Supplemental information #1: A new strategy to characterize the domain architecture structure of proteins of the innate inmune system in tunicate species

Cristian A. Velandia-Huerto*, Ernesto Parra, Federico D. Brown, Adriaan Gittenberger, Peter F. Stadler and Clara I. Bermúdez-Santana

April 16, 2019

Genomic sources

Sub-phylum	**Specie**	**Source**	**Version**
V	*Latimeria chalumnae*	`Ensembl FTP`[1]	Release 81
V	*Danio rerio*	`Ensembl FTP`[2]	Release 81
V	*Petromyzon marinus*	`Ensembl FTP`[3]	Release 81
T	*Ciona savignyi*	`Ensembl FTP`[4]	Release 81
T	*Ciona robusta*	`Ensembl FTP`[5]	Release 81
T	*Didemnum vexillum*	Universidad Nacional[8]	v.2
T	*Botryllus schlosseri*	`ANISEED`[9]	v.1
T	*Botrylloides leachii*	`ANISEED`[10]	v.1
T	*Molgula occidentalis*	`ANISEED`[11]	v.1
T	*Molgula oculata*	`ANISEED`[12]	v.1
T	*Oikopleura dioica*	`OikoBase`[13]	Version 3
C	*Branchiostoma floridae*	`JGI genome portal`[14]	v.1 and v.2
H	*Saccoglossus kowalevshii*	`NCBI FTP`[15]	Skow_1.1
E	*Patiria miniata*	`Echinobase`[16]	v2.0
E	*Strongylocentrotus purpuratus*	`Echinobase`[17]	v4.2

Table 1: Genomic data source. Described labels for Subphylums: **V**: vertebrates, **T**: tunicates, **C**: cephalochordates, **H**: hemichordates and **E**: echinoderms.

[1] ftp://ftp.ensembl.org/pub/release-81/fasta/latimeria_chalumnae/pep/Latimeria_chalumnae.LatCha1.pep.all.fa.gz

[2] ftp://ftp.ensembl.org/pub/release-81/fasta/danio_rerio/pep/Danio_rerio.GRCz10.pep.all.fa.gz

[3] ftp://ftp.ensembl.org/pub/release-81/fasta/petromyzon_marinus/dna/Petromyzon_marinus.Pmarinus_7.0.dna.toplevel.fa.gz

[4] ftp://ftp.ensembl.org/pub/release-81/fasta/ciona_savignyi/pep/Ciona_savignyi.CSAV2.0.pep.all.fa.gz

[5] ftp://ftp.ensembl.org/pub/release-81/fasta/ciona_intestinalis/pep/Ciona_intestinalis.KH.pep.all.fa.gz

[6] http://tunicatadvexillum.bioinf.uni-leipzig.de/

[7] https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fbs%2Fbotryllus_protein_fasta.zip

[8] https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fboleac%2FBoleac_proteins_v4_fasta.zip

[9] https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fmoocci%2Fmolgula_occidentalis_protein_fasta.zip

[10] https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fmoocul%2Fmolgula_oculata_protein_fasta.zip

[11] http://oikoarrays.biology.uiowa.edu/Oiko/Downloads.html

[12] https://genome.jgi.doe.gov/Brafl1/Brafl1.home.html

[13] ftp://ftp.ncbi.nlm.nih.gov/genomes/Saccoglossus_kowalevskii/protein/protein.fa.gz

[14] http://www.echinobase.org/Echinobase/PmDownload/pmin_proteins_v2.0.fa

[15] ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/235/GCF_000002235.4_Spur_4.2/GCF_000002235.4_Spur_4.2_protein.faa.gz