compgen final project

Final Project

Hi C! (Get it?!!)

Introduction

This is a final project for the Comparative Genomics seminar in the spring of 2019.

While biologists characterize most animals by what they have, non-bilaterians are sometimes characterized by what they don't have. In sponges, traditionally these absences have been considered ancestral, an assumption contributing to their primitive image.

One timely discussion is whether sponges have lost a nervous system. The flip side of this debate, that ctenophores have a convergent one, also rests on assumptions of absence - because ctenophores lack many fundamental 'neural' genes they must use a very different mechanism. For both animals, it may be productive to first understand whether these absences are ancestral or secondary. This would be valuable for inferring whether the LCA of metazoans may have had a bilaterian-like sensory/coordinating system.

Synteny analysis provides one method of determining loss. In particular, the ghost locus hypothesis suggests that in the case of gene loss, the synteny surrounding the locus of the gene may be preserved even in the absence of the gene itself [1]. I will look for the ghost loci of absent 'neural' genes in sponges and ctenophores using synteny analysis.

The goal

Our goal is to classify whether absences for particular 'neural' genes in sponges and ctenophores are ancestral (i.e. gene was never there) or secondary.

I must first determine whether my genes of interest are found within conserved blocks of synteny in other animals. If absence in a particular taxon is ancestral, I would expect that the genes defining the synteny block not to be co-localized. If absence is secondary, I expect to find conservation of the synteny block in the absence of the gene of interest.

The data

- Data source: Published, and unpublished but publically available
  So far have: [NOTE: couldn't add to public fork via git lfs]
  -Mnemiopsis leidyi genome [2]
  -Pleurobrachia bachei genome [3]
  -Amphimedon queenslandica genome [4]
  -Oscarella carmela [5]. See presence/absence of 'neural' gene table in Riesgo et al., 2014 [6]

These genomes were selected because there are presence/absence lists of 'neural' genes readily available for them.

Other genomes to use (probably): Salpingoeca rosetta, Monosiga brevicollis, Trichoplax (can use as positive control?), Nematostella vectensis, Hydra magnipapillata, Homo sapiens.

- Data structure: FASTA files, GFF3 files, contingency table of presence/absence of 'neural' genes. More?

Background

Methods

Synteny analysis to search for ghost loci of absent 'neural' genes in sponges and ctenophores. I don't know how to do this yet…but I will learn!!

Results

Discussion

Some major assumptions:
- surrounding syntenic block does not exist before the emergence of the gene itself.
- there IS conservation of synteny around these genes across taxa.

Assessment

Was it successful in achieving the initial goal?

What are the main obstacles encountered?

What would you have done differently?

What are future directions this could go in?

References

1. Ramos OM, Barker D, Ferrier DE. Ghost loci imply hox and parahox existence in the last common ancestor of animals. Current biology. Elsevier; 2012;22: 1951–1956.

2. Joseph F. Ryan CES Kevin Pang. The genome of the ctenophore mnemiopsis leidyi and its implications for cell type evolution. Science. American Association for the Advancement of Science; 2013;342: 1242592. doi:10.1126/science.1242592

3. Leonid L. Moroz MRC Kevin M. Kocot. The ctenophore genome and the evolutionary origins of neural systems. Nature. Nature Publishing Group; 2014;510: 109–114. doi:10.1038/nature13400

4. Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier ME, Mitros T, et al. The amphimedon queenslandica genome and the evolution of animal complexity. Nature. Nature Publishing Group; 2010;466: 720.

5. Nichols SA, Roberts BW, Richter DJ, Fairclough SR, King N. Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/β-catenin complex. Proceedings of the National Academy of Sciences. National Acad Sciences; 2012;109: 13046–13051.

6. Ana Riesgo PJW Nathan Farrar. The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. Molecular Biology and Evolution. Oxford Academic; 2014;31: 1102–1120. doi:10.1093/molbev/msu057