

PhyMet²: phylogenetic tree construction

Supplementary Materials 3

Michał Burdukiewicz, Przemysław Gagat, Sławomir Jabłoński,
Jarosław Chilimoniuk, Michał Gaworski, Paweł Mackiewicz, Marcin Łukaszewicz

To infer the phylogenetic tree of methanogens displayed in PhyMet² database (Fig.1, 2), we used the 16S rRNA sequences extracted from GenBank (<https://www.ncbi.nlm.nih.gov/>) and Silva (<https://www.arb-silva.de/>) databases. The sequences were aligned in MAFFT using slow and accurate algorithm L-INS-i with 1,000 cycles of iterative refinement [1]. Sites suitable for phylogenetic studies from the alignments were selected in GBlocks [2]. The phylogenetic tree was constructed using MrBayes [3]. We applied the mixed model of nucleotide substitutions to specify appropriate substitution models across the large parameter space [4]. Moreover, we assumed the heterogeneity rate across sites described by the invariant model and the discrete gamma model with five category rates. Two independent runs starting from random trees, each using eight Markov chains, were applied. The trees were sampled every 100 generations for 20,000,000 generations. In the final analysis, we selected the last 85,720 trees that reached the stationary phase and convergence, i.e. when the standard deviation of split frequencies stabilized and was lower than the proposed threshold of 0.01.

- [1] Katoh K, Standley DM: *MAFFT multiple sequence alignment software version 7: improvements in performance and usability*. Molecular Biology and Evolution 2013, 30(4):772-780.
- [2] Talavera G, Castresana J: *Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments*. Systematic Biology 2007, 56(4):564-577.
- [3] Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP: *MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space*. Systematic Biology 2012, 61(3):539-542.
- [4] Huelsenbeck JP, Larget B, Alfaro ME: *Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo*. Molecular Biology and Evolution 2004, 21(6):1123-1133

