WORKSHOP

PRE-CONGRES

Du 7 au 9 avril 2025 au CIGASS

Analyse des données génétiques, épi-génomiques et transcriptomiques massives : Introduction aux principes théoriques et ateliers pratiques

------Programme Final------

Lundi 7 Avril 2025

8h30	**Accueil des participants**
9h00	**Introduction par les organisateurs**
9h30-13h00	**Session 1 : Le séquençage d'exome**

Objectif pédagogique de la session

Pendant cette session, nous allons aborder l'approche génétique de l'étude du génome codant ou exome. Une première partie sera consacrée à quelques définitions et rappel théoriques sur cette approche très populaire en génétique des maladies dites monogénique ou oligo-génique, souvent rares à très rares. Les participants seront initiés aux critères et choix à appliquer pour tenter d'identifier efficacement les variants génétiques les plus prometteurs pour expliquer une maladie génétique.

1.1	Enjeux et objectifs du séquençage d'exome (Définitions)
1.2	Contrôle qualité des données exomes (Théorie)
1.3	Annotation des variants et identification des variants d'intérêt (Théorie)
1.4	Critères de définitions : variants génétiques délétères

11h00-11h30 **Pause-café-thé**

1.5 Exercice : Filtrations des variants issus d'exomes

1.6. Exercice : stratégie d'analyse de ségrégation

13h00-14h00 **Pause déjeuner**

14h00-17h00 **Session 2 : Théorie des études d'association génomique (GWAS)**

Objectif pédagogique de la session

Dans cette session, l'objectif sera de fournir une compréhension solide des études d'association pangénomiques (Genome-wide association studies : GWAS) et une initiation à la visualisation de données à l'aide de R et ggplot2. Nous aborderons quelques définitions de cette approche qui a révolutionné l'étude génétique des maladies communes et multifactorielles qui représente la majorité écrasante des maladies humaines.

2.1 Introduction à l'approche GWAS (Définition, contexte et exemples)

2.2 Théorie de l'association génomique (Conception épidemio-génétique)

2.3 Concepts clés (déséquilibre de liaison, données génétiques, imputation)

15h30-16h00 **Pause-café-thé**

2.4 Contrôle qualité des données et tests statistiques

2.5 Limitations et défis des GWAS

17h - Fin de la journée 1– Diner libre

Mardi 8 Avril 2025

9h00- 12h30 Session 3 : Accès et visualisation des données génétiques

Objectif pédagogique de la session

Les GWAS restent une approche couteuse, qui nécessitent d'avoir accès à des cohortes de taille relativement grande, et dans un design cas témoins. Dans cette session, nous décrirons le contenu et son utilisation de plusieurs sites où se trouvent des ressources importantes de données GWAS libres d'accès le plus souvent. Dans la partie pratique, les participants seront initiés à la visualisation des résultats spécifiques aux données GWAS et comment les interpréter.

3.1 Les bases GWAS et LocusZoom

3.2 Importation et interprétation de résultats résumés (summary statistics)

10h30-11h00 Pause-café-thé

3.3 Création des graphiques basiques avec ggplot2

3.4 Création d'un Manhattan plot avec ggplot2

12h30-13h30 Pause déjeuner

13h30-16h30 Session 4 : **Rôle des variants génétiques non-codants identifiés par GWAS**

Objectif pédagogique de la session

Dans cette session, nous aborderons une notion importante d'analyse post-GWAS. Les variants non-codants étant majoritaires parmi ceux identifiés par GWAS, il existe de nombreuses approches analytiques comprenant la position génomique, l'annotation épi-génomique et l'association avec les niveau d'expression des gènes à proximité. Ici nous expliquerons comment aborder l'ensemble d'outils web afin de définir *in silico* le caractère fonctionnel des hits GWAS.

4-1. Notions de eQTL: corrélation génotype expression

4-2. Épigénétique et régulation transcriptionnelle

4-3. Lien entre épigénétique et génétique

4-4. Les études épi-génomiques: RNA-Seq et ATAC-Seq

15h00-15h30 **Pause-café-thé et Photo de groupe**

4-5. Choix des données épi-génomiques pertinent (ENCODE)

4-6. Visualisation des données sous R

4-7. Utilisation *UCSC Genome Browser* pour visualiser les données

17h - Fin de la journée 2– Diner libre

Mercredi 9 Avril 2025

9h00- 12h30 **Session 5: Applications épidémio-génétiques des GWAS**

Objectif pédagogique de la session

Dans cette dernière session, il sera question d'aborder quelques-unes des nombreuses applications épidémio-génétiques des données obtenues par GWAS. Cette session fait un focus particulier sur les score polygéniques (PRS pour polygenic risk scores), dont la promesse est importante dans la mise en place d'une future médecine génomique et de précision.

5-1. Introduction aux applications épidémiologiques des données GWAS

5-2. Présentation des scores génétiques

5-3. Méthodes de calcul et applications

10h30-11h00 **Pause-café-thé**

5-4. Calcul d'un PRS pour prédire le risque de maladie à partir de GWAS

5-5. Limitations et défis des PRS

12h30-13h00 **Session de Clôture**

Questions des Participants et Sondage

13h - Fin du Workshop- Déjeuner Libre

Intervenants

Nabila Bouatia-Naji, docteure en génétique humaine (Université de Lille), est directrice de recherche (DR1) et chef d'équipe au PARCC-Inserm. Spécialiste de la génétique des maladies cardiovasculaires, elle étudie notamment la SCAD (Spontaneous Coronary Artery Dissection), entre autres maladies prédominantes chez les femmes jeunes. Ses travaux, menés avec une équipe multidisciplinaire, ont permis d'identifier plusieurs causes génétiques. Reconnue mondialement, elle a reçu des distinctions prestigieuses, dont l'ERC et le prix Jean-Paul Binet. Engagée pour l'égalité des sexes en science, elle siège au conseil de *l'European Society of Human Genetics* (ESHG) et de la *Science For Africa Foundation* (SFA).

Asraa Esmael, doctorante à l'Université Patis Cité, est titulaire d'un Master en bio-informatique de l'Université de Rennes. Elle développe actuellement un projet de thèse visant à mettre en place des stratégies d'intégration de données multi-omiques dans le contexte de maladies génétiques complexes comme la SCAD, grâce notamment à des méthodes *Machine learning*.

Takiy-Eddine Berrandou, Docteur en génétique épidémiologique (Université Paris-Saclay), il a étudié les interactions gènes-environnement dans le cancer du sein. Il a ensuite rejoint l'Inserm (PARCC) pour étudier la base génétique de maladies cardiovasculaires prédominantes chez les femmes. Lors de son postdoctorat à l'Université d'Aarhus (Danemark), il a travaillé sur le développement de méthodes en génétique statistique. Il développe et applique actuellement au PARCC-Inserm des méthodologies d'analyse génétique à l'étude de maladies cardiovasculaires complexe et peu étudiées.

Adrien Georges, docteur en biologie cellulaire (Université Paris-Diderot), a étudié la régulation transcriptionnelle dans les maladies rares et la levure. Il a rejoint le PARCC-Inserm pour explorer la régulation génomique et transcriptomique dans le risque polygénique des maladies cardiovasculaires. Actuellement chargé de recherche (CRCN) à l'Inserm, il mène des travaux sur la régulation liée au risque génétique de la SCAD, entre autres maladies cardiovasculaires.