Planned analyses for investigating Innexin clusters

Jennifer Ortiz, Melissa DeBiasse, and Joseph Ryan

v.1.0

27 June 2019

LIST OF ABBREVIATIONS

ML	maximum likelihood

1   INTRODUCTION: BACKGROUND INFORMATION AND SCIENTIFIC RATIONALE

1.1   BACKGROUND INFORMATION

Invertebrate gap junctions are formed by innexins. These genes have not been characterized in detail in the ctenophore *Mnemiopsis leidyi*. Preliminary BLAST analyses suggest that *M. leidyi* has a cluster of four unrelated innexins that show correlated expression.

1.2   RATIONALE

A detailed comparative phylogenetic analyses of innexins will allow us to better understand the function of these genes in ctenophores and role of clustering in *M. leidyi* innexins.

1.3   HYPOTHESES

The innexin cluster in *M. leidyi* is ancient and was present in the last common ancestor of *Beroe* and *Mnemiopsis*.

1.4   OBJECTIVES

Phylogenetically classify innexin genes in *M. leidyi*, *Beroe ovata*, and *Pleurobrachia bachei*.

2   STUDY DESIGN AND ENDPOINTS

2.1 Identify putative innexins within genomes

We will use the protein sequence of innexin 2 (isoform D) from *Drosophila melanogaster* (Abascal and Zardoya 2012) as a query sequence to identify putative innexins in the protein gene models of the following: *H. vulgaris*, *P. bachei*, *M. leidyi*, and *B. ovata*. Protein models with BLAST E-values below 1e-3 will be retained.

2.1.1 Building a database

We will search putative innexin genes against Pfam to identify the protein domain. If partial innexins occur, we will search for the complete sequence in unfiltered gene models and gene transcripts. Innexin predictions will be extended to the full length of the domain if found to be incomplete by 5 or fewer amino acids at the N or C terminus of the domain.

Innexin phylotocol

2.1.2 Filter duplicate copies of *B. ovata* innexin gene models

Our current genome assembly of the diploid *B. ovata* includes scaffolds from both haplotypes, with some regions of the genome represented by one collapsed sequence and others represented by both haplotypes. To be sure that we only include one representation of each *B. ovata* innexin, we will build an ML tree of the innexins found in *B. ovata* and use this tree to identify and remove any duplicate innexins present on phased haplotig scaffolds.

```
iqtree-omp –s [Bova_innexins.mafft-gb] –nt AUTO –bb 1000 –m TEST –pre [output
prefix] > iq.out 2> iq.err
```

2.2 Alignment of putative innexins

We will align putative innexin sequences with MAFFT using default parameters.

```
mafft [fasta_file] > [fasta_file].mafft
```

2.3 Innexin gene trees

We will infer the phylogenetic relationships among innexins by estimating gene trees.

2.3.1 IQTREE

```
iqtree-omp –s [infile.mafft-gb] –nt AUTO –bb 1000 –m TEST –pre [output
prefix] > iq.out 2> iq.err
```

2.3.2 RAxML with 25 starting parsimony trees and 25 random starting trees;

```
raxmlHPC–SSE3.PTHREADS –f a –T 25 –p [random_number] –# 25 –m PROTGAMMA[best-
fit_model] –s [alignment_file] –n [name]_mp
```

```
raxmlHPC–SSE3.PTHREDAS –f a –T 25 –d –p [random_number] –# 25 –m
PROTGAMMA[best-fit_model] –s [alignment_file] –n [name]_rt
```

2.3.3 MrBayes

We will run the following Bayesian analyses using MrBayes v3.2.6 with the following execution block (best model will be determined from IQ-TREE run):

```
prset aamodelpr = fixed(BEST_MODEL); lset rates = gamma;
```

2.4 Choose best tree to be the main figure

3   WORK COMPLETED SO FAR WITH DATES

We have identified the putative innexin cluster using BLAST.

4   LITERATURE REFERENCES

Abascal F, Zardoya R. Evolutionary analyses of gap junction protein families. *BBA.* 2012

Kuznetsov *et al*. Gap junctions in nematodes. *Russian Journal of Nematology*. 2016

Innexin phylotocol

Ryan *et al*. The Genome of the Ctenophore Mnemiopsis leidyi and Its Implications for Cell Type Evolution. *Science*. 2013

Sebé-Pedros *et al.* Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nature Ecology & Evolution*. 2018

Slivko-Koltchik G *et al.* Evolution of Pannexin/Innexin Gap Junction Protein Family. *Homo* 3: 70.

5   PHYLOTOCOL AMENDMENT HISTORY

Version	Date	Significant Revisions