

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)
K2+G	57	3095.652	2682.807	-1284.085	n/a	0.14	2.24	0.250	0.250	0.250	0.250	0.039	0.039	0.173	0.039	0.173	0.039	0.039
T92+G	58	3098.420	2678.343	-1280.842	n/a	0.14	2.32	0.222	0.222	0.278	0.278	0.033	0.041	0.195	0.033	0.195	0.041	0.033
K2+I	57	3101.071	2688.226	-1286.795	0.75	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
T92+I	58	3103.558	2683.480	-1283.411	0.75	n/a	2.29	0.222	0.222	0.278	0.278	0.033	0.042	0.194	0.033	0.194	0.042	0.033
K2+G+I	58	3104.724	2684.647	-1283.994	0.71	2.68	2.26	0.250	0.250	0.250	0.250	0.038	0.038	0.173	0.038	0.173	0.038	0.038
T92+G+I	59	3105.896	2678.588	-1279.953	0.71	2.66	2.34	0.222	0.222	0.278	0.278	0.033	0.041	0.195	0.033	0.195	0.041	0.033
TN93+G	61	3113.096	2671.326	-1274.299	n/a	0.13	2.46	0.193	0.251	0.310	0.246	0.035	0.043	0.282	0.027	0.120	0.034	0.027
TN93+I	61	3118.122	2676.352	-1276.812	0.75	n/a	2.37	0.193	0.251	0.310	0.246	0.036	0.045	0.270	0.028	0.127	0.035	0.028
HKY+G	60	3120.090	2685.551	-1282.423	n/a	0.14	2.34	0.193	0.251	0.310	0.246	0.038	0.047	0.172	0.029	0.216	0.037	0.029
TN93+G+I	62	3122.356	2673.356	-1274.302	0.63	0.92	2.47	0.193	0.251	0.310	0.246	0.035	0.043	0.278	0.027	0.125	0.034	0.027
HKY+I	60	3125.614	2691.074	-1285.185	0.75	n/a	2.29	0.193	0.251	0.310	0.246	0.039	0.048	0.170	0.030	0.214	0.038	0.030
HKY+G+I	61	3127.418	2685.648	-1281.460	0.71	2.61	2.36	0.193	0.251	0.310	0.246	0.038	0.047	0.172	0.029	0.217	0.037	0.029
GTR+G	64	3134.391	2670.932	-1271.065	n/a	0.13	2.45	0.193	0.251	0.310	0.246	0.028	0.057	0.282	0.022	0.118	0.053	0.036
GTR+I	64	3139.612	2676.153	-1273.676	0.75	n/a	2.34	0.193	0.251	0.310	0.246	0.029	0.056	0.268	0.022	0.126	0.057	0.035
GTR+G+I	65	3144.205	2673.517	-1271.345	0.60	0.75	2.46	0.193	0.251	0.310	0.246	0.026	0.057	0.276	0.020	0.124	0.053	0.036
JC+G	56	3153.104	2747.490	-1317.438	n/a	0.15	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	56	3158.070	2752.457	-1319.921	0.75	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	57	3162.943	2750.097	-1317.730	0.71	3.29	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2	56	3224.398	2818.784	-1353.085	n/a	n/a	2.00	0.250	0.250	0.250	0.250	0.042	0.042	0.167	0.042	0.167	0.042	0.042
T92	57	3231.595	2818.750	-1352.057	n/a	n/a	2.01	0.222	0.222	0.278	0.278	0.037	0.046	0.186	0.037	0.186	0.046	0.037
TN93	60	3251.754	2817.214	-1348.255	n/a	n/a	2.00	0.193	0.251	0.310	0.246	0.041	0.051	0.231	0.032	0.144	0.040	0.032
HKY	59	3253.777	2826.469	-1353.894	n/a	n/a	2.00	0.193	0.251	0.310	0.246	0.042	0.052	0.163	0.033	0.205	0.041	0.033
GTR	63	3275.584	2819.354	-1346.289	n/a	n/a	2.01	0.193	0.251	0.310	0.246	0.027	0.063	0.231	0.021	0.144	0.056	0.039
JC	55	3278.580	2880.199	-1384.803	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

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 29 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 368 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.