

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)	r(
T92	45	1221.984	949.630	-429.166	n/a	n/a	2.19	0.180	0.180	0.320	0.320	0.027	0.047	0.226	0.027	0.226	0.047	0.027	0.
T92+I	46	1227.682	949.306	-427.974	0.40	n/a	2.22	0.180	0.180	0.320	0.320	0.026	0.047	0.226	0.026	0.226	0.047	0.026	0.
T92+G	46	1228.291	949.915	-428.279	n/a	1.08	2.22	0.180	0.180	0.320	0.320	0.026	0.047	0.227	0.026	0.227	0.047	0.026	0.
K2	44	1230.555	964.225	-437.492	n/a	n/a	2.19	0.250	0.250	0.250	0.250	0.039	0.039	0.172	0.039	0.172	0.039	0.039	0.
HKY	47	1235.219	950.821	-427.702	n/a	n/a	2.19	0.229	0.130	0.310	0.331	0.020	0.047	0.232	0.034	0.217	0.050	0.034	0.
K2+I	45	1235.963	963.610	-436.155	0.42	n/a	2.21	0.250	0.250	0.250	0.250	0.039	0.039	0.172	0.039	0.172	0.039	0.039	0.
T92+G+I	47	1236.373	951.974	-428.279	0.00	1.08	2.22	0.180	0.180	0.320	0.320	0.026	0.047	0.227	0.026	0.227	0.047	0.026	0.
K2+G	45	1236.655	964.302	-436.501	n/a	0.99	2.21	0.250	0.250	0.250	0.250	0.039	0.039	0.172	0.039	0.172	0.039	0.039	0.
HKY+I	48	1241.210	950.792	-426.657	0.38	n/a	2.20	0.229	0.130	0.310	0.331	0.020	0.046	0.232	0.034	0.217	0.050	0.034	0.
HKY+G	48	1241.826	951.408	-426.965	n/a	1.22	2.20	0.229	0.130	0.310	0.331	0.020	0.046	0.232	0.034	0.217	0.050	0.034	0.
TN93	48	1243.032	952.613	-427.568	n/a	n/a	2.19	0.229	0.130	0.310	0.331	0.019	0.046	0.212	0.034	0.243	0.049	0.034	0.
JC	43	1243.108	982.804	-447.809	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	0.
K2+G+I	46	1244.736	966.360	-436.501	0.00	0.99	2.21	0.250	0.250	0.250	0.250	0.039	0.039	0.172	0.039	0.172	0.039	0.039	0.
JC+I	44	1248.408	982.079	-446.418	0.43	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	0.
TN93+I	49	1249.020	952.582	-426.521	0.38	n/a	2.20	0.229	0.130	0.310	0.331	0.019	0.046	0.211	0.034	0.245	0.049	0.034	0.
JC+G	44	1249.068	982.739	-446.748	n/a	0.94	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.
TN93+G	49	1249.568	953.130	-426.796	n/a	1.18	2.20	0.229	0.130	0.310	0.331	0.019	0.046	0.208	0.034	0.249	0.049	0.034	0.
HKY+G+I	49	1249.907	953.470	-426.965	0.00	1.22	2.20	0.229	0.130	0.310	0.331	0.020	0.046	0.232	0.034	0.217	0.050	0.034	0.
JC+G+I	45	1257.149	984.796	-446.748	0.00	0.94	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.
TN93+G+I	50	1257.649	955.194	-426.796	0.00	1.18	2.20	0.229	0.130	0.310	0.331	0.019	0.046	0.208	0.034	0.249	0.049	0.034	0.
GTR	51	1262.728	954.256	-425.294	n/a	n/a	2.19	0.229	0.130	0.310	0.331	0.012	0.012	0.211	0.022	0.242	0.087	0.009	0.
GTR+I	52	1268.980	954.493	-424.380	0.36	n/a	2.21	0.229	0.130	0.310	0.331	0.012	0.012	0.210	0.022	0.246	0.084	0.009	0.
GTR+G	52	1269.482	954.995	-424.631	n/a	1.30	2.20	0.229	0.130	0.310	0.331	0.012	0.012	0.207	0.022	0.250	0.085	0.009	0.
GTR+G+I	53	1277.563	957.062	-424.631	0.00	1.30	2.20	0.229	0.130	0.310	0.331	0.012	0.012	0.207	0.022	0.250	0.085	0.009	0.

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 (InL), 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 23 nucleotide sequences. The partial deletion option was applied to eliminate all positions with less than 95% site coverage resulting in a final data set comprising 144 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., Petersen 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.

Disclaimer: While every effort has been made to ensure the accuracy of the caption, it is provided "as is" without warranty of any kind. Users are advised to thoroughly review the caption before its use for any purpose and report any errors or issues to the authors at www.megasoftware.net. The authors and their employers disclaim any liability for damages, including but not limited to special or consequential damages. Additionally, the authors expressly disclaim all other warranties, whether expressed or implied, including the suitability of the caption text for a specific purpose, use, or application.