University of Tennessee at Chattanooga

UTC Scholar

Apr 14th, 1:00 PM - 3:00 PM

Revised phylogenetic hypothesis and comparative evolution of mitochondrial genomes of Neotropical weakly electric fishes (Gymnotiformes)

Tiffany Harvey
University of Tennessee at Chattanooga

Fernando Alda-Pons
University of Tennessee at Chattanooga

Follow this and additional works at: https://scholar.utc.edu/research-dialogues

Departmental Honors Project Proposal

Revised phylogenetic hypothesis and comparative evolution of mitochondrial genomes of

Neotropical weakly electric fishes (Gymnotiformes)

Submitted by: Tiffany Harvey

Director: Fernando Alda-Pons

Introduction/Background:

The ability to generate and detect electric fields has evolved in several groups of fishes as a means of communication, navigation, and occasionally, predation (Elbassiouny et al., 2019). Despite the burden required for the daily energy expenditure of these fishes, molecular adaptations that enable electric fish to meet the metabolic demands of bioelectrogenesis remain unknown. The foundation of this metabolic adaptation lies in the capacity of mitochondria to supply enough energy to maintain these novel functions at the cellular level (Hill, 2014). Of the six lineages of fishes that have independently evolved specialized electric organs (EOs), the Gymnotiformes of South America is considered the most diverse group with 244 species, which together with the independent evolution of bioelectrogenesis in this group, makes them an ideal system for investigating patterns in molecular evolution associated with the origin of this novel sensory modality (Elbassiouny et al., 2019).

The main objective of this proposal is to test accelerated rates of molecular evolution in mitochondrial genes involved with energy metabolism in the Gymnotiformes, which are expected to have evolved high energy requirements related to their electrogenic capabilities in contrast to related fish orders of ostariophysan fishes that are not electrogenic. I will use comparative approaches of sequence analysis that allow the estimation of selective pressures

acting on protein-coding genes (Bromham, 2011), and phylogenetic-based statistical methods to detect changes in evolutionary rates across lineages in response to ecological or evolutionary processes (Torres-Dowdall et al., 2015).

Most importantly, all of the above comparative methods rely on accurate phylogenetic hypotheses. Therefore, to accomplish these major aims I first need to infer a robust phylogeny for this group of fishes that will be used as a framework for our comparative analyses; this task, however, has remained elusive to date (Arcila et al., 2017, Chakrabarty et al., 2017, Alda et al., 2019). I will recover complete mitochondrial genomes (mitogenomes) of Neotropical weakly electric fishes (Gymnotiformes), and report mitogenomes from 39 species in all families in the order, 90% of which had never been sequenced. This sequencing effort surpasses, both in the number of species and genes, previous works that have attempted to provide a phylogenetic hypothesis for the order but have not included a comprehensive number of species and the mitogenomes associated with them (Aguilar et al., 2019, Elbassiouny et al., 2019). Increasing the molecular data and taxa analyzed will allow us to develop a more robust hypothesis for comparison to previous studies.

First, I will collect next-generation sequencing data and assemble them to generate complete and annotated mitochondrial genomes. Second, I will infer a phylogenetic hypothesis including the new data and all mitochondrial genomes available from other orders of ostariophysan fishes. Finally, I will use this phylogenetic hypothesis to carry out comparative analyses of evolutionary rates across lineages to test hypotheses concerning the molecular changes accompanying major evolutionary innovations such as electrogenesis.

Methodology:

Collection and assembly of the next-generation raw reads into mitochondrial genomes will be completed with Geneious 2019.2.1 (Kearse et al., 2012) pairwise alignment tool. To carry out the development of the phylogenetic hypothesis with our new data, I will obtain from Genbank previously published mitochondrial genomes from all the other orders of Ostariophysi: Characiformes, Siluriformes and Cypriniformes. I will also download all genomes available from the order Gonorynchiformes—the sister group of Ostariophysi—that will be used as outgroups in my analyses. I will do a mass extraction of all the mitochondrial genes and align them using the program MUSCLE (Edgar, 2004) implemented in Geneious, following visual examination of the exported sequences. I will then estimate the phylogenetic relationships for each alignment, for a total of 37 gene trees (13 protein coding genes, 22 tRNA genes, 2 rRNA genes), and lastly, I will concatenate all the mitochondrial data to infer a new phylogenetic hypothesis for the complete dataset. Phylogenetic inference will be carried out using Maximum Likelihood and Bayesian Inference methods using IQTree (Nguyen et al, 2015) and MrBayes (Ronquist et al. 2012), respectively.

Finally, I will use this phylogenetic hypothesis to carry out comparative analyses of evolutionary rates across lineages to test hypotheses concerning the molecular changes accompanying major evolutionary innovations such as electrogenesis. These comparative analyses will allow me to infer new phylogenetic relationships among the orders of Ostariophysi and test for varying evolutionary rates in the mitochondrial genomes involved in the energy production of the Gymnotiformes and other ostariophysan lineages.

I will investigate the extent to which selection regulates gene sequence evolution using the ratio of divergence at non-synonymous (d_N) and synonymous (d_S) substitutions. Over long time scales, selection is generally inferred from evolutionary change between divergent lineages

that arose after a distant population split or speciation event. Each lineage is then represented by one stereotypic genome sequence, where sequence comparison allows quantifying evolutionary change at orthologous positions. The calculated d_N/d_S ratios among lineages in our dataset will allow me to investigate possible codon usage bias which will expose the differences in the frequency of occurrence of synonymous codons in coding DNA.

I will infer and compare the dynamics of evolution between mitochondrial genes that are involved in energy production and those that are not, using the program BAMM (Rabosky, 2014). BAMM (Bayesian Analysis of Macroevolutionary Mixtures) is a statistical framework that uses reversible jump MCMC to infer complex macroevolutionary dynamics of diversification and phenotypic evolution on phylogenetic trees. A set of rate shifts and associated parameter values will be mapped to a set of branches on the inferred phylogeny of Gymnotiformes thus defining a particular shift configuration. Using rjMCMC, BAMM simulates a posterior distribution of shift configurations conditioned on an observed phylogenetic dataset (Rabosky, 2014). This method provides statistical evidence for the number and location of rate shifts across phylogenies that I will use to evaluate if they are associated with those genes involved in energy production in the Gymnotiformes.

Timeline:

The collection and assembly of the next-generation raw read sequences into mitochondrial genomes, mass exportation and alignment of all the mitochondrial sequences, and phylogenetic inference are on track to be completed at the end of this semester. The comparative analyses of evolutionary rates across lineages will be expected to begin at the start of the fall semester if not before the end of this semester, time willing. The cumulative research completed from the fall semester will be compiled into a poster for the Fall 2020 National Collegiate

Honors Council Conference in Dallas, TX, which I will present. The project as a whole will be presented at the Spring 2021 Research Dialogues by me. Dr. Alda has expressed tentative plans of writing a scientific paper for publication based on the results of the project.

Works Cited

Aguilar, C., Miller, M. J., Loaiza, J. R., González, R., Krahe, R., & León, L. F. D. (2019). Tempo and mode of allopatric divergence in the weakly electric fish Sternopygus dariensis in the Isthmus of Panama. *Scientific Reports*, *9*(1). doi: 10.1038/s41598-019-55336-y

Alda, F., Tagliacollo, V. A., Bernt, M. J., Waltz, B. T., Ludt, W. B., Faircloth, B. C., … Chakrabarty, P. (2018). Resolving Deep Nodes in an Ancient Radiation of Neotropical Fishes in the Presence of Conflicting Signals from Incomplete Lineage Sorting. *Systematic Biology*, *68*(4), 573–593. doi: 10.1093/sysbio/syy085

Arcila, D., Ortí, G., Vari, R., Armbruster, J. W., Stiassny, M. L. J., Ko, K. D., … Betancur-R., R. (2017). Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nature Ecology & Evolution*, *1*(2). doi: 10.1038/s41559-016-0020

Bromham, L. (2011). The genome as a life-history character: why rate of molecular evolution varies between mammal species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1577), 2503–2513. doi: 10.1098/rstb.2011.0014

Chakrabarty, P., Faircloth, B. C., Alda, F., Ludt, W. B., Mcmahan, C. D., Near, T. J., … Alfaro, M. E. (2017). Phylogenomic Systematics of Ostariophysan Fishes: Ultraconserved Elements Support the Surprising Non-Monophyly of Characiformes. *Systematic Biology*, *66*(6), 881–895. doi: 10.1093/sysbio/syx038

Elbassiouny, A. A., Lovejoy, N. R., & Chang, B. S. W. (2019). Convergent patterns of evolution of mitochondrial oxidative phosphorylation (OXPHOS) genes in electric fishes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1790). doi: 10.1098/rstb.2019.0179

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics, 5(1), 113.

Hill, G. E. (2014). Cellular Respiration: The Nexus of Stress, Condition, and Ornamentation. *Integrative and Comparative Biology*, *54*(4), 645–657. doi: 10.1093/icb/icu029

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647–1649. doi: 10.1093/bioinformatics/bts199

Nguyen, L.-T., Schmidt, H. A., Haeseler, A. V., & Minh, B. Q. (2014). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution, 32(1), 268–274. doi: 10.1093/molbev/msu300

Torres-Dowdall, J., Henning, F., Elmer, K. R., & Meyer, A. (2015). Ecological and Lineage-Specific Factors Drive the Molecular Evolution of Rhodopsin in Cichlid Fishes. *Molecular Biology and Evolution*, *32*(11), 2876–2882. doi: 10.1093/molbev/msv159

Rabosky, D. L. (2014). Automatic Detection of Key Innovations, Rate Shifts, and Diversity-Dependence on Phylogenetic Trees. PLoS ONE, 9(2). doi: 10.1371/journal.pone.0089543

Ronquist, F., Teslenko, M., Mark, P. V. D., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology, 61(3), 539–542. doi: 10.1093/sysbio/sys029