one for each internal node. Node height probability densities are identical apart from noise, indicating that StarBEAST2 is mathematically correct.
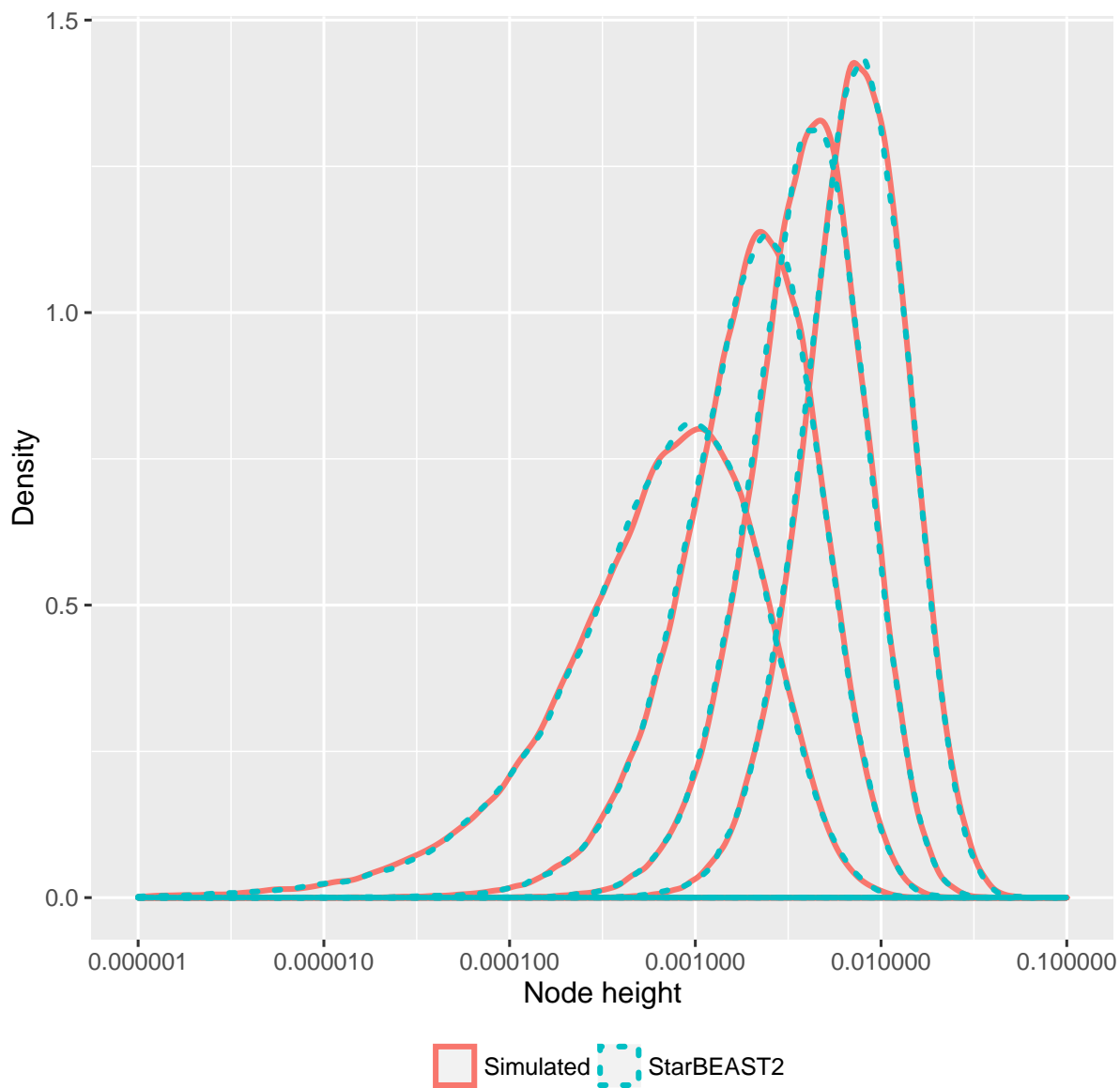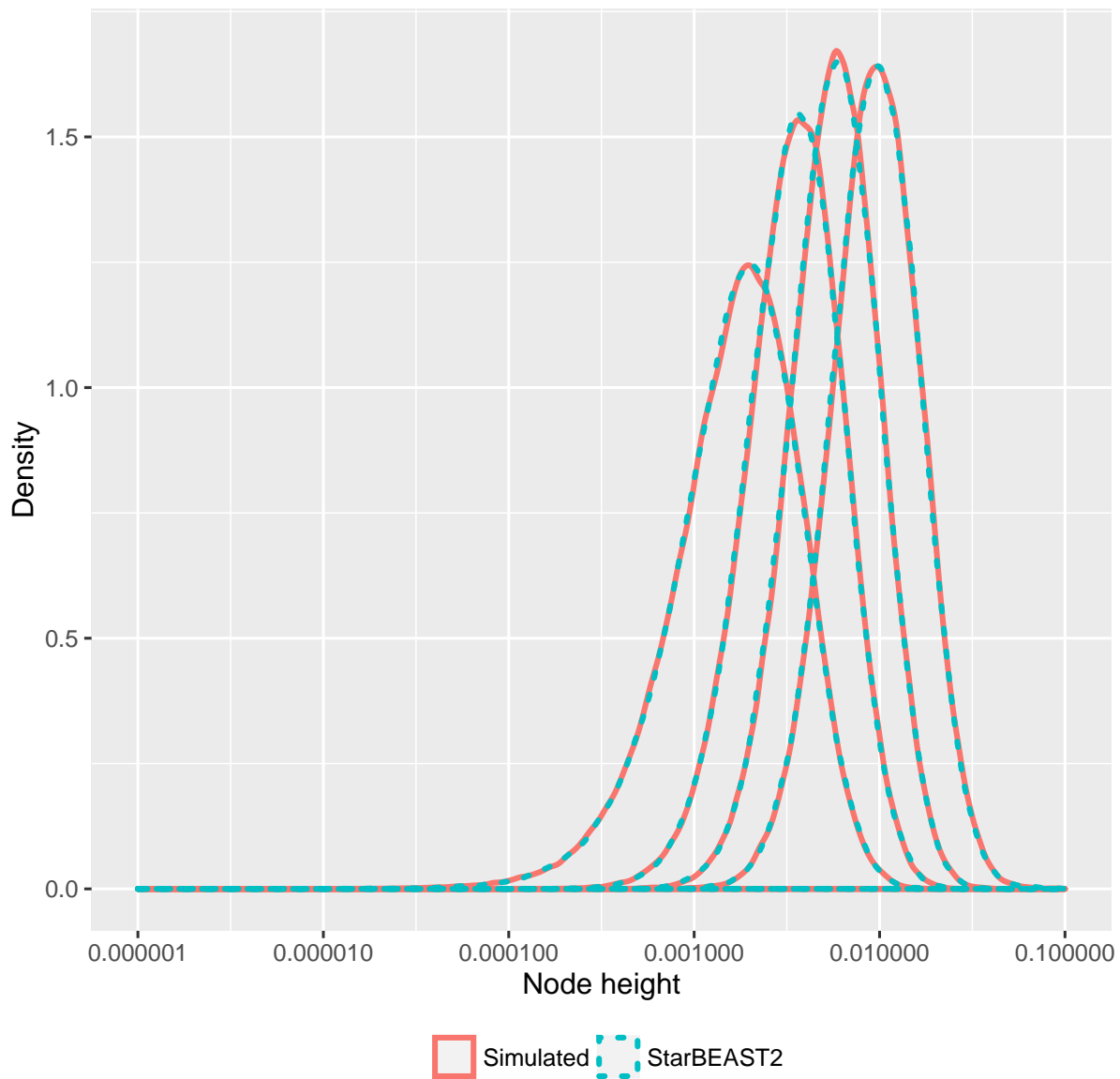
Figure 4: Probability densities of five-taxon gene tree node heights sampled from a multispecies coalescent prior distribution. Node heights were sampled by simulating gene trees within species trees using biopy (red), or by using a StarBEAST2 MCMC chain (blue). Gene trees were sampled with two clock rates, 0.5 and 2.0, and node heights from both sets of gene trees were combined as clock rate should not affect topology or node heights in units of time. Probability densities are plotted separately for each ranked node giving four peaks, one for each internal node. Node height probability densities are identical apart from noise, indicating that StarBEAST2 is mathematically correct.

Figure 5: Probability densities of five-taxon gene tree branch rates produced by a species tree uncorrelated lognormal (UCLN) relaxed clock. Branch rates were sampled by simulating gene trees within species trees using biopy and simulating species tree branch rates using scipy (red), or by using a Star-BEAST2 MCMC chain (blue). Gene trees were sampled with two clock rates, 0.5 and 2.0, resulting in two peaks. Branch rate probability densities are identical apart from noise, indicating that StarBEAST2 is mathematically correct.
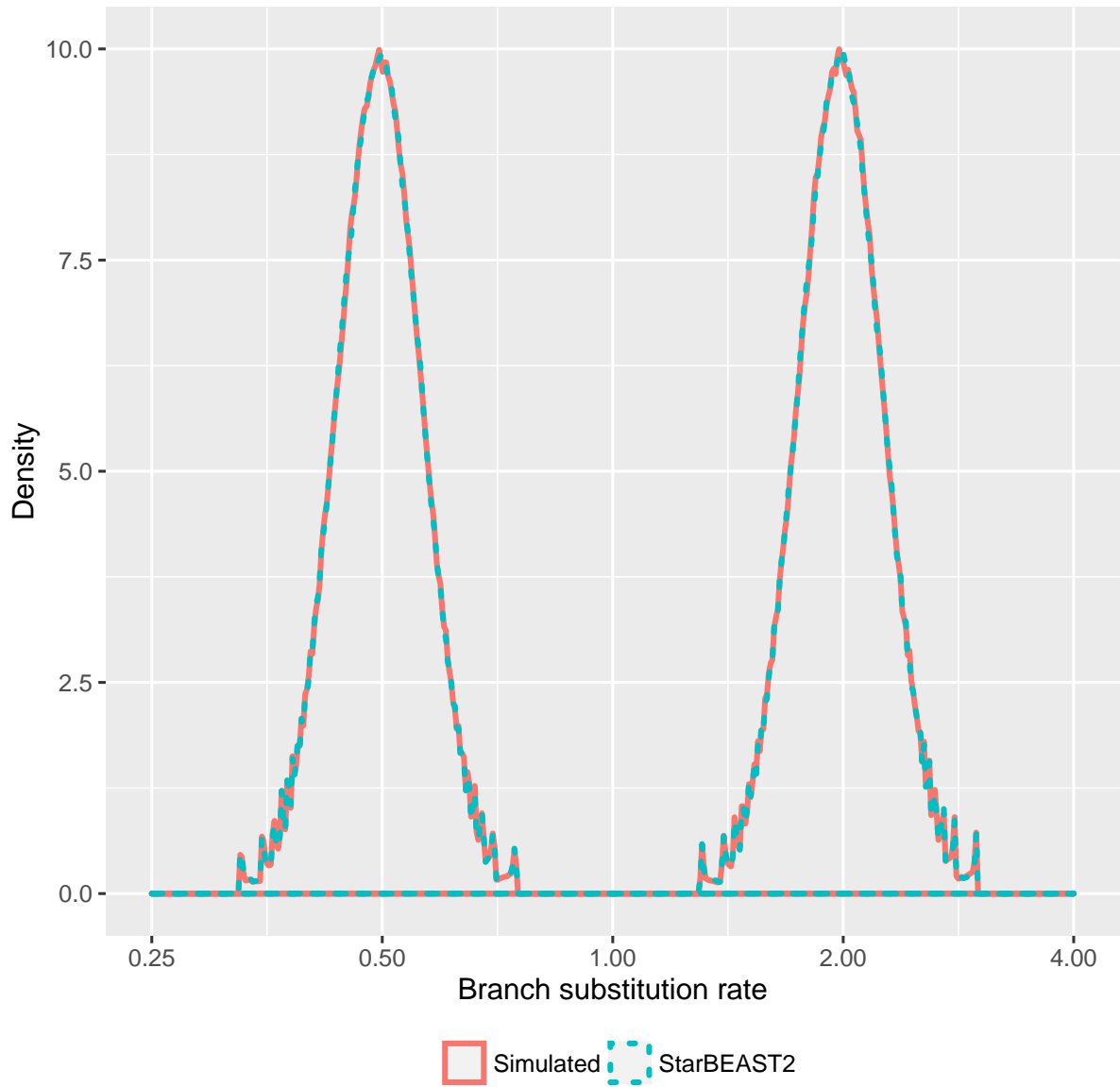
Figure 6: Probability densities of five-taxon gene tree branch rates produced by a species tree uncorrelated exponential (UCED) relaxed clock. Branch rates were sampled by simulating gene trees within species trees using biopy and simulating species tree branch rates using scipy (red), or by using a StarBEAST2 MCMC chain (blue). Gene trees were sampled with two clock rates, 0.5 and 2.0, resulting in two peaks. Branch rate probability densities are identical apart from noise, indicating that StarBEAST2 is mathematically correct.
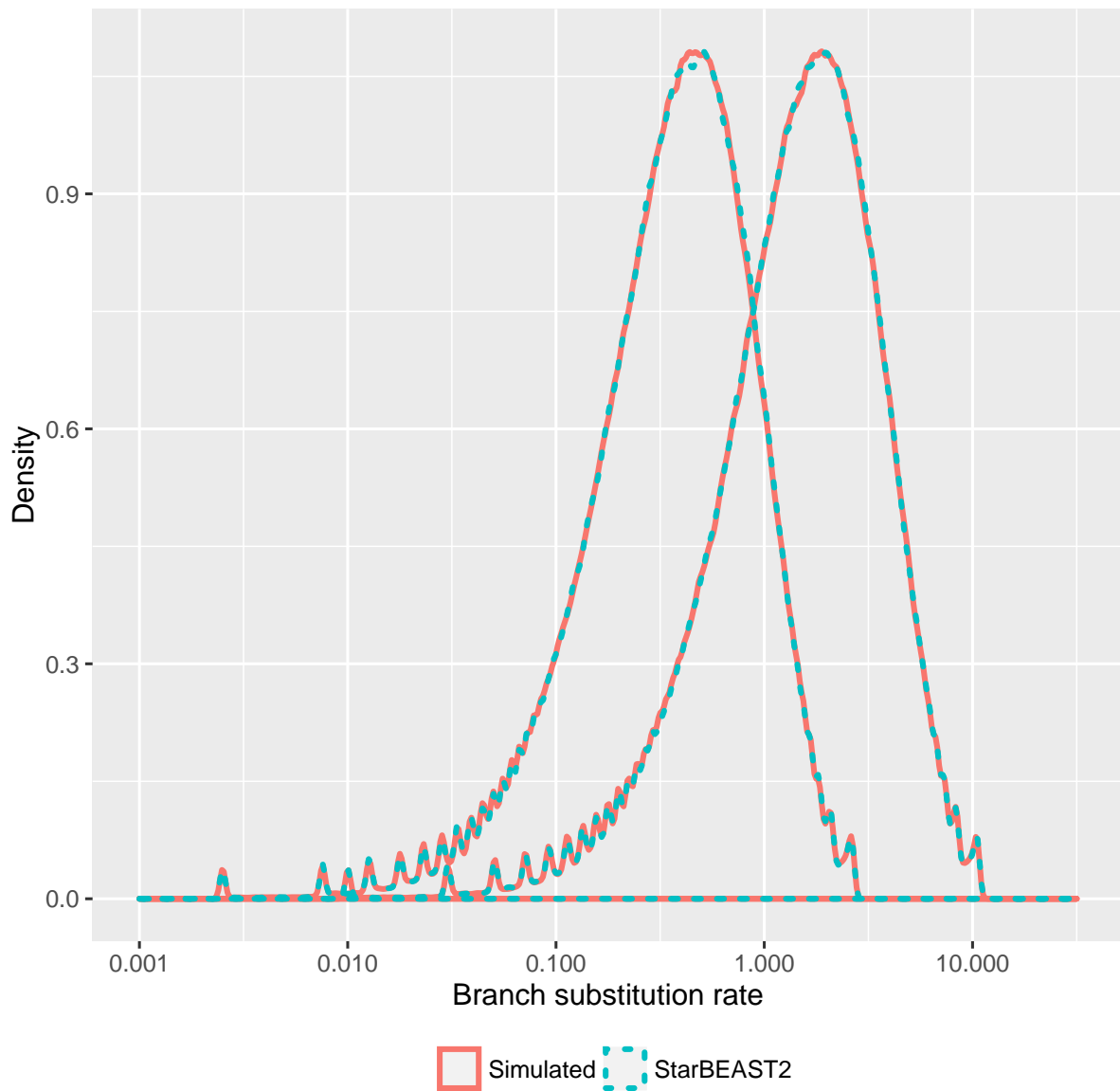
Table 1: Average ESS per hour convergence for *Pseudacris* data.

	Birth-death probability	Clock rate stdev	Extinction fraction	Among-site rate variation	Transition/ Transversion / κ	Phylogenetic likelihood	Diversification rate	Population Mean	Log-posterior	Log-prior	Species tree height	Species tree length	Coalescent probability
◇◇◇◇	114.8 99.1−135.1	241.3 210.8−249.8	221.5 197.2−230.9	308.9 263.4−329.0	309.4 260.2−333.1	21.9 18.0−26.4	218.7 192.5−235.1	30.2 26.2−34.2	26.2 22.8−29.7	30.1 27.8−32.7	9.3 7.1−11.8	84.1 74.2−95.9	19.9 16.6−24.4
◇◇◇◆	133.1 124.0−139.5	281.4 251.7−314.6	226.0 210.3−234.0	481.1 428.1−560.0	524.3 479.3−633.0	28.0 23.1−34.4	227.7 208.2−244.9	38.3 35.5−42.5	31.9 28.9−37.1	37.9 34.2−42.7	75.6 69.0−86.9	104.7 97.5−110.8	23.1 21.2−27.8
◇◇◆◇	85.2 73.8−97.5	192.5 170.1−204.4	180.8 162.5−185.7	245.4 205.7−257.7	252.6 208.3−263.5	18.2 16.0−21.0	184.6 164.3−193.4	23.9 20.8−26.1	17.2 15.0−20.8	25.9 23.0−29.1	7.4 5.5−10.3	68.6 60.5−75.9	16.3 14.0−19.4
◇◇◆◆	89.3 80.7−99.2	225.5 201.9−243.1	179.3 167.9−191.0	390.1 300.9−435.4	407.2 316.5−456.1	21.4 18.8−24.9	188.8 177.1−199.2	32.8 26.9−36.5	26.4 24.6−29.0	28.5 24.4−32.4	62.0 58.9−67.3	73.9 67.3−82.2	20.4 19.2−23.0
◇◆◇◇	50.1 38.8−61.4	165.7 159.2−173.6	146.4 137.0−152.1	187.6 180.6−193.7	199.7 189.8−204.2	14.3 10.0−16.7	143.0 133.6−150.3	22.5 19.6−24.7	19.8 16.2−22.3	20.8 18.9−22.7	1.8 1.4−2.3	34.4 23.9−42.2	17.8 14.2−19.4
◇◆◇◆	89.3 82.1−95.6	230.9 217.2−258.1	164.0 157.0−168.3	364.9 307.1−450.8	405.3 342.4−485.8	22.5 17.8−27.2	172.7 167.6−179.7	30.4 27.0−37.3	29.8 27.1−35.7	24.9 23.4−27.0	46.4 42.1−51.4	72.9 66.8−79.2	25.2 21.6−29.4
◇◆◆◇	42.2 32.3−54.6	180.4 138.8−218.5	140.4 124.9−159.5	229.5 170.5−289.8	231.4 162.6−304.3	15.6 12.6−17.6	139.5 127.4−154.0	21.2 17.5−24.5	16.1 13.6−19.2	19.4 17.6−21.0	2.8 1.8−4.0	29.9 19.6−38.2	13.7 11.7−16.2
◇◆◆◆	85.6 78.1−92.1	202.9 190.5−226.7	142.6 131.0−153.1	317.3 262.2−417.1	338.8 303.0−414.7	18.2 14.9−22.5	151.5 141.6−159.7	27.3 24.6−29.5	24.8 22.3−28.9	24.4 22.4−27.0	40.2 35.1−46.5	67.0 63.0−74.5	19.0 16.9−21.5
◆◇◇◇	125.2 120.7−131.9	233.4 218.0−241.3	221.6 209.4−231.0	295.4 280.6−309.0	315.5 303.1−322.2	23.5 19.5−28.7	230.8 221.7−234.4	51.6 45.6−57.3	36.4 33.0−41.5	184.6 166.3−202.8	11.2 8.6−13.9	108.3 99.4−112.9	33.9 30.9−37.2
◆◇◇◆	144.5 124.8−157.6	324.2 299.2−354.0	224.8 212.4−240.4	678.8 578.4−756.1	746.5 637.7−819.1	45.4 34.3−57.1	243.2 233.2−256.6	79.9 71.6−88.6	53.8 48.9−61.6	239.8 221.2−261.3	104.3 93.7−115.5	116.5 99.9−130.8	45.8 40.3−51.0
◆◇◆◇	91.9 84.0−98.5	191.4 177.5−198.0	169.0 160.2−176.5	235.3 216.0−246.2	249.4 229.7−255.1	20.6 17.3−25.8	173.9 161.8−183.0	44.8 39.2−51.9	32.8 27.8−37.2	155.8 140.7−167.7	8.7 6.2−11.5	75.4 69.3−83.6	26.1 23.0−31.1
◆◇◆◆	113.4 107.1−123.2	262.1 246.7−279.2	199.5 183.2−210.3	582.7 467.5−667.5	624.9 480.9−732.3	33.6 30.1−36.6	216.2 203.7−227.3	70.0 64.1−74.8	47.6 43.2−51.8	200.6 185.6−216.0	84.1 78.7−89.4	95.5 89.8−103.6	37.0 33.2−41.6
◆◆◇◇	73.7 68.7−78.1	169.4 161.1−175.4	145.5 137.5−152.2	178.7 171.3−188.8	197.6 189.7−202.8	14.8 10.1−19.8	148.7 141.8−156.1	41.7 31.6−52.3	32.6 25.3−37.5	121.6 113.2−130.0	3.6 2.8−4.6	50.4 35.4−60.2	27.1 22.5−32.5
◆◆◇◆	95.3 86.2−101.5	233.6 219.0−260.9	160.3 147.7−172.8	454.3 384.5−516.4	496.8 391.0−575.7	26.3 24.1−29.8	170.0 159.4−180.8	65.0 56.7−73.0	46.1 41.3−50.4	205.5 185.9−226.0	55.2 49.6−60.2	77.4 70.1−83.1	35.7 32.7−40.0
◆◆◆◇	61.6 50.5−75.8	158.2 130.1−197.4	130.9 119.2−147.1	210.6 167.8−266.1	228.3 163.2−294.5	15.6 11.3−20.9	132.4 113.9−152.3	38.0 30.8−46.3	27.6 22.5−31.9	118.6 110.0−132.9	3.1 2.3−4.0	40.0 30.4−55.0	24.6 20.5−27.6
◆◆◆◆	80.6 71.2−88.9	231.7 196.5−257.2	146.0 137.1−157.1	398.2 339.0−432.3	424.1 345.5−487.7	22.7 20.3−25.4	153.8 141.1−166.3	54.8 49.3−62.4	40.6 35.1−44.9	181.2 166.0−192.8	49.5 44.3−53.9	69.1 59.9−76.4	31.3 27.3−35.7

Top numbers show trimmed means (25% trim was applied to reduce the influence of outliers)

Bottom numbers show 95% confidence intervals for the trimmed mean calculated by bootstrapping

Filled diamonds indicate which new features are enabled. From left-to-right; analytical integration of population sizes, species tree relaxed clocks, coordinated topology operators, and coordinated height operators.

Table 2: Average ESS per million states convergence for *Pseudacris* data.

	Birth-death probability	Clock rate stdev	Extinction fraction	Among-site rate variation	Transition/ Transversion κ	Phylogenetic likelihood	Diversification rate	Population Mean	Log-posterior	Log-prior	Species tree height	Coalescent probability
◇◇◇◇	7.29 / 6.48−8.12	17.20 / 14.38−20.40	15.75 / 13.51−18.03	19.04 / 14.95−23.81	20.83 / 15.80−26.27	1.72 / 1.54−1.99	16.21 / 13.69−18.64	2.36 / 2.16−2.56	1.94 / 1.74−2.13	2.61 / 2.43−2.81	0.81 / 0.66−0.95	1.68 / 1.52−1.84
◇◇◇◆	12.41 / 11.33−13.43	25.19 / 22.47−28.14	21.38 / 20.11−22.81	37.90 / 31.28−45.25	39.27 / 31.72−46.37	2.79 / 2.43−3.27	22.38 / 21.08−24.03	3.66 / 3.34−3.98	3.08 / 2.85−3.33	3.35 / 3.03−3.69	2.07 / 1.79−2.36	2.52 / 2.31−2.72
◇◇◆◇	7.43 / 6.74−8.08	14.59 / 12.76−17.29	15.00 / 12.99−17.27	17.43 / 14.29−21.63	17.62 / 14.06−22.55	1.96 / 1.75−2.16	14.72 / 12.62−17.34	2.25 / 2.06−2.49	1.86 / 1.64−2.12	2.71 / 2.53−2.92	0.79 / 0.63−0.99	1.63 / 1.47−1.80
◇◇◆◆	10.88 / 10.00−11.74	23.12 / 21.37−25.42	21.06 / 20.56−21.70	31.02 / 27.32−37.84	32.84 / 28.83−40.30	2.83 / 2.60−3.23	21.29 / 20.74−22.48	3.38 / 3.11−3.75	2.80 / 2.54−3.09	3.50 / 3.17−3.83	1.71 / 1.51−1.99	2.33 / 2.13−2.52
◇◆◇◇	4.58 / 3.92−5.34	8.02 / 7.28−9.92	7.94 / 7.15−9.95	8.23 / 7.43−10.22	8.17 / 7.33−10.25	1.72 / 1.58−1.88	8.08 / 7.15−9.96	2.80 / 2.62−2.97	2.34 / 2.12−2.51	2.56 / 2.39−2.73	0.39 / 0.34−0.45	2.11 / 1.96−2.25
◇◆◇◆	9.72 / 8.78−10.61	19.72 / 16.27−23.47	17.74 / 15.03−20.52	22.02 / 17.38−27.09	22.32 / 17.33−27.33	2.94 / 2.71−3.18	17.51 / 15.10−20.24	4.01 / 3.65−4.36	3.65 / 3.42−3.90	3.74 / 3.42−3.99	0.98 / 0.83−1.20	2.89 / 2.71−3.07
◇◆◆◇	5.68 / 4.69−6.74	11.48 / 8.99−13.83	11.11 / 9.00−13.39	11.73 / 9.48−14.42	11.71 / 9.37−14.30	1.82 / 1.61−2.02	10.96 / 8.82−13.40	2.88 / 2.69−3.14	2.31 / 2.12−2.55	2.77 / 2.53−3.04	0.48 / 0.38−0.60	2.09 / 1.93−2.28
◇◆◆◆	9.17 / 8.64−10.07	18.70 / 14.98−23.14	17.02 / 14.64−19.78	20.96 / 15.73−27.94	21.33 / 16.86−28.89	2.96 / 2.71−3.15	16.90 / 14.19−20.14	4.07 / 3.73−4.41	3.82 / 3.59−4.13	3.80 / 3.59−4.00	0.94 / 0.81−1.16	3.15 / 2.95−3.38
◆◇◇◇	9.68 / 8.99−10.49	16.72 / 13.90−19.88	15.99 / 13.97−18.20	18.78 / 15.31−23.23	19.20 / 15.12−23.85	2.17 / 1.76−2.62	16.43 / 14.25−18.73	4.14 / 3.81−4.53	2.80 / 2.57−3.13	13.76 / 12.32−15.66	0.94 / 0.83−1.09	2.69 / 2.54−2.90
◆◇◇◆	13.56 / 12.55−14.39	27.89 / 24.72−31.09	22.93 / 21.70−24.02	46.06 / 37.98−55.69	49.79 / 41.99−60.70	3.53 / 3.18−3.84	23.58 / 21.96−25.50	7.26 / 6.65−7.88	5.06 / 4.60−5.55	21.23 / 19.74−22.85	2.72 / 2.07−3.42	4.08 / 3.70−4.47
◆◇◆◇	9.28 / 8.24−10.24	16.30 / 13.32−19.47	16.08 / 13.54−18.39	19.41 / 14.81−24.11	19.39 / 14.29−25.19	2.20 / 1.87−2.69	15.89 / 13.35−18.66	4.19 / 3.76−4.68	2.82 / 2.49−3.13	13.65 / 11.43−16.08	0.87 / 0.68−1.12	2.51 / 2.23−2.76
◆◇◆◆	13.12 / 12.24−13.95	27.94 / 25.73−29.69	23.19 / 21.87−24.43	48.66 / 41.48−52.03	53.59 / 45.96−56.30	3.73 / 3.37−4.17	24.28 / 23.11−25.51	7.76 / 6.80−8.60	5.24 / 4.75−5.86	21.44 / 20.03−23.17	2.80 / 2.32−3.27	4.39 / 3.98−4.77
◆◆◇◇	5.16 / 4.23−6.40	8.61 / 7.30−11.21	8.47 / 7.29−11.06	8.56 / 7.19−11.53	8.75 / 7.30−11.84	2.12 / 1.82−2.41	8.46 / 7.29−10.60	4.51 / 4.06−5.04	3.43 / 3.12−3.87	7.29 / 6.32−9.72	0.40 / 0.33−0.52	3.15 / 2.84−3.43
◆◆◇◆	11.65 / 10.57−12.74	22.04 / 18.46−26.57	19.77 / 16.91−22.33	25.52 / 20.23−34.12	25.76 / 20.89−34.16	3.44 / 3.16−3.86	19.69 / 17.20−22.26	7.69 / 7.18−8.28	5.97 / 5.62−6.37	18.90 / 16.08−22.37	1.02 / 0.91−1.24	5.11 / 4.85−5.38
◆◆◆◇	5.54 / 4.62−6.81	10.19 / 7.90−14.53	9.88 / 7.90−13.45	10.30 / 7.95−14.34	10.69 / 8.36−15.57	2.36 / 2.15−2.55	9.99 / 7.68−13.32	4.79 / 4.34−5.50	3.46 / 3.07−3.86	8.57 / 7.12−11.06	0.45 / 0.37−0.62	3.07 / 2.78−3.47
◆◆◆◆	11.44 / 10.23−12.52	20.88 / 17.23−24.72	19.17 / 16.43−21.87	23.53 / 19.02−28.38	24.06 / 18.97−29.09	3.59 / 3.36−3.90	18.55 / 16.07−21.22	7.89 / 7.39−8.38	6.09 / 5.71−6.54	17.95 / 15.16−20.81	0.99 / 0.87−1.16	5.12 / 4.80−5.49

Top numbers show trimmed means (25% trim was applied to reduce the influence of outliers)
Bottom numbers show 95% confidence intervals for the trimmed mean calculated by bootstrapping
Filled diamonds indicate which new features are enabled. From left-to-right; analytical integration of population sizes, species tree relaxed clocks, coordinated topology operators, and coordinated height operators.

Table 3: Average ESS per hour convergence for simulated data.

Statistic	Concatenation	Concatenation 10X	*BEAST	StarBEAST2 GT-UCLN	StarBEAST2 ST-UCLN
Birth-death probability	898.4 (743.2–1048.8)	21.4 (18.3–27.2)	140.5 (129.3–154.2)	169.6 (154.0–188.3)	126.9 (108.5–144.7)
Clock rates σ	1383.9 (1272.9–1502.7)	58.5 (52.5–64.8)	329.5 (304.2–352.0)	324.2 (309.6–342.9)	278.1 (260.3–293.3)
Coalescent probability	NA	NA	30.2 (26.5–34.8)	60.8 (53.8–71.6)	49.5 (43.8–56.6)
Extinction fraction	834.0 (758.9–901.7)	97.8 (89.8–104.5)	281.5 (269.0–291.4)	245.7 (234.8–252.6)	189.1 (180.8–198.6)
Log-posterior	930.7 (870.0–999.5)	9.6 (8.6–10.7)	36.6 (31.4–41.4)	67.6 (60.1–79.7)	57.4 (50.0–66.4)
Log-prior	1389.6 (1283.0–1491.8)	21.7 (19.4–23.5)	39.5 (34.8–43.5)	251.8 (235.0–272.0)	245.0 (225.5–259.7)
Net diversification rate	920.7 (815.2–1000.2)	98.1 (89.4–105.9)	285.1 (268.3–298.6)	249.5 (235.2–263.7)	181.6 (172.5–194.1)
Phylogenetic likelihood	918.5 (859.4–974.3)	9.4 (8.3–10.7)	72.0 (54.2–92.0)	127.6 (115.0–143.4)	83.2 (72.0–96.7)
Population mean	NA	NA	42.0 (37.0–48.2)	105.0 (86.4–127.2)	86.8 (74.4–100.9)
Species tree height	788.0 (645.4–942.6)	11.4 (9.4–14.0)	11.9 (9.6–14.7)	130.0 (112.3–145.3)	75.5 (60.1–90.6)
Species tree length	754.5 (618.9–898.4)	9.5 (8.2–11.9)	107.8 (95.3–120.6)	138.9 (127.1–155.6)	103.6 (86.1–119.3)
Substitution model α	2924.4 (2490.0–3436.6)	334.3 (313.8–354.5)	625.8 (608.0–644.7)	816.6 (768.3–869.8)	659.0 (610.2–710.3)
Substitution model κ	4431.6 (4160.5–4685.2)	383.8 (374.9–393.4)	671.8 (658.0–689.0)	955.1 (910.2–1024.4)	781.4 (689.2–873.2)

25% trim was applied to reduce the influence of outliers

Numbers in parentheses show 95% confidence intervals calculated by bootstrapping

Table 4: Average ESS per million states convergence for simulated data.

Statistic	Concatenation	Concatenation 10X	*BEAST	StarBEAST2 ST-UCLN	StarBEAST2 GT-UCLN
Birth-death probability	269.60 (251.00–292.08)	53.17 (49.80–56.54)	6.24 (5.77–6.77)	13.19 (11.63–14.65)	13.53 (12.56–14.50)
Clock rates σ	297.00 (279.18–313.90)	73.63 (70.37–77.05)	11.29 (9.99–12.93)	22.56 (19.75–25.59)	23.02 (21.53–24.37)
Coalescent probability	NA	NA	1.39 (1.26–1.52)	4.52 (4.17–5.01)	3.97 (3.63–4.33)
Extinction fraction	192.51 (185.51–200.33)	150.34 (146.81–153.64)	11.49 (10.24–12.80)	19.38 (17.43–21.57)	21.23 (20.03–22.28)
Log-posterior	147.60 (142.03–153.04)	12.16 (11.75–12.65)	1.53 (1.38–1.67)	5.02 (4.62–5.53)	4.34 (3.97–4.76)
Log-prior	314.72 (289.39–335.46)	43.77 (42.41–44.93)	2.40 (2.26–2.54)	19.09 (16.79–21.24)	19.09 (17.58–20.49)
Net diversification rate	222.97 (213.38–231.93)	149.85 (145.79–153.59)	11.41 (10.20–12.68)	19.68 (17.46–22.16)	21.91 (20.77–22.95)
Phylogenetic likelihood	144.43 (140.04–149.13)	11.74 (11.31–12.23)	3.71 (3.40–4.03)	7.71 (7.10–8.38)	8.72 (8.20–9.27)
Population mean	NA	NA	2.08 (1.90–2.26)	7.04 (6.43–7.74)	6.55 (6.00–7.19)
Species tree height	253.66 (233.31–274.76)	27.09 (24.13–30.11)	0.54 (0.47–0.63)	1.36 (1.13–1.61)	1.92 (1.67–2.22)
Species tree length	221.92 (199.22–254.28)	24.70 (21.98–28.13)	4.54 (3.80–5.56)	10.13 (7.03–12.80)	11.12 (9.52–12.77)
Substitution model α	467.99 (440.77–504.04)	224.65 (218.83–229.63)	12.46 (11.04–14.88)	30.41 (24.35–36.49)	39.71 (35.21–44.65)
Substitution model κ	1013.70 (985.60–1040.98)	241.60 (239.18–243.39)	12.73 (11.23–14.94)	32.65 (25.39–39.76)	44.35 (38.19–50.56)

25% trim was applied to reduce the influence of outliers
Numbers in parentheses show 95% confidence intervals calculated by bootstrapping