VENOMOUS LIAISONS: INVESTIGATING THE MOLECULAR BASIS OF THE TROPHIC RELATIONSHIP BETWEEN CORALS AND CORALLIVOROUS SNAILS

PhD Project, Elisa Nocella
Curriculum: Animal Biology

Many marine organisms subdue preys or deter predators using venoms, complex bioactive secretions mostly containing proteins and peptides, produced and delivered through specialized anatomical structures. Marine realm offers many examples of predator-prey interactions involving the use of venomous secretions, which in some cases are produced both in the predators and in the preys. Investigating diversification of toxin genes can shed light on the role of predator-prey interactions in generating species adaptations. The Cnidaria is a highly diversified group of marine venomous organisms, producing in their nematocysts and in other cell types a venom; despite their toxicity, they are consumed as food by several organisms, and in particular by multiple gastropod lineages. For example, the Caenogastropoda families Muricidae, Ovulidae, and Epitoniidae, are exposed to the cnidarian venom, since exploded nematocysts are commonly retrieved into their stomach, thus raising interesting questions on the mechanisms employed by the snails to counteract the effects of prey venom.

The overarching aims of this PhD project are to elucidate the level of specificity, evolutionary patterns and underlying molecular mechanisms, at the same time highlighting bioactive peptides and proteins worth of subsequent biotechnological development.
In particular, two main research purposes have been set:

- To understand the evolution of corallivory in the different snail lineages, mapping trophic ecology as a trait on the phylogeny, reconstructing ancestral trophic ecology for each lineage, dating the acquisition of selected trophic adaptations, and testing the correlation between trophic specialization and phyletic diversification.
- To identify genes expressed in tentacles of Anthozoa, and in the salivary glands of the corallivorous snails to better define the potential activity and the molecular interactions of bioactive peptides and proteins.

Especially, this PhD project will be focused on three lineages of caenogastropod corallivores: the Coralliophilinae Muricidae (that can feed both on all Anthozoans), the

Ovulidae,(associated to soft corals) and the Epitoniidae. (on stony corals only). These snails are equipped with one or two pairs of salivary that discharge at the proboscis tip, and are likely to play a role in feeding, e.g. chemically inactivating prey venom components. Beside a few reports on a reduced number of species, generally no information is available on the degree of species specificity of predator-prey relationship, especially for deep-water taxa, since they are generally collected through dredging or trawling, causing the detachment of the snails.

A clear definition of the species specificity of the trophic associations between corallivorous snails and their preys is a fundamental prerequisite to understand its evolutionary significance and evaluate the potential diversity of the underlying bioactive secretions. In this PhD, the trophic association between snails and corals will be indirectly inferred through DNA barcoding of snails' stomach content. Despite the identification of the corals at the species level can be limited by the unavailability of reference material in sequence databases, molecular species delimitation tools can be applied to delimit species boundaries and count the number of MOTUs, without hampering the expected outcome (identification of prey diversity). Dissections of stomach and DNA-barcoding will be carried out on a range of species as broad as possible, and on multiple ethanol preserved specimens, to evaluate the degree of polyphagy in opposite to specificity of corallivorous snail.

We will rely mostly on the outstanding collection of the Museum National d'Histoire Naturelle (MNHN Paris, France), which includes about 900 ethanol-preserved specimens of Coralliophilinae Muricidae, 1300 of Ovulidae, 800 of Epitoniidae, with a high proportion of deep-water material. This material will be integrated through loans from collaborating institutions. the sampling of shallow-water species for RNA analysis will take place along the Italian Mediterranean coasts.

The first year will be dedicated to stomach dissections from suitable specimens, DNA extraction using the EZNA Mollusc DNA Kit (Omega BioTek), and DNA barcoding of stomach content using the ITS2 marker that will be amplified using universal primers. This marker was chosen since it has been demonstrated to have a different length in cnidarian vs. molluscs, allowing an easy isolation. The obtained sequences will be blasted against NCBI nucleotide database to identify stomach content at the lowest taxonomic rank possible (down to species level, when reference sequences will be available). Sequences will also be aligned used MAFFT algorithm available in Geneious 11, and the alignment will be used to delimit species through both similarity-based approaches. The Shannon diversity index will be calculated to estimate dietary breadth in polyphagous species.

Sampling, RNA/protein extraction and A HTS RnaSeq approach will be carried out during the second year of the PhD, while the third year will be dedicated to comparative evolution analyses and evaluation of biotechnological potential. To evaluate the degree of

convergence in salivary biochemical specialization in the different lineages of corallivore caenogastropods, we will focus on shallow-water species, chosen in order to maximize trophic overlapping: *Simnia spelta* (Ovulidae) and *Coralliophila brevis* (Muricidae Coralliophilinae) on the octocoral *Eunicella singularis*; *Epitonium commutatum* (Epitoniidae) and *Coralliophila meyendorffi* on the hexacoral *Anemonia sulcata*. These predators and preys are locally common and easy to collect in the Mediterranean Sea. Since transcriptomes and proteomes of *C. brevis* and its prey *E. singularis* are already available, during this PhD 8 further tissue-specific transcriptomes and proteomes will be sequenced. We will focus on the salivary glands, since preliminary work on *C. meyendorffi* has revealed a highly prey-specific gene expression in these tissues. Additionally, salivary glands are well developed in all the target lineages and correspond to homologous tissues, compared with other foregut glandular structures which are reduced in some of the target lineages and whose homology is often unclear.

Host corals transcriptomes and proteomes will also be investigated, to elucidate the nature of the adaptive forces acting at the molecular level on salivary secretions of the corallivorous snails. A HTS RnaSeq approach will be used to identify genes expressed in tentacles of Anthozoa, and in the salivary glands of the corallivorous snails. The implementation of a differential expression analysis will allow for the comparison between secretory tissues and the whole body for each species, identifying enriched tissue-specific transcripts even when they do not display similarities with known venom proteins. To complement transcriptomic data, a proteomic approach involving LC-MS and LC-MS/MS spectrometry will be pursued, allowing for the confirmation of assembled NGS sequences, the detection of post-translational modification (PTMs) of primary transcripts and for a more complete evaluation of venom complexity. Snail salivary glands (SG) and cnidarian tentacles (TE) will be dissected from 6-12 specimens of the target species kept in an aquarium at the IBMM animal facility and pooled in 3 replicates for RNA extraction. Additionally, RNA will be extracted from 3 replicates of one whole-body (WB) samples each. RNA will be used to construct barcoded libraries that will be sequenced in MinION device at Sapienza University of Roma. Raw reads will be processed and assembled using standard bioinformatics analyses. The assembly will be subjected to similarity search against online databases using the BLAST algorithm. The obtained final set of sequences will be used as reference for read mapping and differential expression (DE) analyses that will compare the target tissues (SG for snails, TE for cnidarians) and the WB in a pairwise manner; functional annotation of the DE transcripts will be carried out in Blast2GO. Transcripts significantly enriched in target tissues compared with WB will be singularly screened even when they will not display similarity with known proteins and manually annotated. We will rely on the support of the SZN BIOINFORMA platform for bioinformatics analyses. Proteomic analyses will be carried out at the proteomic platform

FPP (Montpellier, France). Liquid chromatography and tandem electrospray mass spectrometry analyses will be performed on a SYNAPT G2-Si High Definition Mass Spectrometry instrument (Waters). If necessary, additional MALDI experiments will be performed on an Ultraflex III (Bruker). Diluted samples of each species extracts will be directly subjected to LC-ESI-MS in order to obtain a complete mass list of underivatised peptides. Information Dependent Acquisition (IDA) will be performed on the reduced and reduced/alkylated venom samples. A sequence database comprising the entire raw cDNA reads from each transcriptome will be used to match MS/MS data with transcriptomic sequences utilizing PEAKS software. The detected peptide fragments will be manually inspected and validated. The output of these analyses will constitute the basis for reconstructing the molecular evolution of bioactive peptides. To highlight the role of convergence in shaping the biochemical profile of snail salivary secretion we will use a phylogenetic and comparative approach. Toxin genes identified in transcriptomic analyses will be attributed to gene families basing on their similarity with known proteins. Aminoacidic sequences of homologous toxin genes will be aligned, and their percentage sequence identity (PID) and similarity (PSI) will be calculated using Geneious 11 for each family. Orthologous genes will be identified via a phylogenetic approach. Toxin genes that in a ML phylogeny occur in highly supported clades including one sequence for each predator (and for each prey) will be considered orthologous. Tests for positive selection will be carried out estimating the pairwise non-synonymous substitutions per non-synonymous sites (dN) and synonymous substitutions per synonymous sites (dS) on the mature region of orthologous toxins using the ML approach implemented in PAML. Diversity of the toxin peptides for each family in the predator and prey species will be calculated using a Shannon diversity index, while the evenness of the toxin peptides will be calculated for each family using Shannon equitability. Expression levels will be calculated for each toxin family in each predator/prey species, as the sum of transcript-specific TPM values obtained with RSEM. The biotechnological potential of selected peptides and proteins will be evaluated through homology with molecules with known activity, and the potential interaction between coral toxins and snail proteins/peptide will be inferred through *in silico* molecular docking approaches.

Below the Gantt chart: