Ecole Doctorale COMPLEXITE DU VIVANT – Fiche Projet CONCOURS

Nom et prénom du directeur de thèse: Olivier VALLON
Coordonnées Tel : 01 5841 5058 e-mail : ovallon@ibpc.fr

Y-a-t-il un candidat déjà identifié pour le projet: OUI

Nom et prénom du responsable de l'équipe : Angela FALCIATORE

Intitulé de l'équipe : Biologie du chloroplaste et perception de la lumière chez les micro-algues.
Nombre de chercheurs et enseignants-chercheurs statutaires de l'équipe titulaires d'une HDR (ou équivalent) : 7

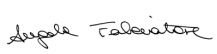
Nom et prénom du responsable d'UMR ou de département: Angela FALCIATORE

Intitulé et N° d'UMR ou de département : UMR7141

Titre du projet de thèse *: Chloroplastic OPR-RAPs as sequence-specific endoribonucelases / Les OPR-RAPs du chloroplast comme endoribonucléases séquence-spécifiques*

Spécialité : Biologie moléculaire

Signature du directeur d'UMR ou de département (vaut avis favorable pour le dépôt du projet) :



Résumé du projet de thèse *(1 page maximum, en anglais)*

Plastid endosymbiosis, whereby a cyanobacterium was engulfed into a eukaryotic cell and made it phototrophic, was the starting point of a long evolutionary history of gene loss and gene transfer to the nucleus. Today, the tiny chloroplast (Cp) genome only encodes about 100 proteins, including subunits of the photosynthetic enzymes. While the basic gene expression machinery of the Cp has kept many ancestral prokaryotic traits, gene regulation now depends on nucleus-encoded factors acting post-transcriptionally by binding in a sequence-specific manner to Cp transcripts. Stabilization of the mRNAs depends on M-factors, while translation initiation requires T-factors. In our laboratory, genetic screens using the unicellular green alga Chlamydomonas reinhardtii have unraveled many such factors. They belong mostly to two unrelated families of modular proteins: the PPR and OPR (respectively: Penta- and Octo-tricoPeptide Repeat) families. While the former predominates in higher plants (450 PPRs in Arabidopsis), the latter is more developed in green algae (138 OPRs in *C. reinhardtii*). In both families, tandem repeats of a degenerated motif (of 35 and 38 residues, respectively), folded as a pair of antiparallel α-helices, stack to form a super-helix that accommodates the target RNA in its inner groove. Each motif interacts with one nucleotide via specificity-determining amino acids at defined positions within the repeat. In the case of PPRs, a recognition code has been proposed that links the sequence of the repeat with the nucleotide it binds to. For OPRs, Yves Choquet in our lab has established a preliminary version of the code, based on a few M-factors for which the target is known.

The project deals with a specific class of OPRs that harbor at their C-terminus a RAP domain. This domain of unknown function is structurally similar and probably homologous to the VsR and DUF559 endonucleases. We have obtained dominant missense mutations in the OPR domain of two such OPR-RAPs, NCC1 and NCC2. The mutations affect residues that we suspect control binding specificity. In the *ncc1* and *ncc2* mutants, respectively the *atpA* and *petA* mRNAs are destabilized by binding of the mutant proteins. We propose that OPR-RAPs act as endoribonucleases, binding to the target mRNA via their OPR domain and cleaving it downstream using their RAP domain. Interestingly, most of the OPR-RAPs in Chlamydomonas belong to the NCL subfamily, a series of 48 closely-related genes found in a large cluster on chromosome 15. This cluster has evolved rapidly by vigorous gene duplication, neofunctionalization and pseudogenization, and analysis of synonymous and non-synonymous substitutions shows that a strong positive (= diversifying) selection is exerted on the residues that we think specify the bound nucleotide. In addition, mapping of Illumina sequences obtained from a variety of

wild interfertile isolates of *C. reinhardtii* shows numerous sequence variants in *NCL* genes, suggesting that diversification is also favored in natural populations. An unusually high variability is also observed in the Cp genomes of these wild isolates.

The first objective of the thesis will be to characterize the mechanisms of target recognition by OPR-RAPs and extend the "OPR code". The student will first establish fluorescent reporter systems, in the Chlamydomonas chloroplast and in the *E. coli* cytosol, meant to rapidly measure the efficiency of target cleavage by an NCL. The target site will be placed in the mRNA coding for the fluorescent protein, in such a way that cleavage will abolish expression. A Cp-encoded mCherry construct is already available, and the heterologous GFP-based *E. coli* reporter system is being developed in the lab. Using synthetic biology tools such as Golden Gate cloning and MoClo toolkits, the student will generate a panel of transgenic algae and bacteria expressing NCL proteins of native or tailored specificity. The latter will include variants of a fully synthetic NCL, comprising 10 repeats of the OPR consensus motif. Once the interaction between an NCL and its target has been reconstituted, mutations will be introduced in the target to disrupt the interaction and abolish cleavage. Then complementary mutations will then be introduced in the OPR, as predicted by the code, to restore the interaction.

If time allows, the student will also explore the biological function of the NCL cluster. Our current hypothesis is that OPR-RAPs differ in target specificity between populations of *C. reinhardtii*: this could cause nucleo-chloroplastic incompatibilities, constituting a post-meiotic barrier to genome admixing in the wild. The student will analyze the meiotic progenies obtained by crossing laboratory strains holding appropriate genetic markers with field isolates showing interesting polymorphisms in the predicted specificity of their NCL genes. In a cross, the Cp genome will be inherited solely from the mt+ parent, while the NCL cluster will segregate 2:2 in a Mendelian fashion. Genetic incompatibilities will be identified by fitness defects and state-of-the-art chlorophyll *a* fluorescence imaging which can quickly reveal even subtle changes in the photosynthetic apparatus. Incompatibilities due to NCL genes will be identified by testing co-segregation with the resistance markers. Their molecular basis will be established using RNA-Seq and validated by reconstitution in the reporter systems.

Faisabilité du projet de thèse (*1/2 page maximum, en anglais)*

The thesis is tightly connected to major scientific projects currently led in our lab. The student will collaborate closely with Y. Choquet, and with another PhD student already working on the ANR-funded ARAMIS project aimed at deciphering the mechanisms of gene expression in the Chlamydomonas mitochondrion, in particular with respect to OPRs. In addition, an French-German ANR/DFG project led by O. Vallon is under review, aimed at addressing the functions of OPRs in 4 model organisms. O. Vallon is also leading a Community Science Project with the Joint Genome Institute, which will assemble and annotate the genomes of 20 field isolates of *C. reinhardtii* so as to fully describe the intraspecific diversity of this premier model organism. In addition to its interest for population genetics, this "Chlamydomonas Pan-Genome" project will reveal a wide panel of NCL variants and guide the genetics studies described above.

This highly-disruptive project presents the student with an opportunity to contribute to a major breakthrough, the establishment of OPR-RAPs as a novel type of RNases. The modular nature of the OPR domain renders them potentially useful as sequence-specific RNA modifiers, from the design of genetic switches for synthetic biology to diagnostic and therapeutic applications. The student will have access to all the necessary equipment for genetic manipulations and molecular/functional analyses. The nurturing pluridisciplinary environment of our lab, structured as a single research team, should be attractive to individuals with high scientific ambitions.

Thèses actuellement en cours dans l'équipe

Tous les encadrements doivent être indiqués (y compris les co-directions avec un autre HDR pour des doctorants d'une autre ED, et les encadrements dans le cadre de programmes doctoraux tels qu'IPV, FDV...)

Ecole Doctorale COMPLEXITE DU VIVANT – Fiche Projet CONCOURS

Nom et Prénom du doctorant	Directeur(s) de thèse	Année de 1ère inscription	ED	Financement
Amel Zaoui	Olivier Vallon	2021	515	ANR
Mathieu Mustas	Yves Choquet	2018	515	Labex
Marcio Rodrigues	Catherine de vitry	2018	SdV	ministère
Croteau Dany	Benjamin Bailleul	2020	515	ERC
Céline Cattelin	Ingrid Lafontaine	2019	515	80prime
Alessandre Manzotti	Angela Falcialtore	2018	515	ministère
Carole Duchène	Angela Falcialtore	2018	515	ANR

Trois publications récentes du directeur de thèse (du co-directeur ou du co-encadrant s'il y a lieu). Mettre en gras le nom du directeur de thèse.

1. Chaux-Jukic, F., S. O'Donnell, R.J. Craig, S. Eberhard, **O. Vallon**, and Z. Xu (2021) Architecture and evolution of subtelomeres in the unicellular green alga Chlamydomonas reinhardtii. Nucleic Acids Res. 49(13): p. 7571-87.
2. Salinas-Giégé, T., M. Cavaiuolo, V. Cognat, E. Ubrig, C. Remacle, A.M. Duchene, **O. Vallon**, and L. Marechal-Drouard (2017) Polycytidylation of mitochondrial mRNAs in Chlamydomonas reinhardtii. Nucleic Acids Res. 45(22): p. 12963-73.
3. Cavaiuolo, M., R. Kuras, F.A. Wollman, Y. Choquet, and **O. Vallon** (2017) Small RNA profiling in Chlamydomonas: insights into chloroplast RNA metabolism. Nucleic Acids Res. 45(18): p. 10783-99.