



**3IEME COLLOQUE DE L'INFRASTRUCTURE INRAE GENOMICS**  
18, 19 et 20 juin 2025



Hôtel Novotel Clermont Ferrand

32-34 rue Georges Besse, Le Brezet  
63100 CLERMONT FERRAND  
FRANCE  
GPS:45.786567, 3.14186

[colloque-inraegenomics@inrae.fr](mailto:colloque-inraegenomics@inrae.fr)

# PROGRAMME

## MERCREDI 18 JUIN 2025

---

12H – 13H30 DEJEUNER (BUFFET)

14H – 14H30 PRESENTATION DE L'INFRASTRUCTURE INRAE GENOMICS

14H30 – 16H SESSION 1 - EXPLORATION DE LA DIVERSITE : WHOLE GENOME

- PROPOSITIONS DES PLATEFORMES DANS CETTE THEMATIQUE
- Davide Degli Esposti : The genome of *G. fossarum*: a genomic toolkit to investigate endocrine and metabolic disruption in amphipod sentinel species
- Ludovic Duvaux : Apport du séquençage longues lectures pour la pangénomique des chênes blancs européens
- Aurélie Manicki : Diversité truite des Kerguelen (approche Pool-seq MGI et low input Illumina)

16H – 16H30 PAUSE

16H30 – 18H00 SESSION 1 SUITE - EXPLORATION DE LA DIVERSITE : WHOLE GENOME

- Christophe Klopp : Assemblage de génomes auto-tétraploïdes et création de pangénomes
- Olivier Garsmeur : Inferring genomic ancestries using repeated k-mers from whole-genome sequencing data: Application to polyploid Sugarcane
- Hélène Rimbert : LowPass long read sequencing blé Revio Pacific Biosciences

19H30 DINER

## JEUDI 19 JUIN 2025

---

8H30 – 10H30 SESSION 2 - EXPLORATION DE LA DIVERSITE : GENE

- PROPOSITIONS DES PLATEFORMES DANS CETTE THEMATIQUE
- Frédéric Choulet : Génomique de l'immunité chez le blé par l'analyse de données transcriptomiques Pacbio Kinnex
- Nathalie Boissot : Nanopore Adaptive sampling to identify the NLR-gene family in melon (*cucumis melo* L.)
- Raphael Minguella : Estimation des fréquences haplotypiques à partir de tGBS en pool
- Guillaume Besnard : Travaux sur le locus d'auto incompatibilité

10H30 – 10H50 PAUSE

10H50 – 12H30 SESSION 3 - METAGENOMIQUE

- PROPOSITIONS DES PLATEFORMES DANS CETTE THEMATIQUE
- Yanis Bouchenak-Khelladi : Towards evidence for plant-bacteria coevolution: an ecophylogenetic framework reveals selection through niche construction
- Bérangère Baillet : Projet LAMINET : L'ADNe pour décrypter la structure trophique des forêts de Laminaires
- Thibault Leroy : Le séquençage des miels, un regard génomique sur l'agroécologie?

12H30 – 14H00 DEJEUNER

14H00 – 17H SESSION 4 - EXPLORATION DE LA DIVERSITE : LOCUS

- PROPOSITIONS DES PLATEFORMES DANS CETTE THEMATIQUE
- Grégoire Aubert : Développement d'une puce SNP multi-espèces pour caractériser la diversité génétique chez les protéagineux (pois, féverole, lupin blanc)
- Kritika Adhikari : Assessing the diversity of the Espelette pepper landrace in the French Basque country

15H15 – 15H35 PAUSE

- Antoine Bacon : Capture de région ciblée en Twist séquençage PacBio
- Sabrina Delaunay : Comment les marqueurs KASPAR permettent de suivre l'introgression de la fétuque glauscence dans le RGI
- Etienne Delannoy : étude combinant transcriptomique et métabolomique sur la transition évolutive vers la perte de la photosynthèse chez les orchidées européennes

17H – 19H30 TEMPS LIBRE (PISCINE, JACUZZI...)

19H30 DINER AVEC ANIMATIONS SURPRISES

**9H00 – 10H30 SESSION 5 - REGULATION DE L'EXPRESSION DES GENOMES**

- **PROPOSITIONS DES PLATEFORMES DANS CETTE THEMATIQUE**
- Hervé Acloque : Single-cell omics reveal distinct gene regulatory dynamics underpinning embryonic and extraembryonic lineage functions during pig blastocyst development
- Julie Demars : Discovering parent-of-origin expression and methylation in the pig highlights evolutionary and molecular mechanisms shaping genomic imprinting

10H30 – 11H00 PAUSE

**11H00 – 12H SESSION 6 - QUELS FUTURS POUR LA GENOMIQUE POUR LES AGROBIOSCIENCES ?**

*Retour sur l'enquête France Génomique 2024 Denis Milan*

*iA : utilisation de l'iA en génomique par Mahendra Mariadassou (visio)*

*Discussion*

12H00 – 14H DEJEUNER

**FIN DU COLLOQUE**

# ACCES AU COLLOQUE

## LOCALISATION

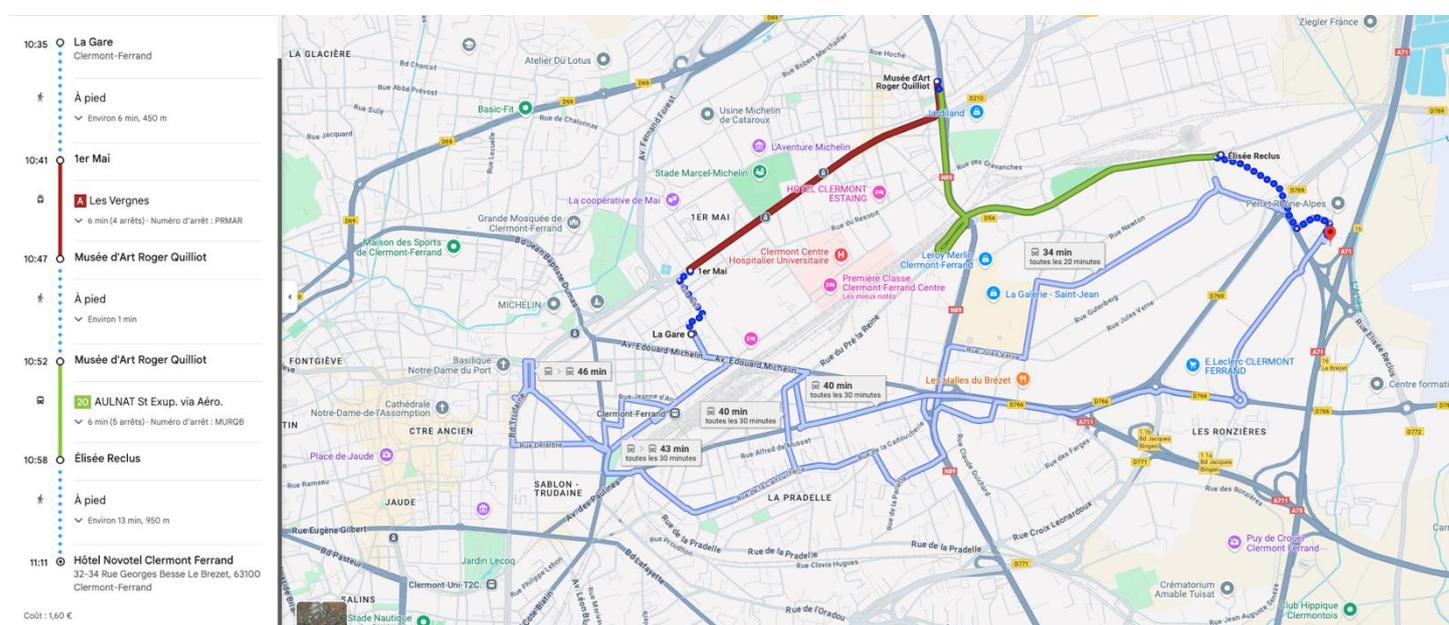
Hôtel Novotel Clermont-Ferrand  
32-34 rue Georges Besse, Le Brézet  
63000 CLERMONT-FERRAND  
France  
Coordonnées GPS:45.786567, 3.14186

## TRANSPORTS

Gare de Clermont-Ferrand à 4 km, Aéroport à 2 km (taxi)

Depuis la gare~36 min en transport en commun dont 14 min à pied.

<https://www.t2c.fr/votre-trajet/votre-itineraire>



## INFORMATIONS UTILES

### ACCUEIL ET REMISE DES CLES

L'accueil et la remise des badges se feront dans le hall de l'hôtel à partir de 11h30.

La remise des clés des chambres se fera en fin d'après-midi, les bagages pourront être laissés en attendant dans une salle dédiée.

Le parking de l'hôtel est accessible aux participants (gratuit).

### REPAS

Les repas et les pauses sont pris en charge du déjeuner du mercredi 18 midi au déjeuner du vendredi 20 inclus.

Le déjeuner du mercredi 18 juin est servi sous forme d'un buffet accessible de 12h à 13h30.

Les consommations hors repas et pauses sont à la charge du participant et sont à régler directement à l'hôtel.

### CONTACT

[inraegenomique@colloque-inrae.fr](mailto:inraegenomique@colloque-inrae.fr)

## AVEC LA PARTICIPATION :

De la **Délégation aux Infrastructures Scientifiques Collectives d'INRAE (DISC)**

Des départements INRAE :

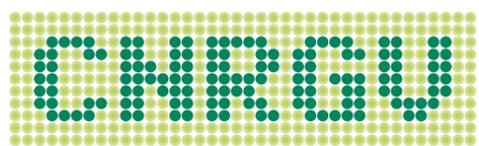
**Biologie et Amélioration des Plantes (BAP)**

**Ecologie et biodiversité des milieux forestiers, prairiaux et aquatiques (ECODIV)**

**Génétique Animale (GA)**

Des **plateformes INRAE GENOMICS**

**INRAE  
GENOMICS**

  
**PLANT GENOMIC CENTER**

  
**GENTYANE**

**EPGV**  
Etude du Polymorphisme  
des **G**énomes **V**égétaux

 **PGTB**  
SEQUENCAGE | GENOTYPAGE

 **Genotoul  
GeT**

**MERCI A TOUS :**  
**ORATEURS, ANIMATEURS ET**  
**PARTICIPANTS !**

# RESUMES DES PRESENTATIONS

## SESSION 1 - EXPLORATION DE LA DIVERSITE : WHOLE GENOME

### **The genome of *G. fossarum*: a genomic toolkit to investigate endocrine and metabolic disruption in amphipod sentinel species**

**Davide Degli Esposti<sup>1</sup>, Damien Hinsinger<sup>2</sup>, João Sousa<sup>1,3</sup>, Patrice Baa-Puyoulet<sup>4</sup>, Aurélie Berard<sup>2</sup>, Raquel Ruivo<sup>3</sup>, Teresa Neuparth<sup>3</sup>, Patricia Faivre-Rampant<sup>2</sup>, Federica Calevro<sup>4</sup>, Hubert Charles<sup>4</sup>, Arnaud Chaumot<sup>1</sup>, Olivier Geffard<sup>1</sup>**

1. INRAE, UR RIVERLY, Villeurbanne, France
2. Université Paris-Saclay, Centre INRAE Île-de-France Versailles-Saclay, EPGV, Evry, 91057, France
3. Interdisciplinary Centre of Marine and Environmental Research (CIMAR/CIIMAR), University of Porto, Matosinhos, Portugal
4. INRAE, INSA Lyon, BF2I, UMR203, 69621 Villeurbanne, France

**Background/Introduction.** Amphipods, small crustaceans commonly found in freshwater and marine habitats, play a critical role in nutrient cycling, sediment turnover and as a food source for higher trophic levels. Among amphipods, the Gammaridae represent a large family of species, some of which, such as *Gammarus fossarum*, are widely used as sentinel species for assessing water quality and as animal models for ecotoxicology. While some transcriptomes are available for several gammarid species, the lack of good quality genomic resources is a major research gap that hinders our understanding of their physiology and response to aquatic pollution. Here, we report the first contiguous assembly of the genome of a single amphipod of the species *Gammarus fossarum* and the first insights from this unique genomic resource for the family Gammaridae.

**Methods.** This genome project was designed to take advantage of recent advances in high-throughput long-read sequencing technologies and to streamline the bioinformatics assembly pipeline. High-molecular-weight DNA extracted from a morphologically determined male *Gammarus fossarum* was used for subsequent sequencing. A single molecule real time (SMRT) sequencing platform (PacBio) was used to sequence three different libraries on three HiFi SMRT cells. A first assembly was performed using the Hifiasm de novo assembler.

**Results.** The main results of the assembly indicate a haploid genome size of 3.2 Gb with a sequencing depth of 40x. The N50 is 239 kb and the number of contigs is 50,732, with a BUSCO completeness value of 88.3%, indicating the most contiguous gammarid genome. Phylogenetic analysis comparing the COI sequences of several populations of continental gammarid populations confirmed that the organism belonged to a *Gammarus fossarum* population (subgroup B). Here, we present the results of the systematic annotation of the metabolic pathways obtained using the CycADS pipeline and the identification of the genes belonging to the ecdysone signalling axis.

**Conclusion.** We provide a first high quality genome for the species *Gammarus fossarum*. This resource will be useful for both ecological and ecotoxicological communities, allowing a better characterisation of the diversity of gammarid populations and the study of metabolic and endocrine disruption caused by chemical contaminants.

## **LUDOVIC DUVAUX : CONTRIBUTION OF LONG-READ SEQUENCING TO THE PANGENOMICS OF EUROPEAN WHITE OAKS**

The 15 species of European white oaks began to diverge more than 10 million years ago and are ecologically, phenologically, and morphologically distinct. However, most of them still interbreed, forming a species complex with incomplete reproductive isolation. The genetic basis of this isolation has been studied using comparative genomics approaches, but neither major regions of differentiation nor diagnostic SNPs have been detected. Recent literature has shown the involvement of structural variants (SV), particularly inversions, in reproductive isolation by reducing local recombination. We therefore adopted a pangenome variation graph approach to study the role of SV in the divergence of European white oaks.

We assembled the genomes of 16 individuals — 12 using PacBio HiFi and 4 using ONT data — belonging to 8 different (sub)species, yielding a total of 28 haplotypes. These haplotypes were then used to construct a pangenome variation graph of the European white oaks using PGGB. Visual inspection and distance matrices derived from graph nodes revealed contrasting chromosome behaviours with respect to SV content and species discrimination. The difference is particularly striking and representative between the two large chromosomes 2 and 4. Chromosome 2 has a low SV content, a large proportion of core nodes, and strongly discriminates species — carrying many of the rare known near-diagnostic SNPs. In contrast, chromosome 4 has a much higher SV content and shows little to no resolution for species discrimination.

The lowest SV density observed in the most discriminating chromosomes suggests reduced local recombination rates and/or stronger selective constraints. These hypotheses will be tested by analyzing linkage disequilibrium, genetic divergence, and recombination rates along chromosomes, particularly around SV. Annotation work is underway to examine the gene and transposable element content of candidate regions and identify potentially involved biological functions. Finally, to interpret these results in the macroevolutionary context of long-lived species, we are currently developing a coalescent-based population pangenome simulator. This simulator will allow us to test various demographic and historical speciation scenarios for the species complex, while explicitly modeling the role of SVs in the process.

**A Manicki, G. Brahy, C. Klopp, A. Berard, D. Hinsinger, M. Minjou, C. Ducos, F. Gueraud, S. Muratorio.**

### **Diversité et structure génomique d'une population de truite en expansion aux Iles Kerguelen : mise au point d'une approche de Poolseq**

#### **Résumé :**

Notre étude vise à comprendre comment les processus démographiques et évolutifs (dérive génétique, migration, sélection, mutation) façonnent la diversité génomique et sa structure dans une population de truite commune (*Salmo trutta*) en expansion. Introduite aux Iles Kerguelen dans les années 1960, *S. trutta* a aujourd'hui colonisé 35 cours d'eaux et continue encore son expansion. Ces populations font l'objet d'un suivi à long terme depuis 1962 avec des campagnes d'échantillonnage régulières, qui nous permettent de disposer d'une riche collection de plus de 50 000 échantillons (écailles) répartis le long du front d'expansion sur un gradient temporel de 50 ans.

Nous avons séquencé 13 pools d'individus (25 individus/pool), correspondant à huit populations, dont 3 populations anciennes introduites dans les années 60 et échantillonnées sur 2 à 4 dates, et 3 populations récentes, ayant colonisé des cours d'eaux dans les années 2000-2010 et échantillonnées une seule fois. L'ADN a été extrait à partir de cellules épidermiques fixées aux échantillons d'écailles d'alevins âgés de 1 an. Le séquençage a été effectué sur séquenceur MGI (2 x 150 bp, objectif 50X de profondeur) par l'EPGV avec une étude préliminaire visant à savoir si la préparation de la librairie avec et sans PCR affecte le séquençage. La préparation des librairies a également été adaptée aux pools anciens composés d'ADN dégradés et peu concentrés, jusqu'à 2 ng/μL.

Les premiers résultats des analyses bioinformatiques montrent que la préparation des librairies avec PCR permet d'obtenir 10 X de profondeur supplémentaire sans réduire la performance du séquençage (couverture, duplication). Ce gain n'est pas négligeable notamment pour les pools anciens. Sur les 13 pools, la profondeur varie de 35 à 62X, et le nombre de reads de 500 à 769 millions. L'étape de SNP calling est en cours.

Ainsi cette étude révèle que l'extraction d'ADN et le poolseq sont réalisables à partir d'échantillons d'écailles d'alevin, même ancien et dégradés, et permettent d'obtenir des données génomiques de qualité. Cette étude inclura à terme 32 pools supplémentaires répartis sur le même gradient spatio-temporel (45 pools au total). Les perspectives pour l'identification des SNPs sous sélections seront présentées.

## CHRISTOPHE KLOPP : ASSEMBLAGE DE GENOMES AUTO-TETRAPLOÏDES ET CREATION DE PANGENOMES

Dans le cadre du PEPR Agroécologie et numérique, au sein programme BReIF, nous avons mis en place une plate-forme de service pour aider les biologistes à assembler, annoter des génomes et construire des pangénomes pour leurs espèces d'intérêt. Les deux premiers génomes que nous avons traités sont la luzerne et le dactyle qui sont auto-tétraploïdes. L'autopolyploïdie est courante chez les plantes. Elle complique l'assemblage du génome à cause de la ressemblance partielle des différents haplotypes mais aussi à cause du manque de logiciels adaptés à ces configurations de génomes. La situation évolue et dans ses dernières versions hifiasm permet, en présence de données Hi-C, de produire différents haplotypes. Malheureusement ceux-ci sont souvent mal-équilibrés. Nous avons développé et vous présenterons une approche pour améliorer ces résultats. Nous avons aussi élaboré une chaîne de traitement pour simplifier la génération des pangénomes avec une interface de visualisation facilitant l'analyse de la qualité des pangénomes par les biologistes. Elles feront l'objet du second volet de la présentation.

## Inferring genomic ancestries using repeated k-mers from whole-genome sequencing data: Application to polyploid Sugarcane

**Olivier Garsmeur<sup>1,2</sup>, Simon Rio<sup>1,2</sup>, Nicolas Pompidor<sup>1,2</sup>, Anna Lipzen<sup>3</sup>, Catherine Hervouet<sup>1,2</sup>, Théo Durand<sup>1,2</sup>, Kerrie Barry<sup>3</sup>, Jeremy Schmutz<sup>3</sup>, Angélique D'Hont<sup>1,2</sup>**

<sup>1</sup>CIRAD, UMR AGAP Institut, F-34398 Montpellier, France

<sup>2</sup>UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

<sup>3</sup>DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

The origin of sugarcane (*Saccharum* sp.), along with the evolutionary history and taxonomy of *Saccharum* species, remain largely unresolved. This lack of knowledge is mainly due to the complex polyploid nature of their genomes and a reticulate evolution involving frequent interspecific hybridizations.

To characterize the contribution of the two wild *saccharum* species (*S. robustum* and *S. spontaneum*) to cultivated sugarcane, we generated whole-genome-sequencing (WGS) data for 305 representative accessions, including few closely related genera. We developed novel approaches based on repeated k-mers, used here as proxies for transposable elements (TEs). These approaches take advantage of the TE characteristics, particularly long terminal repeat (LTR) retrotransposons, which amplify and disperse across the nuclear genome, and thus become durable markers that reflect recent evolutionary history. These methods enabled us to identify large sets of repeated k-mers representative of genera, species or subgroups within species. These species- and subgroups-specific k-mers were aligned to the genome assembly of the sugarcane cultivar R570. The results showed that they covered the entire genome assembly and allowed to determine the origin of all chromosome segments. In addition, we could verify that most of these repeated k-mers (95%) corresponded to annotated LTR-retrotransposons. These k-mers were also used to quantify the contribution of wild species and subgroups to all hybrids in our sample. Interestingly, this method allowed to identify contributions from an unknown ancestor for which we had no pure accessions in our sample.

Our novel approaches based on repeated k-mers could be particularly useful for characterizing other genomes and populations combining admixture and polyploidy.

## SESSION 2 - EXPLORATION DE LA DIVERSITE : GENE

### Génomique de l'immunité chez le blé par l'analyse de données transcriptomiques Pacbio Kinnex

**Frédéric Choulet<sup>1</sup>**

<sup>1</sup>GDEC, INRAE, Université Clermont Auvergne, Clermont-Ferrand, 63000, France.

La génomique chez le blé a toujours été très impactée par les avancées technologiques dans le domaine du séquençage. C'est particulièrement prégnant chez cette espèce car son génome est très grand, 15 Gb, et hexaploïde. Cette complexité impose des limites fortes sur notre capacité à caractériser son génome avec précision, notamment les gènes de résistance souvent très répétés. La taille des lectures et les débits des séquenceurs sont donc des facteurs déterminants. Nous avons produit des données transcriptomiques Pacbio Kinnex (plateforme Gentyane) dans l'objectif d'identifier trois gènes de résistance (*Stb*) à *Zymoseptoria tritici*,

un pathogène fongique. 57 millions de lectures ont été obtenues sur 15 échantillons. Les aspects méthodologiques et les résultats des analyses bioinformatiques seront présentés. Les pistes ouvertes par cette technologie dans la génomique de l'immunité des plantes seront discutées.

### SESSION 3 – METAGENOMIQUE

#### **Projet LAMINET : L'ADNe pour décrypter la structure trophique des forêts de Laminaires**

**Baillet B., Quéméré E., Leclerc J-C.**

Les forêts de laminaires sont un des écosystèmes sous-marins les plus emblématiques du littoral, tant pour leur valeur écologique, économique et culturelle. Elles constituent un habitat privilégié pour un cortège d'espèces d'algues, de méiofaune et de macrofaune, qui sont des ressources importantes pour les pêcheries.

Au niveau mondial, ces écosystèmes sont fragilisés par des épisodes de dénudation à large échelle du fait d'une conjonction de facteurs tels que la pollution, la surpêche, l'exploitation goémonière, des invasions biologiques ou des anomalies de température liées au changement climatique. Les disparitions massives de laminaires s'accompagnent souvent de dérèglements dans leur dynamique trophique pouvant avoir des effets en cascade sur toute la communauté. Pour comprendre les liens entre biodiversité et fonctionnement des forêts de laminaires et en particulier les processus de « compensation biotique » qui favorisent leur résilience et stabilité à long terme, il est essentiel de connaître les réseaux d'interactions qui relient les espèces au sein des communautés. Les approches utilisées traditionnellement pour déterminer les préférences alimentaires (isotopie, micro-histologie des contenus stomacaux) sont difficiles à développer sur un nombre important d'individus/taxons et restent peu précises. Ainsi, on manque de données sur les espèces qui dépendent directement ou indirectement de la fraction détritique des laminaires ou encore sur les consommateurs secondaires (poissons, macro invertébrés) qui interviennent dans la régulation des populations d'herbivores. Plus généralement, on manque de descripteurs de la qualité fonctionnelle des habitats pour évaluer l'impact des pressions anthropiques.

Co-financé par l'OFB et la Région Bretagne, le projet LAMINET vise à Identifier les interactions trophiques entre les organismes associés aux forêts de laminaire (phoque, oiseaux marins, poissons, invertébrés et algues) dans deux aires marine protégées (Parc National Marin d'Iroise et Réserve Naturelle Nationale des Sept Iles) en développant une approche génétique basées sur l'ADN environnemental (ADNe) dans les fèces et contenus stomacaux (métabarcoding ADN trophique). Dans un deuxième temps, le projet vise à identifier les liens entre la structure trophique (régime alimentaire d'espèce bio-indicatrices), la biodiversité (benthique et ichtyologique) et la production secondaire (tailles et abondances des poissons) dans des herbiers de laminaires en plus ou moins « bonne santé ».

#### **Thibault Leroy : Le séquençage des miels, un regard génomique sur l'agroécologie?**

Au cours de cette présentation, je présenterai les résultats du projet MALLAURIE (2023-2024), qui visait à étudier l'intérêt d'un séquençage shotgun directement basé sur les produits de la ruche (miel, cire, propolis, etc.) pour la caractérisation du continuum agroécologique plantes-abeilles mellifères, en y incluant les ADN des végétaux visités et des abeilles elles-mêmes, ainsi que pour sur les perspectives offertes par cette approche pour le suivi sanitaire des colonies et pour l'amélioration de la traçabilité des produits de la ruche.

## Développement d'une puce SNP multi-espèces pour caractériser la diversité génétique chez les protéagineux (pois, féverole, lupin blanc)

Grégoire Aubert<sup>1\*</sup>, Jonathan Kreplak<sup>1</sup>, Mathieu Chanis<sup>1</sup>, Margaux Briffaut<sup>2</sup>, Carole Confolent<sup>2</sup>, Lucy Garaud<sup>2</sup>, Lydia Jaffrello<sup>2</sup>, Anthony Klein<sup>1</sup>, Jean-Bernard Magnin-Robert<sup>1</sup>, Marianne Chabert-Martinello<sup>1</sup>, Karen Boucherot<sup>1</sup>, Céline Rond<sup>1</sup>, Florence Naudé<sup>1</sup>, Damien Ollivier<sup>1</sup>, Hélène Pidon<sup>3</sup>, Charles Poncet<sup>2</sup>, Nadim Tayeh<sup>1</sup>, and Judith Burstin<sup>1</sup>

<sup>1</sup> UMR Agroécologie, INRAE, Université Bourgogne Europe, Institut Agro Dijon, Dijon, France

<sup>2</sup> UMR GDEC Gentyane, INRAE, Université Clermont Auvergne, Clermont-Ferrand, France

<sup>3</sup> UMR AGAP Institut, INRAE, Université de Montpellier, CIRAD, Institut Agro, Montpellier, France

Le pois (*Pisum sativum* L.), la féverole (*Vicia faba* L.) et le lupin blanc (*Lupinus albus* L.) sont des cultures de grand intérêt pour l'alimentation humaine et animale, grâce à la forte teneur en protéines de leurs graines. De plus, leur capacité à établir des symbioses pour fixer l'azote atmosphérique en font des espèces de choix pour les systèmes de culture agroécologiques. Elles sont cependant confrontées à de nombreux stress biotiques et abiotiques, amplifiés par le changement climatique. Leur amélioration variétale est cruciale et l'exploitation de la diversité génétique présente dans ces espèces en est un levier majeur. La conservation et la caractérisation des ressources génétiques au niveau génotypique sont donc des enjeux prioritaires.

Le CRB PROTEA, géré par l'UMR Agroécologie à Dijon, maintient ainsi des collections de ressources génétiques comportant respectivement 3378, 1580 et 1209 accessions patrimoniales actives de pois, féverole et lupin blanc. Les outils de génotypage disponibles jusqu'ici étaient trop onéreux pour caractériser l'entièreté des collections, et les techniques de génotypage par séquençage sont peu adaptées à ces espèces en raison de la grande taille de leur génome.

Dans le cadre du projet PEPR Agroécologie et numérique AgroDIV, une puce de génotypage multi-espèces de technologie Axiom<sup>®</sup>, contenant respectivement 17941, 17595, et 4846 marqueurs SNP pour le pois, la féverole et le lupin blanc, a été développée à partir des ressources génomiques disponibles. En parallèle, une méthode d'extraction d'ADN non destructrice à partir de graines individuelles a été mise au point. Une ou plusieurs graines ont été génotypées par accession suivant le statut des accessions (lignée ou population).

Les premiers résultats de génotypage montrent que la grande majorité des marqueurs est de haute qualité avec très peu de données manquantes. Les données de génotypage obtenues vont permettre d'éliminer les doublons et de structurer la diversité au sein des collections. Ce travail a ainsi d'ores et déjà permis de définir une core-collection chez le lupin blanc, actuellement en phénotypage au champ dans le cadre du projet PULSAR.

**Mots clés :** puce à SNP, légumineuses, marqueurs moléculaires, ressources génétiques

### Adhikari Kritika : Assessing the diversity of the Espelette pepper landrace in the French Basque country

Espelette pepper (*Capsicum annuum* L.) is a traditional landrace cultivated in the Basque region of France and represents a unique genetic resource shaped by local practices and environments. Currently, this landrace faces increasing vulnerability to biotic stresses, such as the cucumber mosaic virus (CMV), which threaten its yield stability and long-term viability. This study aims to investigate how diversity within the Espelette population influences its adaptive response to CMV and other biotic stresses. Utilizing existing phenotypic and genomic data from 575 Espelette plants sampled in 23 farms distributed on the Espelette cultivated PDO (Protected Designation of Origin) area (~299 Ha), we aim to explore population structure, diversity patterns, and potential genotype–phenotype associations related to biotic stress tolerance. By considering diversity not only as a target of conservation but as a functional mechanism for adaptation, this research could contribute to a broader understanding of how intraspecific variation influences pathogen adaptation and promotes crop resilience. As a part of the COMBINE project, the research insights will support the development of sustainable management practices, such as varietal mixtures or resistance combinations, that minimize epidemics and

enhance the durability of resistance through diversity. Moreover, the findings may encourage integrative strategies that valorise traditional varieties as important contributors to breeding programs.

### **SABRINA DELAUNAY : Comment les marqueurs KASPAR permettent de suivre l'introgression de la fétuque glaucescence dans le RGI**

Dans une optique d'amélioration des graminées de prairies semées *via* la production d'hybrides *Festulolium*, des données issues du génotypage par séquençage de quatre espèces de graminées fourragères allogames constituées de mélanges (bulks) de *L. multiflorum*, *F. pratensis*, *F. arundinacea* et *F. arundinacea var. glaucescens* ont été étudiées.

Une fois les génomes séquencés et alignés sur une séquence consensus de *L. perenne*, les données de mapping, de distance génétique et de fréquence d'hétérozygotie ont conduit à l'étude de leur niveau de polymorphisme intraspécifique.

Dans un second temps, ces données ont servi à extraire 295 SNP interspécifiques localisés sur les sept groupes de liaison des graminées étudiées et permettant de discriminer chez un hybride, l'origine du matériel génétique issu de l'une ou d'autre des espèces parentes. Pour cela plusieurs tris, notamment sur le nombre de reads, la fréquence des bases majoritaires à l'échelle de l'espèce, la distance des positions d'intérêt entre elles et le positionnement sur la carte de Byrne *et al.*, 2015, ont été réalisés.

Ces SNP ont par la suite été utilisés pour designer des marqueurs KASPar pour l'étude d'hybrides *Festulolium* et ont été partagés à une entreprise de sélection privés qui s'en sert comme outils de sélection assistée par marqueurs.

**Mots-clés :** marqueurs moléculaires, SNP, Kaspar, *Festulolium*, plantes fourragères, sélection variétale.

## **SESSION 5 - REGULATION DE L'EXPRESSION DES GENOMES**

### **Hervé Acloque : Single-cell omics reveal distinct gene regulatory dynamics underpinning embryonic and extraembryonic lineage functions during pig blastocyst development**

The late-stage development of the blastocyst before implantation is a unique feature of ungulates. During this period, the epiblast proliferates and remains pluripotent for several days before gastrulation begins. Simultaneously, extra-embryonic tissues undergo significant growth, elongating to several tens of centimeters. However, the mechanisms that regulate and synchronize these processes remain poorly understood. In this study, we analyzed transcriptomic and epigenomic data at the single-cell level from early and late porcine blastocyst cells, spanning from the hatched blastocyst stage (E7) to early (E9) and late ovoid blastocyst (E11). We characterized 15,370 blastocyst cells, clustering them into distinct embryonic and extra-embryonic populations with specific chromatin accessibility landscapes. For each population, we inferred gene regulatory networks based on enhancer-driven gene regulation modules (eRegulons) and validated them through motif occupancy visualization. Our findings reveal strong dynamics in gene regulatory module activity within extra-embryonic tissues at the onset of blastocyst elongation, while gene regulatory activity in the epiblast remains relatively stable.

### **Julie Demars**

### **Discovering parent-of-origin expression and methylation in the pig highlights evolutionary and molecular mechanisms shaping genomic imprinting**

The development of custom methylation arrays and panels specifically targeting genomic imprinted regions promote a routine access to imprintomes, which can be seen as the fraction of the methylome regulating genomic imprinting. A genome-wide focus on this key regulatory layer should provide information for better understanding a range of mechanisms at play in imprinting at different scales, from macroevolutionary processes to genomic region-specific possibilities of gene dosage control. Here we have undertaken a detailed exploration of parental Differentially Methylated Regions (DMRs) in the pig relying on the bisulfite-free enzymatic methylome sequencing of reciprocally crossed individuals. Such a profiling assay, applicable to other species, allowed the identification of candidate parental DMRs. This firstly contributes to a better understanding of imprinting regulations in the pig, especially in large and complex imprinted domains. From

an evolutionary point of view, accessing a conserved fraction of the parental DMRs illustrates possible trajectories of genome diversification following retrovirus domestication. Other loci of evolutionary interest but still incompletely characterized, at least in the pig, show less heterogeneity across species. Interestingly, we also identify likely pig-specific DMRs. The set of original results highlighted here suggests that an in-depth analysis of parent-of-origin expression and methylation can help clarify a range of regulations at play in genomic imprinting.

#### POSTERS :

### Identification and Evolution of Metabolite Biosynthetic Gene Clusters in the Cannabaceae Family

O Lrhcha<sup>1</sup>, M Da Rocha<sup>1</sup>, P Favier-Rampant<sup>2</sup>, D Hinsinger<sup>2</sup>, A Gravot<sup>3</sup>, A Bouchereau<sup>3</sup>, Leo Andruszkow<sup>3</sup>, F Mathis<sup>4</sup> and F Hilliou<sup>1</sup>

1: INRAE, Institut Sophia Agrobiotech Université Côte D'Azur Nice France

2: Université Paris-Saclay, Centre INRAE Île-de-France Versailles-Saclay, EPGV, Evry, F-91057, France.

3: IGEPP-P2M2, INRAE, Institut Agro, Université de Rennes, 35000 Rennes, France.

4: HEMP-it ADN, 49250, Beaufort-en-Anjou, France.

The reduction in the use of chemical pesticides due to the banning of certain molecules such as neonicotinoids and organophosphates (e.g. Phosmet), as well as the development of pesticide resistance mechanisms in insects, has led to the need for new control methods and plant varieties.

In recent years, flea beetles have attacked industrial hemp (fiber) crops, particularly those cultivated for seed production. These attacks occur in the spring, at the beginning of the growing season, on young hemp shoots. The flea beetle is also a pest of hop, another plant in the Cannabaceae family. Different levels of tolerance to this insect have been observed in plots with different genetic backgrounds (Hemp-it-DNA). The PlantAlliance ACRA project aims to leverage hemp's genetic diversity to identify natural variabilities that could be selected to thwart phytophagous pests.

We are working to understand which volatile organic compounds (VOC) make hemp more repulsive to the flea beetle. Genes that encode biosynthetic enzymes for specialized metabolic pathways are not randomly distributed throughout the genome, instead they are physically linked in structures called biosynthetic gene clusters (BGC). We compare the BGCs organization in hemp, hop and medicinal cannabis genomes. Using plantiSMASH, we detect BGCs in each genome and analyze the microsynteny among the three.

### **Génomique de l'adaptation : stratégie d'étude de la tolérance à la chaleur et à la sécheresse chez les vanilliers** **Rinah H. Ravelonanosy**

Le changement climatique menace la pérennité des plantations de vanille Bourbon (*Vanilla planifolia*) dans le monde, en particulier à Madagascar et à La Réunion. Les populations de vanilliers sauvages aphylls de Madagascar, qui se développent dans des conditions arides, pourraient représenter une ressource génétique précieuse pour renforcer la tolérance à la sécheresse et à la chaleur des vanilliers cultivés. Dans cette étude, des technologies de séquençage nouvelle génération (NGS) sont utilisées pour identifier des marqueurs génomiques de l'adaptation à la sécheresse et à la chaleur chez les vanilliers. Tout d'abord, les signatures génomiques de l'adaptation climatique sont recherchées dans des populations naturelles de ces espèces sauvages aphylls de Madagascar, présentes dans des environnements contrastés, par des approches de scans génomiques et d'associations génotype-environnement (GEA). Deux espèces à large distribution ont été échantillonnées à Madagascar pour cette étude : *V. bosseri* et *V. perrieri*, suite à une analyse préliminaire révélant l'existence d'IBE (isolation-by-environment), à savoir l'influence des facteurs bioclimatiques sur la

structuration génétique de ces espèces. Le génome d'une espèce aphyllé (*V. bosseri*, CR0821) est aussi séquencé et assemblé par des approches PacBio HiFi et carte optique. L'endoréplication partielle (PE) du génome caractéristique des vanilliers pouvant influencer la qualité du séquençage et de l'assemblage, l'ADN est extrait à partir de boutons floraux car ils présentent une forte proportion de noyaux non-endorépliqués contrairement aux noeuds et tiges. L'annotation des gènes de ce génome permettra d'identifier les gènes candidats correspondants aux marqueurs SNP identifiés. Enfin, des études de génomique comparative avec le génome publié de *V. planifolia* permettront d'identifier les gènes correspondants chez l'espèce cultivée. Ces résultats seront essentiels à plus long terme pour identifier des variétés de vanillier possédant ces gènes et donc adaptées aux conditions de chaleur et de sécheresse.

#### **Nathalie Iannuccelli :**

La méthylation du génome représente une source importante de régulation de l'expression des gènes. À ce jour, les outils moléculaires permettant d'étudier des régions ciblées du génome sont limités à des régions de 10 Mb. Nous avons donc mis au point un protocole haut-débit d'épi-génotypage des CpGs différenciellement méthylés dans des régions spécifiques du génome. Ce développement a été réalisé dans le cadre du programme de recherche PIPETTE (imPrIntome to PigLET birTh wEight, ANR, J. Demars) dont l'objectif est de déterminer les gènes soumis à empreinte parentale impliqués dans la variabilité du poids de naissance chez le porc.

Ce développement effectué en collaboration avec avec la Plateforme Génomique et Transcriptomique (GeT PlaGe) s'est déroulé en 3 étapes :

- 1) La comparaison de technologies de capture de méthylation - de la référence Agilent® à Twist Bioscience® avec optimisation d'un panel dédié.
- 2) Le changement d'échelles - du développement (n=8) au haut-débit (n=96 puis 384) technologie Twist Bioscience, avec réduction du temps de manipulation et des coûts.
- 3) Le changement de séquenceurs - du NovaSeq à l'Aviti technologie Twist Bioscience, avec amélioration des performances.

Le protocole développé avec la technologie Twist Bioscience fonctionne avec de faibles quantités d'ADN extraites à partir de sang ou de semence, sur un panel dédié complexe. Le protocole est adaptable à n'importe quelle région, tissu ou espèce.

Cette approche pourrait être utilisée pour des études de méthylation de tout organisme disposant d'un génome de référence et également chez l'homme en tant qu'outil de diagnostic à haute densité.