

Supplementary material

	fst	marginal	PRANK	MAFFT	ClustalΩ	MACSE
Avg alignment error (d_{seq})	0.00110	0.00101	0.01010	0.00982	0.01582	0.00932
Perfect alignments	2517	2452	22	2175	1150	1580
Best alignments	3602	3624	155	2763	1609	2081
Imperfect alignments	1071	1136	3566	1413	2438	2008
F1 score of positive selection	91.9%	90.8%	80.5%	73.5%	61.3%	70.6%
F1 score of negative selection	99.2%	99.1%	98.0%	97.2%	96.0%	97.4%

Table 1: Accuracy of COATi fst, COATi marginal, PRANK, MAFFT, ClustalΩ, and MACSE, on 5399 simulated sequence pairs. Perfect alignments have ($d_{seq} = 0$), best alignments have lowest d_{seq} , and imperfect alignments have $d_{seq} > 0$ when at least one aligner found a perfect alignment.