Stage[1] de Master (M2)

Stylins Evolution and Diversity in Hemipterans

Encadrants : Stefano Colella (PHIM, équipe MOVE, CR INRAE)
Emmanuelle Joussellin (CBGP, DR INRAE)

Période (Durée) : à partir de Janvier/Février 2023 (jusqu'à 6 mois)

Context

Aphids are phloem-feeding hemipterans that constitute a major agricultural pest. They are vectors of around 30% of known plant viruses and are the only ones reported to date able to transmit hundreds of non-circulative stylet-borne viral species. Aphids acquire viruses while feeding on plants using their specially adapted mouthparts, called stylets to reach the phloem without damaging the tissues. In particular, the "acrostyle" structure on the apical part of the maxillary stylets has been demonstrated to interact with transmitted non-circulative viruses via cuticular proteins (1). To understand and characterise the molecular mechanisms of non-circulative virus transmission, and particularly virus-vector interactions, several approaches have been used by the MOVE team **to identify virus receptors in the aphid vector**. Currently, the team has characterised **five cuticular proteins named stylins at the virus-vector interface, all receptor candidates of plant viruses**(2). **An investigation of the evolution of stylins throughout** the **diversification of Aphididae** will advance our understanding of the vectoring capacity of aphids. Though cuticular proteins are known to be relatively conserved in arthropods, aphids are the only ones known to retain viruses on the tip of their stylets thanks to these proteins. Furthermore, not all aphid species are virus vectors, and these differences in vectoring capacity might be linked to variation in the repertoire of stylins.

Project Aims

This project is the first step of a larger project aiming to explore the existence and evolutionary history of the cuticular proteins, stylins in particular, in all insects with piercing mouthparts and transmission activity. The availability of multiple insects' genomes offers the possibility to explore the evolution and diversification of cuticular proteins. We will focus here on aphids and related Hemiptera (Psyllidae and Adelgidae) with a double goal. The first aim is to **identify stylins orthologs in sequenced hemipteran genomes and perform a phylogenetic analysis** to characterise their evolution and diversification. A second aim is to **use the sequence data in aphid lineages to design a wet-lab approach to extend the analysis in the Aphididae family by sequencing these genes representatives of aphid main subfamilies available in the INRAE-CBGP aphid collection**[2] (3).

Methods

The project's starting core will include a data mining step, followed by phylogenetic sequence analyses. First, we will BLAST known stylin gene sequences onto available hemipteran genomes with a focus on aphids (21 genomes at the time of writing the project) to retrieve all potential orthologs. In parallel, we will use the MetaPHorS database[3] (4) to explore established orthology relationships for a subset of available genomes. Collected nucleic and amino acid sequence data will be aligned using the MAFFT multiple alignment program

[1] Financement pour étudiants d'établissement d'enseignement supérieur en région Occitanie - https://rivoc.edu.umontpellier.fr/;
[2] https://doi.org/10.15454/D6XAKL; https://aphiddb.supagro.inrae.fr/
[3] http://orthology.phylomedb.org/

v7[4] and the L-INSI method. We will then conduct phylogenetic inferences using both Bayesian approaches (with MrBayes v3.2.) (5) and the maximum likelihood estimation method, implemented in IQ-TREE v1.6.2 (6). We will also inspect gene organisation and sequence duplication in genomes that are well assembled.

Results of the first part of the project will be used to choose and design the most appropriate method to extend the analysis of stylins in the Aphididae subfamilies (using aphid species from CBGP collection resources). We will thus concentrate on the aphid gene sequences. Two potential approaches can be developed to obtain DNA material for high-throughput sequencing of genes of interest: a PCR amplification using degenerated primers or a probe-based capture using a conserved sequence region. We will use the collected sequence data to evaluate the adequacy of these two experimental designs and choose the best strategy. The choice will depend on the possibility of degenerated primers design, the number of gene copies for stylins and the variation within sequences that will determine the design of probes. Finally, depending on the project advancement, we will perform a first wet-lab test with the chosen technique (*i.e.* prepare DNA libraries for NGS sequencing with a Miseq platform available on the Cemeb platform).

Expected Results

The project expected outputs include: the development of a sequence retrieval and storage strategy, the phylogenetic analysis of stylins in Hemiptera, the choice of the best wet-lab strategy to extend the phylogenetic analysis in the aphid lineage and possibly tests of the sample preparation for NGS sequencing-based analysis of stylins in the Aphididae superfamily. Altogether the project will constitute the first step to deciphering how stylin genes vary across aphid species and whether this diversification is associated with known variations in vectoring capacities.

Bibliography

1. Webster,C.G., Pichon,E., Munster,M. van, Monsion,B., Deshoux,M., Gargani,D., Calevro,F., Jimenez,J., Moreno,A., Krenz,B., *et al.* (2018) Identification of Plant Virus Receptor Candidates in the Stylets of Their Aphid Vectors. *J Virol*, 92, e00432-18.
2. Deshoux,M., Masson,V., Arafah,K., Voisin,S., Guschinskaya,N., Munster,M. van, Cayrol,B., Webster,C.G., Rahbé,Y., Blanc,S., *et al.* (2020) Cuticular Structure Proteomics in the Pea Aphid Acyrthosiphon pisum Reveals New Plant Virus Receptor Candidates at the Tip of Maxillary Stylets. *J Proteome Res*, 19, 1319–1337.
3. d'acier,A.C., Cruaud,A., Artige,E., Genson,G., Clamens,A.-L., Pierre,E., Hudaverdian,S., Simon,J.-C., Jousselin,E. and Rasplus,J.-Y. (2014) DNA Barcoding and the Associated PhylAphidB@se Website for the Identification of European Aphids (Insecta: Hemiptera: Aphididae). *Plos One*, 9, e97620.
4. Chorostecki,U., Molina,M., Pryszcz,L.P. and Gabaldón,T. (2020) MetaPhOrs 2.0: integrative, phylogeny-based inference of orthology and paralogy across the tree of life. *Nucleic Acids Res*, 48, W553–W557.
5. Huelsenbeck,J.P., Ronquist,F., Nielsen,R. and Bollback,J.P. (2001) Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, 294, 2310–2314.
6. Minh,B.Q., Schmidt,H.A., Chernomor,O., Schrempf,D., Woodhams,M.D., Haeseler,A. von and Lanfear,R. (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol*, 37, 1530–1534.

Required Skills

- Basic knowledge of genetics and genomics (sequence analysis skills will be a plus).
- Interest is performing *in silico* data analysis and pipeline development.
- Good interpersonal skills and autonomy
- English language skills

Conditions d'accueil et gratification

ATTENTION : sont éligible seulement les étudiant(e)s inscrits dans un établissement d'enseignement supérieur implanté en région Occitanie. Gratification selon taux en vigueur à INRAE. Le stage se déroulera dans l'UMR PHIM sur le Campus CIRAD de Baillarguet (https://umr-phim.cirad.fr/contact-acces/campus-de-baillarguet)

Modalités de candidature

Adresser par mail une lettre de motivation, un CV et éventuellement une/des lettres de recommandation à Stefano Colella (stefano.colella@inrae.fr) & Emmanuelle Jousselin (emmanuelle.jousselin@inrae.fr).

[4] https://mafft.cbrc.jp/alignment/server/