

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	58	6146.838	5686.744	-2785.206	n/a	0.40	3.64	0.214	0.214	0.286	0.286	0.023	0.030	0.225	0.023	0.225	0.030	0.023
T92+G+I	59	6156.777	5688.755	-2785.206	0.00	0.40	3.64	0.214	0.214	0.286	0.286	0.023	0.030	0.225	0.023	0.225	0.030	0.023
HKY+G	60	6159.127	5683.179	-2781.412	n/a	0.40	3.61	0.184	0.245	0.290	0.281	0.026	0.031	0.221	0.020	0.228	0.030	0.020
T92+I	58	6164.684	5704.589	-2794.129	0.58	n/a	3.59	0.214	0.214	0.286	0.286	0.023	0.031	0.224	0.023	0.224	0.031	0.023
K2+G	57	6167.406	5715.239	-2800.459	n/a	0.38	3.65	0.250	0.250	0.250	0.250	0.027	0.027	0.196	0.027	0.196	0.027	0.027
TN93+G	61	6169.038	5685.164	-2781.399	n/a	0.40	3.61	0.184	0.245	0.290	0.281	0.026	0.031	0.224	0.020	0.225	0.030	0.020
HKY+G+I	61	6169.066	5685.191	-2781.412	0.00	0.40	3.61	0.184	0.245	0.290	0.281	0.026	0.031	0.221	0.020	0.228	0.030	0.020
HKY+I	60	6176.563	5700.615	-2790.130	0.57	n/a	3.57	0.184	0.245	0.290	0.281	0.026	0.031	0.220	0.020	0.227	0.030	0.020
K2+G+I	58	6177.344	5717.250	-2800.459	0.00	0.38	3.65	0.250	0.250	0.250	0.250	0.027	0.027	0.196	0.027	0.196	0.027	0.027
TN93+G+I	62	6178.977	5687.176	-2781.399	0.00	0.40	3.61	0.184	0.245	0.290	0.281	0.026	0.031	0.224	0.020	0.225	0.030	0.020
K2+I	57	6184.347	5732.180	-2808.930	0.59	n/a	3.61	0.250	0.250	0.250	0.250	0.027	0.027	0.196	0.027	0.196	0.027	0.027
TN93+I	61	6186.501	5702.626	-2790.130	0.57	n/a	3.57	0.184	0.245	0.290	0.281	0.026	0.031	0.221	0.020	0.227	0.030	0.020
GTR+G	64	6187.308	5679.655	-2775.626	n/a	0.40	3.56	0.184	0.245	0.290	0.281	0.020	0.028	0.226	0.015	0.226	0.016	0.017
GTR+G+I	65	6197.246	5681.667	-2775.626	0.00	0.40	3.56	0.184	0.245	0.290	0.281	0.020	0.028	0.226	0.015	0.226	0.016	0.017
GTR+I	64	6204.265	5696.612	-2784.105	0.58	n/a	3.51	0.184	0.245	0.290	0.281	0.019	0.027	0.222	0.015	0.228	0.016	0.017
T92	57	6253.623	5801.456	-2843.568	n/a	n/a	3.45	0.214	0.214	0.286	0.286	0.024	0.032	0.223	0.024	0.223	0.032	0.024
HKY	59	6263.799	5795.778	-2838.718	n/a	n/a	3.45	0.184	0.245	0.290	0.281	0.027	0.032	0.219	0.020	0.226	0.031	0.020
TN93	60	6274.040	5798.092	-2838.869	n/a	n/a	3.45	0.184	0.245	0.290	0.281	0.027	0.032	0.221	0.020	0.223	0.031	0.020
K2	56	6277.472	5833.232	-2860.461	n/a	n/a	3.45	0.250	0.250	0.250	0.250	0.028	0.028	0.194	0.028	0.194	0.028	0.028
GTR	63	6293.068	5793.342	-2833.475	n/a	n/a	3.45	0.184	0.245	0.290	0.281	0.021	0.031	0.223	0.016	0.225	0.017	0.019
JC+G	56	6420.297	5976.058	-2931.874	n/a	0.41	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	6430.236	5978.069	-2931.874	0.00	0.41	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	6436.263	5992.023	-2939.857	0.58	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	55	6524.673	6088.361	-2989.031	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 725 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.