

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

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

December 17, 2018

Genomic sources

Sub-phylum	Specie	Source	Version
V	<i>Latimeria chalumnae</i>	Ensembl FTP ¹	Release 81
V	<i>Danio rerio</i>	Ensembl FTP ²	Release 81
V	<i>Petromyzon marinus</i>	Ensembl FTP ³	Release 81
T	<i>Ciona savignyi</i>	Ensembl FTP ⁴	Release 81
T	<i>Ciona robusta</i>	Ensembl FTP ⁵	Release 81
T	<i>Perophora viridis</i>	Universidad Nacional ⁶	Draft version
T	<i>Clavelina oblonga</i>	Universidad Nacional ⁷	Draft version
T	<i>Didemnum vexillum</i>	Universidad Nacional ⁸	v.2
T	<i>Botryllus schlosseri</i>	ANISEED ⁹	v.1
T	<i>Botrylloides leachii</i>	ANISEED ¹⁰	v.1
T	<i>Molgula occidentalis</i>	ANISEED ¹¹	v.1
T	<i>Molgula oculata</i>	ANISEED ¹²	v.1
T	<i>Oikopleura dioica</i>	OikoBase ¹³	Version 3
C	<i>Branchiostoma floridae</i>	JGI genome portal ¹⁴	v.1 and v.2
H	<i>Saccoglossus kowalevshii</i>	NCBI FTP ¹⁵	Skow_1.1
E	<i>Patiria miniata</i>	Echinobase ¹⁶	v2.0
E	<i>Strongylocentrotus purpuratus</i>	Echinobase ¹⁷	v4.2

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

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

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

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

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

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

⁶<http://tunicatapviridis.bioinf.uni-leipzig.de/Download.html>

⁷http://tunicatacoblonga.bioinf.uni-leipzig.de/Download_clob.html

⁸<http://tunicata.bioinf.uni-leipzig.de>

⁹https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fbs%2Fbotryllus_protein_fasta.zip

¹⁰https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fboleac%2FBoleac_proteins_v4_fasta.zip

¹¹https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fmoocci%2Fmolgula_occidentalis_protein_fasta.zip

¹²https://www.aniseed.cnrs.fr/aniseed/download/?file=data%2Fmoocul%2Fmolgula_oculata_protein_fasta.zip

¹³<http://oikoarrays.biology.uiowa.edu/Oiko/Downloads.html>

¹⁴<https://genome.jgi.doe.gov/Brafl1/Brafl1.home.html>

¹⁵ftp://ftp.ncbi.nlm.nih.gov/genomes/Saccoglossus_kowalevskii/protein/protein.fa.gz

¹⁶http://www.echinobase.org/Echinobase/PmDownload/pmin_proteins_v2.0.fa

¹⁷ftp://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