SUPPLEMENTARY MATERIALS (SB-2003-29)

TABLES

Table S1.

(a) Mean posterior probability of the true topology across 100 replicates.

Assumed			**Simulated Model**			
Model	**JC**	**K2P**	**HKY**	**GTR**	**GTR+Γ**	**GTR+Γ+I**
JC	0.0683	0.0451	0.0583	0.0310	0.0077	0.0068
K2P	0.0676	0.0426	0.0489	0.0354	0.0130	0.0103
HKY	0.0681	0.0425	0.0471	0.0354	0.0137	0.0113
GTR	0.0681	0.0420	0.0463	0.0319	0.0141	0.0119
GTR+Γ	0.0670	0.0416	0.0453	0.0298	0.0086	0.0074
GTR+Γ+I	0.0667	0.0411	0.0448	0.0290	0.0087	0.0072

(b) Standard deviation of the posterior probability of the true topology.

Assumed			**Simulated Model**			
Model	**JC**	**K2P**	**HKY**	**GTR**	**GTR+Γ**	**GTR+Γ+I**
JC	0.1078	0.0713	0.0876	0.0926	0.0267	0.0296
K2P	0.1093	0.0585	0.0660	0.0932	0.0473	0.0289
HKY	0.1082	0.0576	0.0636	0.0910	0.0480	0.0298
GTR	0.1089	0.0558	0.0629	0.0839	0.0360	0.0293
GTR+Γ	0.1063	0.0543	0.0604	0.0772	0.0143	0.0146
GTR+Γ	0.1063	0.0543	0.0604	0.0772	0.0143	0.0146
GTR+Γ+I	0.1066	0.0544	0.0599	0.0747	0.0140	0.0158

FIGURE CAPTIONS

Figure S1. Tree space sampled by shorter (500,000G) and longer (5,000,000G) runs. Each graph compares the tree space sampled in the shorter run with that sampled in the longer run for the replicate indicated in the upper left hand corner of the graph. We checked 5 replicates from each of four model combinations to see if increasing the run length would affect the posterior distribution sampled. Each point represents one topology found in the 95% credible set of at least one of the two posterior distributions compared. Points are arranged as to maximally reflect the Robinson-Foulds distances between all possible pair-wise combinations. Thus, a pair of points close in the graph represent two topologies that are very similar. The points are colored in reference to whether the topology is present in the 95% credible set of both or just one of the posterior distributions compared.

The color-coding we used is as follows:

	Number of Generations:	
	500,000	5,000,000
Red	upper 95%	not found
Yellow	upper 95%	lower 5%
Green	upper 95%	upper 95%
Blue-Green	lower 5%	upper 95%
Blue	not found	upper 95%

White circles denote the true topology, and stars indicate the position that the true tree would have taken, had either of the two runs sampled the true topology. Empty white circles indicate that the true topology was found in one or both of the runs, but was not present in the 95% credible set of either run.

Figure S2. Examples of incongruence between two topology posterior distributions. These plots were generated by comparing two runs that utilized different data sets. Notice that the two posterior distributions overlap narrowly, as indicated by a narrow strip of green dots separating a population of red dots from a population of blue dots. Color-coding for this graph is the same as that described in Figure S1.

Figure S3. Further evidence that the posterior distributions of the shorter and longer runs were congruent. (a) Here we plot the posterior probability of the topologies found in the longer runs on the x-axis and the posterior probability of the corresponding topologies for the shorter runs on the y-axis. For each graph, both axes have a minimum value of 0.0 and a maximum value equal to the number present in the lower right corner of the graph. Points are colored as described in Figure S1. Note that the points fit tightly to the line y = x, indicating congruence between the shorter and longer runs. Also note that all blue and red dots are located very near the origin, suggesting that the presence of these points is due to increased sample size for the longer runs, and not due to fundamental incongruence between shorter and longer runs. The plots seen in (b) are identical to those seen in (a), except that the axes are scaled to 1/10 that seen in (a).

Figure S4. Examples of incongruence between the posterior distributions of two independent Bayesian analyses. (a) As in Figure S3, we plot the posterior probability of the topologies found in the longer runs on the x-axis and the posterior probability of the corresponding topologies for the shorter runs on the y-axis. Note the low degree of correlation between the posterior probabilities of the topologies found in the two runs. Also note that the red and blue points are often not near the origin, suggesting that the two runs are sampling fundamentally different parts of tree space. The plots seen in (b) are identical to those seen in (a), except that the axes are scaled to 1/10 that seen in (a).

Figure S5. The reliability of bipartition posterior probability estimates as it relates to the number of samples taken. Each graph compares the posterior probabilities of bipartitions sampled in two independent runs under identical conditions. The numbers presented on the right are estimates of the CPU time that it would take to perform the 3600 Bayesian

analyses needed to reproduce Figure 4 (see the manuscript). Note that taking more than 20,000 samples improves the reliability of the estimates very little, while the computational effort required increases linearly.

Figure S6. The effect of increasing sequence length on the bias due to over-parameterization and under-parameterization (see Results section of the manuscript for information regarding how we obtained these data). Here we plot the bipartition posterior probabilities obtained when the correct model was assumed (x-axis) against the bipartition posterior probabilities obtained when an incorrect model was assumed (y-axis). Arrows indicate the presence of points in the corner of a graph. Note that the bias seen when the assumed model is over-parameterized decreases with increased sequence length, whereas the bias seen when the assumed model is under-parameterized is exaggerated with increased sequence length.

Figure S1.

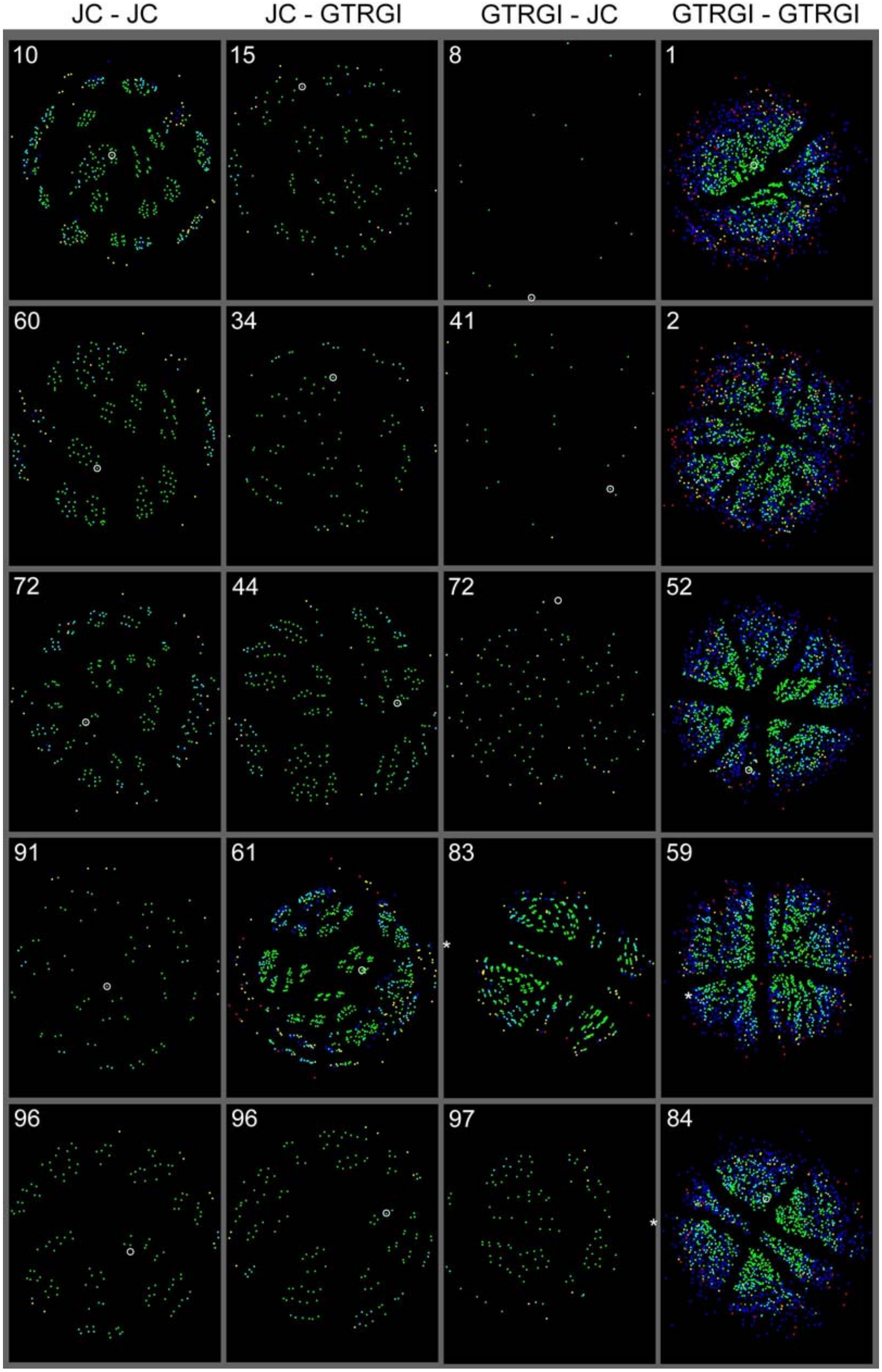

JC - JC
JC - GTRGI
GTRGI - JC
GTRGI - GTRGI
10
15
8
1
60
34
41
2
72
44
72
52
91
61
*
83
59
*
96
96
97
*
84

Figure S2.

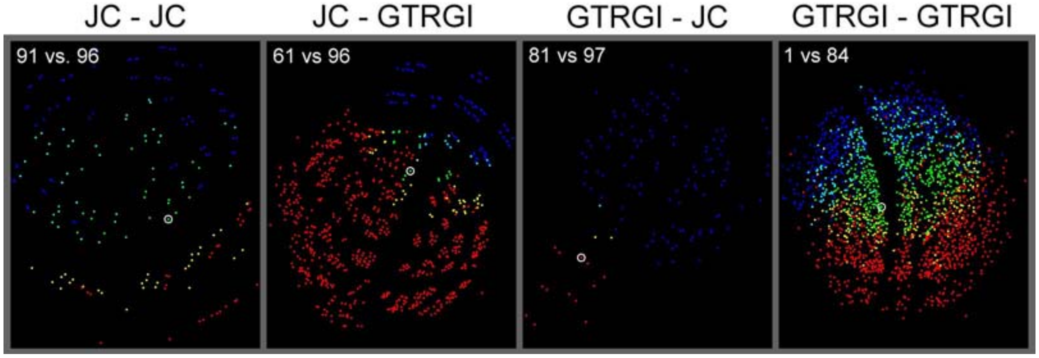

JC - JC
JC - GTRGI
GTRGI - JC
GTRGI - GTRGI
91 vs. 96
61 vs 96
81 vs 97
1 vs 84

Figure S3. (a)

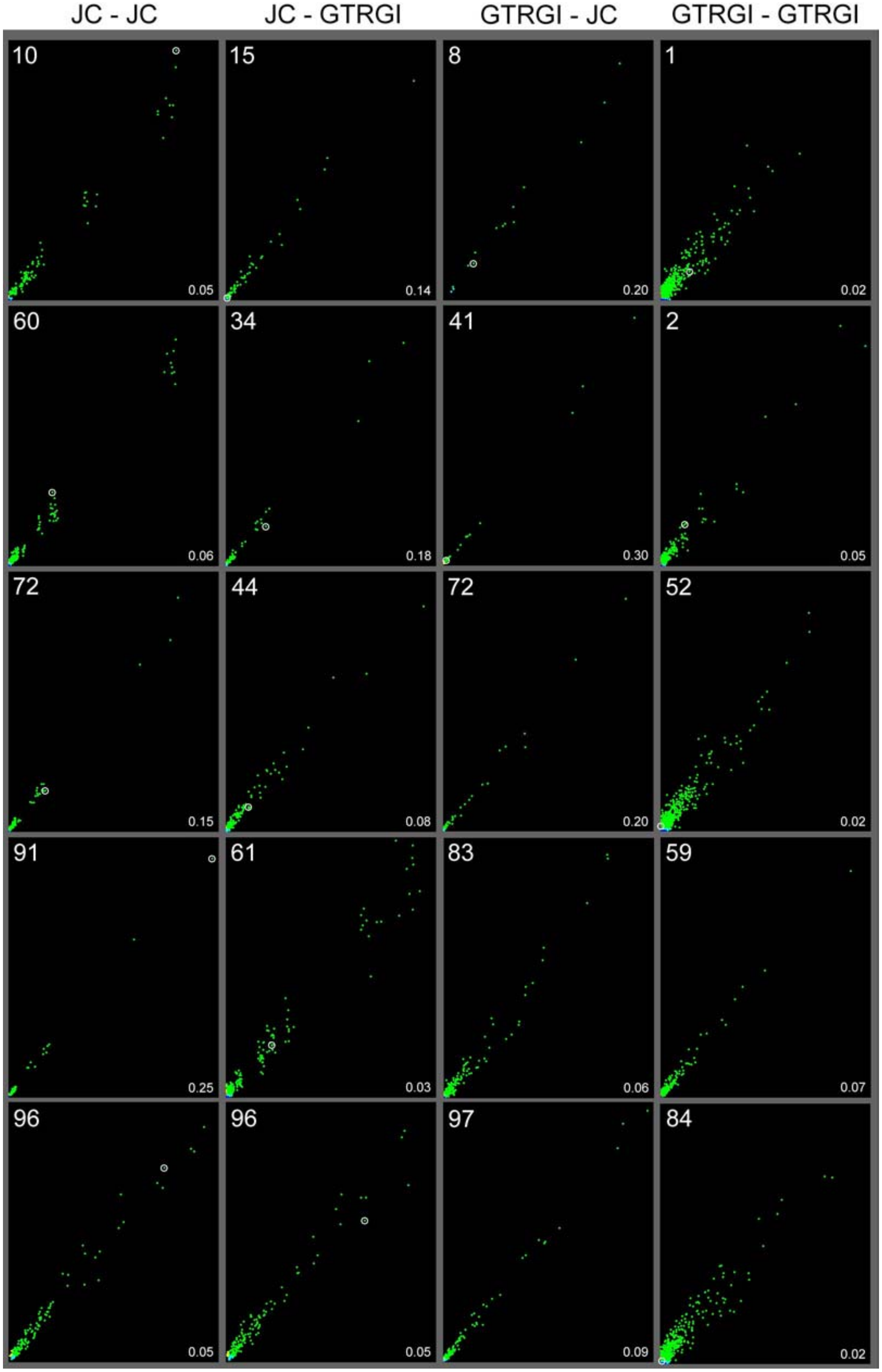
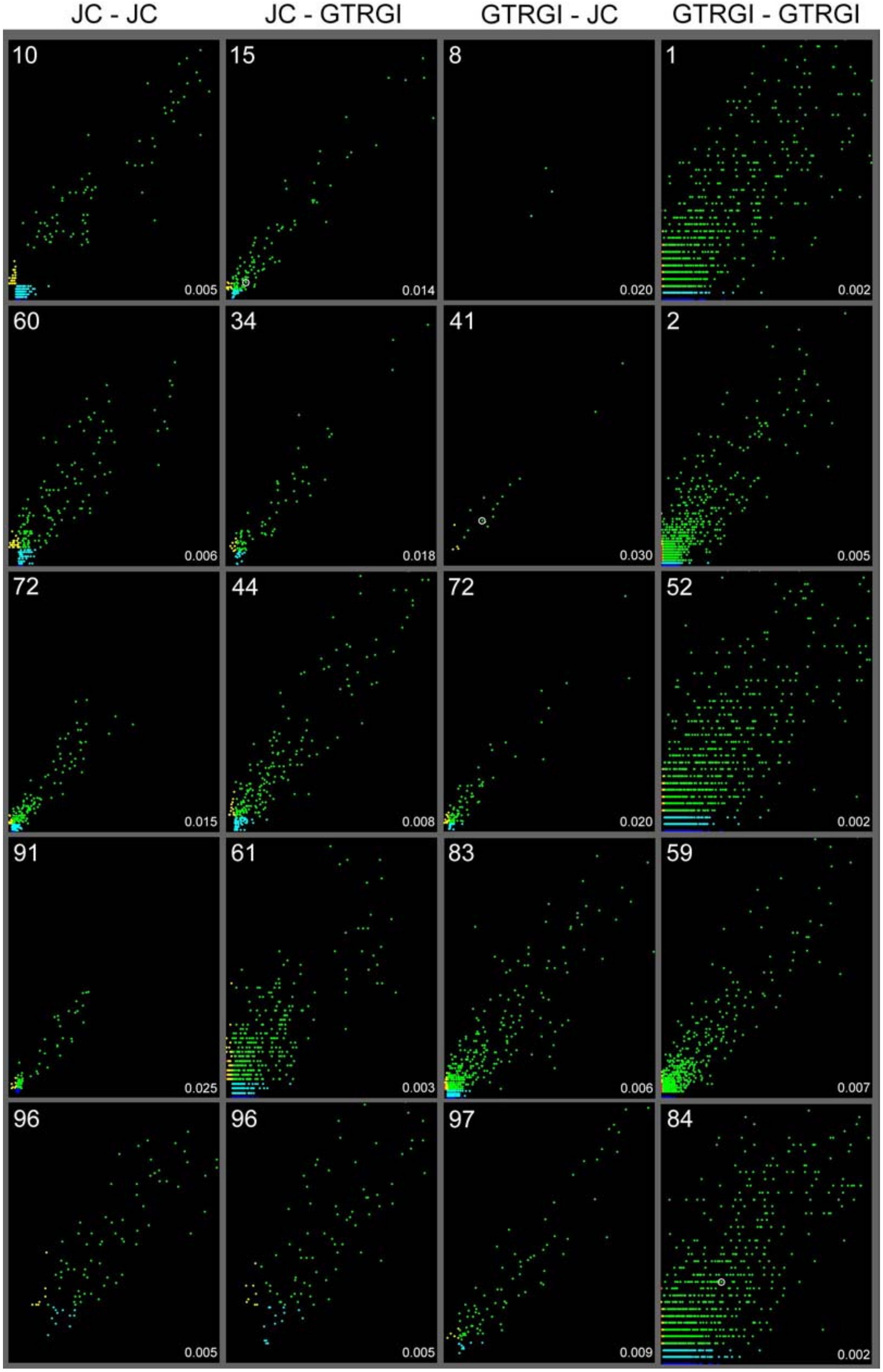

JC - JC
JC - GTRGI
GTRGI - JC
GTRGI - GTRGI
10
0.05
15
0.14
8
0.20
1
0.02
60
0.06
34
0.18
41
0.30
2
0.05
72
0.15
44
0.08
72
0.20
52
0.02
91
0.25
61
0.03
83
0.06
59
0.07
96
0.05
96
0.05
97
0.09
84
0.02

Figure S3. (b)



JC - JC
JC - GTRGI
GTRGI - JC
GTRGI - GTRGI
10
0.005
15
0.014
8
0.020
1
0.002
60
0.006
34
0.018
41
0.030
2
0.005
72
0.015
44
0.008
72
0.020
52
0.002
91
0.025
61
0.003
83
0.006
59
0.007
96
0.005
96
0.005
97
0.009
84
0.002

Figure S4.

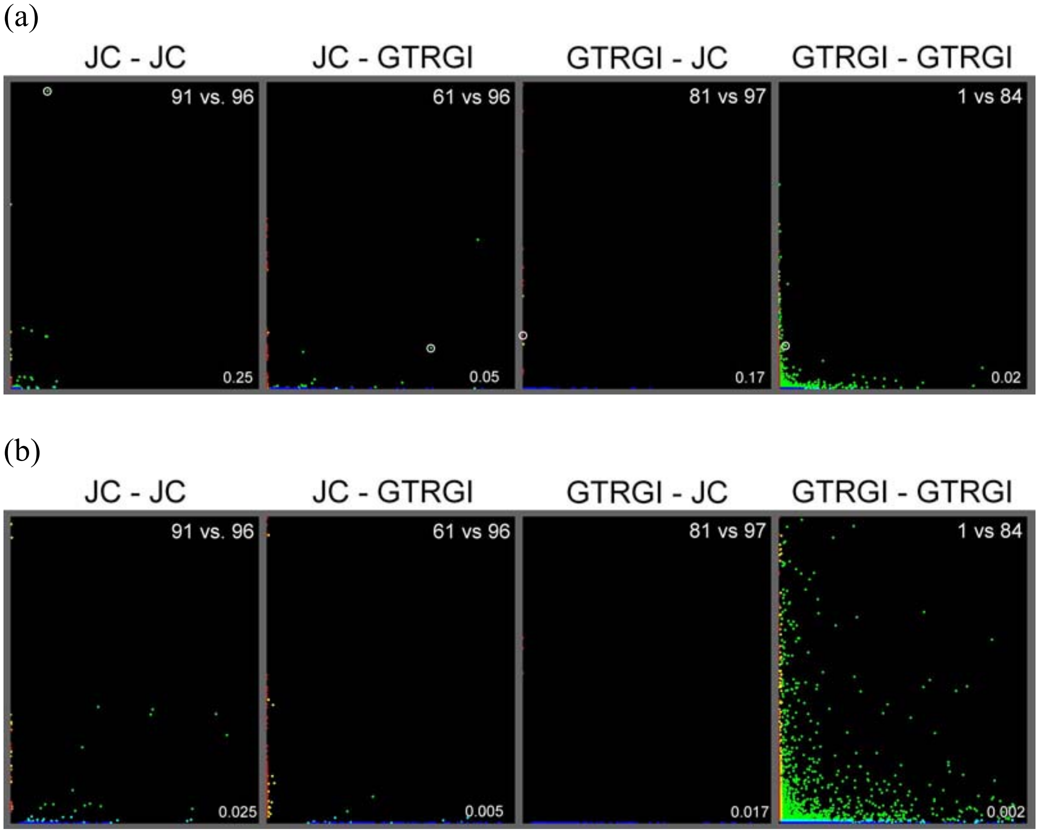

(a)
JC - JC
91 vs. 96
0.25
JC - GTRGI
61 vs 96
0.05
GTRGI - JC
81 vs 97
0.17
GTRGI - GTRGI
1 vs 84
0.02
(b)
JC - JC
91 vs. 96
0.025
JC - GTRGI
61 vs 96
0.005
GTRGI - JC
81 vs 97
0.017
GTRGI - GTRGI
1 vs 84
0.002

Figure S5.

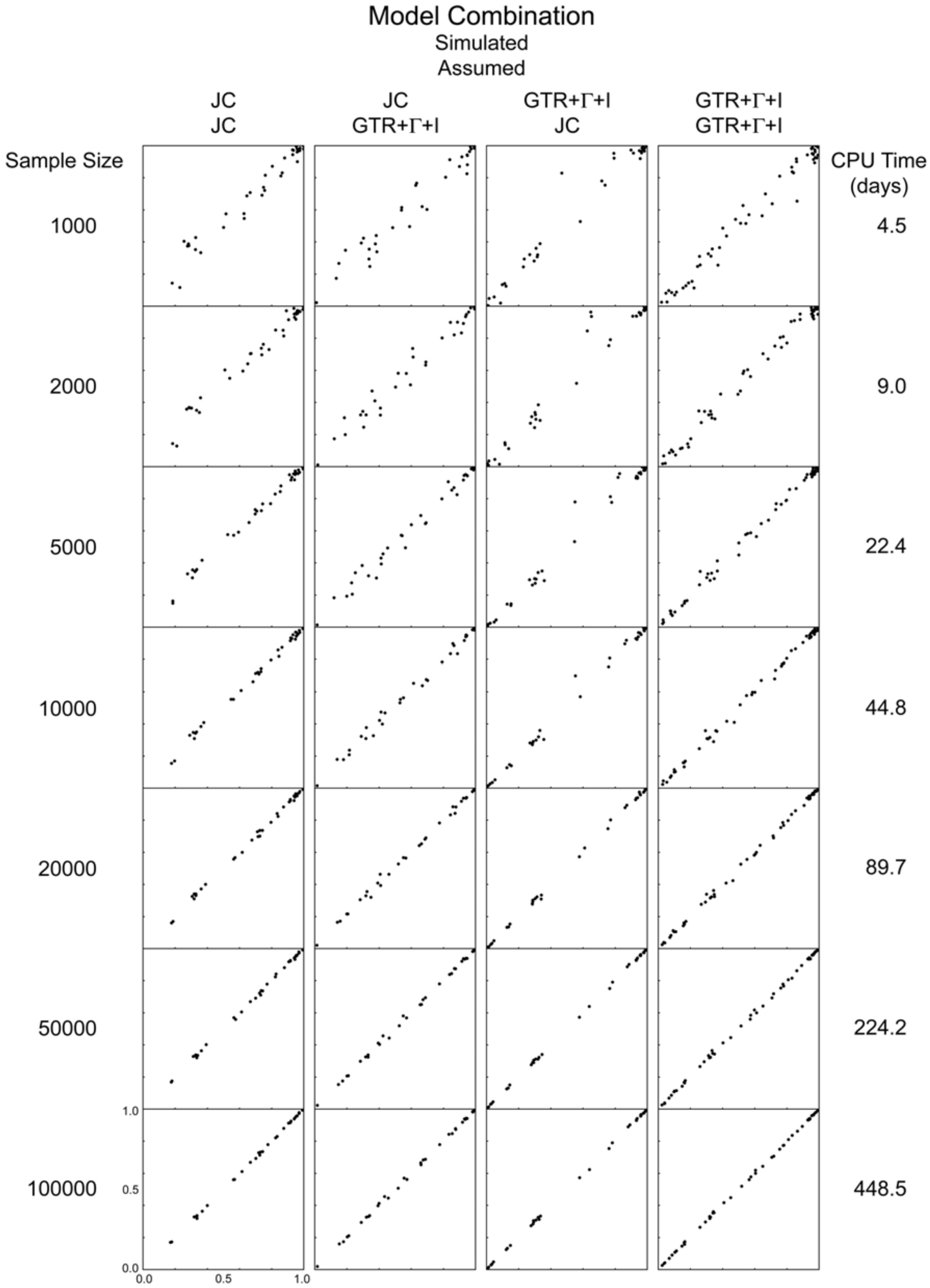
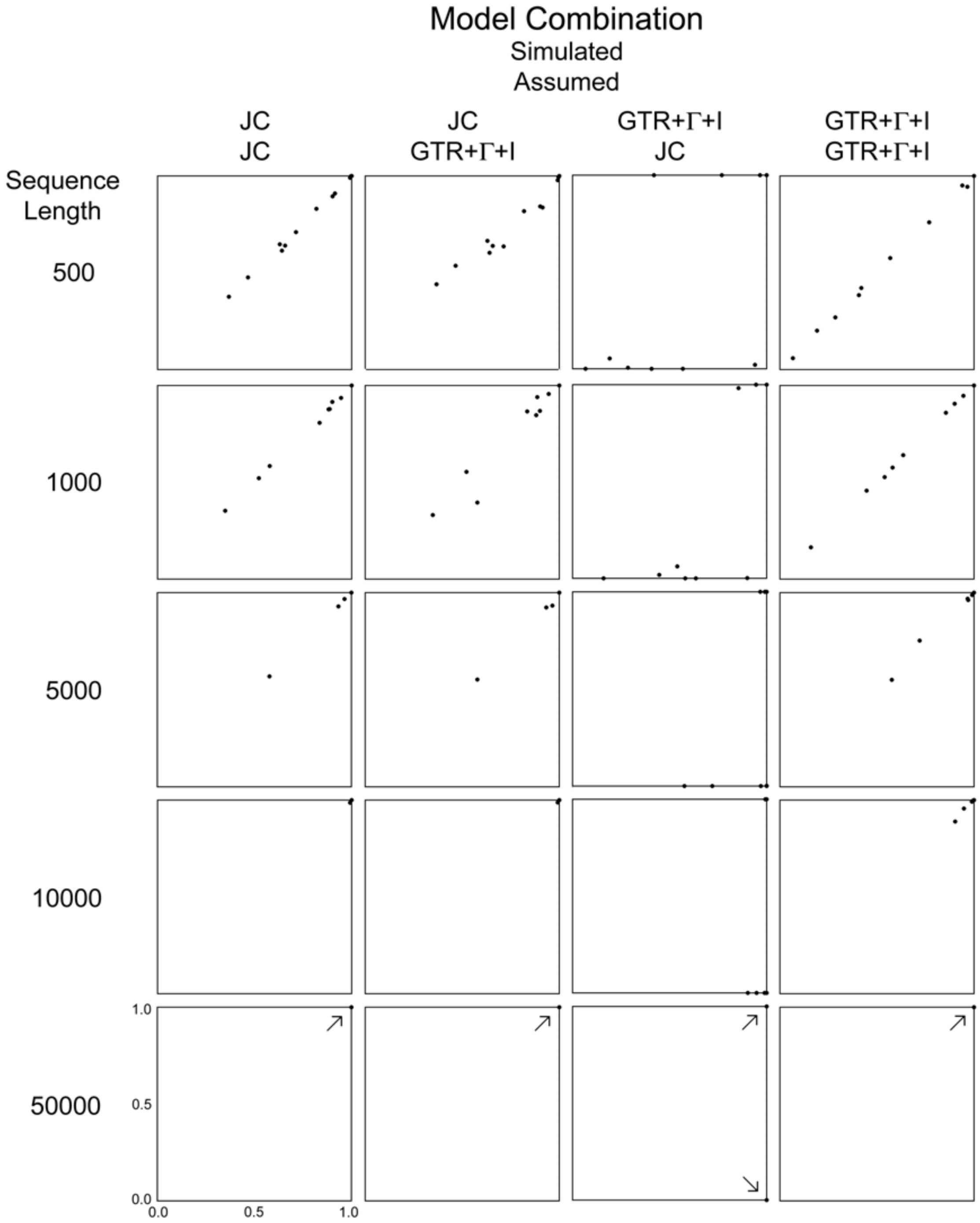

Model Combination
Simulated
Assumed
JC
JC
JC
GTR+Γ+I
GTR+Γ+I
JC
GTR+Γ+I
GTR+Γ+I
Sample Size
CPU Time
(days)
1000
4.5
2000
9.0
5000
22.4
10000
44.8
20000
89.7
50000
224.2
100000
448.5
1.0
0.5
0.0
0.0
0.5
1.0

Figure S6.


Model Combination
Simulated
Assumed
JC
JC
JC
GTR+Γ+I
GTR+Γ+I
JC
GTR+Γ+I
GTR+Γ+I
Sequence
Length
500
1000
5000
10000
50000
1.0
0.5
0.0
0.0
0.5
1.0