Table. Maximum Likelihood analysis of substitution models

Model	Parameters	BIC	AICc	InL	(+I)	(+G)	R	f(A)	f(T)	f(C)	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)
T92+G	68	6149.322	5609.421	-2736.485	n/a	1.99	2.57	0.217	0.217	0.283	0.283	0.030	0.039	0.204	0.030	0.204	0.039	0.030
T92	67	6150.672	5618.704	-2742.133	n/a	n/a	2.51	0.217	0.217	0.283	0.283	0.031	0.040	0.203	0.031	0.203	0.040	0.031
T92+I	68	6151.317	5611.416	-2737.482	0.24	n/a	2.55	0.217	0.217	0.283	0.283	0.030	0.039	0.204	0.030	0.204	0.039	0.030
HKY+G	70	6156.155	5600.387	-2729.955	n/a	2.05	2.57	0.175	0.260	0.282	0.283	0.036	0.039	0.204	0.024	0.204	0.039	0.024
HKY	69	6157.017	5609.183	-2735.359	n/a	n/a	2.52	0.175	0.260	0.282	0.283	0.037	0.040	0.203	0.025	0.203	0.040	0.025
T92+G+I	69	6158.223	5610.389	-2735.962	0.00	1.99	2.57	0.217	0.217	0.283	0.283	0.030	0.039	0.204	0.030	0.204	0.039	0.030
K2+G	67	6158.969	5627.001	-2746.281	n/a	1.99	2.56	0.250	0.250	0.250	0.250	0.035	0.035	0.180	0.035	0.180	0.035	0.035
HKY+I	70	6159.383	5603.616	-2731.569	0.23	n/a	2.55	0.175	0.260	0.282	0.283	0.036	0.039	0.204	0.024	0.204	0.040	0.024
K2	66	6160.251	5636.217	-2751.896	n/a	n/a	2.51	0.250	0.250	0.250	0.250	0.036	0.036	0.179	0.036	0.179	0.036	0.036
K2+I	67	6162.123	5630.155	-2747.859	0.24	n/a	2.55	0.250	0.250	0.250	0.250	0.035	0.035	0.179	0.035	0.179	0.035	0.035
TN93+G	71	6165.106	5601.407	-2729.458	n/a	2.07	2.56	0.175	0.260	0.282	0.283	0.036	0.039	0.221	0.024	0.191	0.039	0.024
TN93	70	6165.779	5610.012	-2734.767	n/a	n/a	2.52	0.175	0.260	0.282	0.283	0.036	0.040	0.220	0.025	0.190	0.040	0.025
HKY+G+I	71	6166.101	5602.401	-2729.955	0.00	2.05	2.57	0.175	0.260	0.282	0.283	0.036	0.039	0.204	0.024	0.204	0.039	0.024
TN93+I	71	6168.378	5604.678	-2731.093	0.23	n/a	2.55	0.175	0.260	0.282	0.283	0.036	0.039	0.220	0.024	0.191	0.039	0.024
K2+G+I	68	6168.915	5629.014	-2746.281	0.00	1.99	2.56	0.250	0.250	0.250	0.250	0.035	0.035	0.180	0.035	0.180	0.035	0.035
TN93+G+I	72	6175.053	5603.421	-2729.458	0.00	2.07	2.56	0.175	0.260	0.282	0.283	0.036	0.039	0.221	0.024	0.191	0.039	0.024
GTR+G	74	6176.343	5588.847	-2720.157	n/a	2.10	2.56	0.175	0.260	0.282	0.283	0.009	0.064	0.221	0.006	0.191	0.041	0.040
GTR	73	6176.741	5597.176	-2725.329	n/a	n/a	2.51	0.175	0.260	0.282	0.283	0.009	0.065	0.220	0.006	0.190	0.041	0.040
GTR+I	74	6179.693	5592.196	-2721.831	0.22	n/a	2.54	0.175	0.260	0.282	0.283	0.009	0.064	0.220	0.006	0.191	0.041	0.040
GTR+G+I	75	6186.290	5590.861	-2720.157	0.00	2.10	2.56	0.175	0.260	0.282	0.283	0.009	0.064	0.221	0.006	0.191	0.041	0.040
JC	65	6349.151	5833.050	-2851.319	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G	66	6349.999	5825.964	-2846.770	n/a	2.26	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+I	66	6352.129	5828.095	-2847.835	0.22	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G+I	67	6359.945	5827.977	-2846.770	0.00	2.26	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083

NOTE.-- This table presents the goodness of fit for multiple substitution models, including their parameter estimates. The Bayesian Information Criterion (BIC) scores are used to select the optimal model. Each model's Maximum Likelihood value (lnL), Akaike Information Criterion, corrected (AICc) value, BIC, and the number of parameters are presented (refer to Chapter 8 in [1] for more information). Evolutionary rate variation among sites was modeled using a discrete Gamma distribution (+G) with 5 rate categories [2] and a fraction of sites were allowed to be evolutionarily invariant (+I). Estimates of gamma shape parameter and fraction of invariant sites are also shown, as appropriate. Assumed or estimated values of transition/transversion bias (R) are shown for each model, as well. They are followed by nucleotide frequencies (f) and rates of base substitutions (r) for each nucleotide pair. The sum of r estimates is set to 1 for each model making comparisons across models easier. The ML analyses were conducted using an automatically generated tree topology [3]. The analytical procedure encompassed 34 nucleotide sequences. The complete deletion option was applied to eliminate positions containing gaps and missing data resulting in a final data set comprising 614 positions. Evolutionary analyses were conducted in MEGA12 [4] utilizing up to 4 parallel computing threads.

Abbreviations: GTR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93: Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor.

1. Nei M. and Kumar S. (**2000**). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
2. Yang Z. (**1994**) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *Journal of Molecular Evolution.*.**39**:306-314.
3. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. (**2011**). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**: 2731-2739.
4. Kumar S., Stecher G., Suleski M., Sanderford M., Sharma S., and Tamura K. (**2024**). Molecular Evolutionary Genetics Analysis Version 12 for adaptive and green computing. *Molecular Biology and Evolution* **41**:1-9.