PoMo- Polymorphism aware phylogenetic models Annotated Bibliography

Phyloseminar 2024

Bhagya Amarasinghe

Wilson, D. J., Hernandez, R. D., Andolfatto, P., & Przeworski, M. (2011). A population genetics-phylogenetics approach to inferring natural selection in coding sequences. *PLoS genetics*, *7*(12), e1002395.

Similar to PoMo in this paper they use patterns of polymorphism within species and divergence between species, but they use Bayesian sliding window approach to model intragenic variation in selection coefficients. This approach combines population genetics and phylogenetics, focusing on how specific genetic codes change over generations (coding sequences).

De Maio, N., Schlötterer, C., & Kosiol, C. (2013). Linking great apes genome evolution across time scales using polymorphism-aware phylogenetic models. *Molecular biology and evolution*, *30*(10), 2249-2262.

This is the first paper to introduce PoMo and they analyzed genome-wide synonymous site alignments of humans, chimpanzees, and two orangutan species (great apes). Unlike previous approaches, PoMo can separate mutation rates from fixation biases in shaping genetic differences between species. This is a good paper to go through if you want to know how the model works.

Schrempf, D., Minh, B. Q., De Maio, N., von Haeseler, A., & Kosiol, C. (2016). Reversible polymorphism-aware phylogenetic models and their application to tree inference. *Journal of theoretical biology*, *407*, 362-370.

The number of populations that could be analyzed with the non-reversible PoMo implementation is limited. So, this paper talks about a reversible mutation model which expands the scope of applications for PoMo. Here they tested revPoMo using simulation studies and real data from great apes, showing that it accurately estimates trees, divergence times, and mutation rates with comparable runtime to standard models. They show revPoMo is a valuable tool for various applications like dating speciation events and reconstructing species trees.

Borges, R., & Kosiol, C. (2020). Consistency and identifiability of the polymorphism-aware phylogenetic models. *Journal of Theoretical Biology*, *486*, 110074.

This is another paper that talks about effectiveness in PoMo at accurately estimating parameters. They show that the maximum a posteriori (MAP) tree estimator of PoMo consistently estimates the species tree. They tested PoMo with simulated data representing natural populations with incomplete lineage sorting and showed that PoMo's MAP tree accurately recovered the true tree with typical genome-wide data sampling.

Borges, R., Kotari, I., Bergman, J., Chase, M. A., Mugal, C. F., & Kosiol, C. (2022). Traditional phylogenetic models fail to account for variations in the effective population size. *bioRxiv*, 2022-09.

In this study, they compared evolutionary rates and times between traditional phylogenetic models and population-genetics-informed models using diverse taxa from animals, plants, and fungi. They talk about how traditional models struggle to establish meaningful distances between closely related taxa due to unaccounted-for variations in effective population size, leading to errors in phylogenetic inference.

Sanderson, M. J., Búrquez, A., Copetti, D., McMahon, M. M., Zeng, Y., & Wojciechowski, M. F. (2022). Origin and diversification of the saguaro cactus (Carnegiea gigantea): a within-species phylogenomic analysis. *Systematic biology*, 71(5), 1178-1194.

Rogers, J., Raveendran, M., Harris, R. A., Mailund, T., Leppälä, K., Athanasiadis, G., ... & Baboon Genome Analysis Consortium. (2019). The comparative genomics and complex population history of Papio baboons. *Science Advances*, 5(1), eaau6947.

These two papers are empirical examples where PoMo is applied. In both of these papers the model organisms show hybridization and thus PoMo is used as a model to account for current and possible ancient polymorphism when doing the phylogenetic analysis.