**Robust detection of convergent evolution at the molecular level**

Convergent evolution is a process in which two organisms develop the same phenotypic functionality which a common ancestor did not have. The recent advances in genome sequencing have allowed the genomes of many organisms to become available and this has led to the search for the molecular basis of convergent phenotypes in order to reveal the genetic basis of complex traits. Moreover, knowing where convergence has occurred on the genome will shed light on the mechanism of convergent evolution. Although recent studies of convergent evolution have focused on the convergence of phenotypes among species, these same techniques could be used to detect the convergent evolution of individual cell lines within an organism, e.g. cancer cells (Gerlinger et al. 2012). Studying convergence among pathogenic cell lines could be a more powerful and accurate method than Genome Wide Association Studies (GWAS) and lead to faster development of targeted therapies for any disease which has a genetic etiology.

Several different methods have been employed for detecting molecular convergence among organisms which display convergent phenotypes. Parker et al. (2013) use a phylogenetic screening mechanism followed by several tests to determine whether the convergence detected is due to chance or some other evolutionary process not related to convergence. This approach has come under attack by two recent papers (Zou and Zhang 2015; Thomas and Hahn 2015) for lacking a strong mechanism for detecting chance molecular convergence. However, the alternative strategies for detecting molecular convergence proposed by both of these articles have little to no power for detecting a small number of important convergent sites. A recent analysis of molecular convergence in marine mammals (Foote et al. 2015) uses an alternative strategy which is more sensitive to small numbers of convergent sites but doesn’t directly test appropriate alternative hypotheses that could generate the convergent signal. We propose a modification of the strategy used by Parker et al. that maintains power to detect small, isolated changes in the genome while controlling for alternative mechanisms of molecular convergence. We combine a phylogeny-based screen of the entire genome with robust tests of alternative hypotheses to find only those sites which are evolutionarily important for the phenotype under investigation. We demonstrate or method on data from fruit flies and marine mammals.

Literature Cited

Foote, AD, Y Liu, GWC Thomas, T Vinar, J Alfoldi, J Deng, S Dugan, CE van Elk, ME Hunter, V Joshi, Z Khan, C Kovar, SL Lee, K Linblad-Toh, A Mancia, R Nielsen, X Qin, Jiaxin Qu, BJ Raney, N Vijay, JBW Wolf, MW Hahn, DM Muzny, KC Worley, MTP Gilbert and Richard A Gibbs. 2015. Convergent evolution of the genomes of marine mammals. Nature Genetics. Advanced online publication; doi:10.1038/ng.3198

Gerlinger, M. et al. Intratumour heterogeneity and branched evolution revealed by multiregion sequencing. N. Engl. J. Med. 366, 883–892 (2012)

Parker J, Tsagkogeorga G, Cotton JA, Liu Y, Provero P, Stupka E, Rossiter SJ. 2013. Genome wide signatures of convergent evolution in echolocating mammals. Nature 502: 228-231

Thomas, GWC and MW Hahn. 2015. Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. Molecular Biology and Evolution. Pre-print

Zou Z, Zhang J. 2015. No genome-wide protein sequence convergence for echolocation. Mol. Biol. Evol. Accepted pending minor revisions.