

## **Stage Master 2 en bio-informatique :**

*Comprendre le rôle des R-loops générées aux cassures doubles brins de l'ADN à l'aide de données de conformation tridimensionnelle du génome (Hi-C)*

### **Environnement de travail, contexte**

Le Centre de Biologie Intégrative (CBI) de Toulouse développe des approches multidisciplinaires, multi-échelles dans le but de comprendre le fonctionnement des organismes vivants. Les équipes du CBI cherchent notamment à décrypter les mécanismes et fonctions impliqués dans la maintenance, l'expression et l'évolution du génome.

Parmi les équipes du CBI, l'équipe "Chromatine et réparation de l'ADN", dirigée par Gaëlle Legube, cherche à comprendre la contribution de l'architecture des chromosomes dans la maintenance de l'intégrité du génome. Pour cela l'équipe a développé une lignée cellulaire humaine appelée DiVA qui permet d'induire des cassures double-brin de l'ADN de façon séquence spécifique à de multiples loci sur le génome.

L'équipe est soutenue pour les analyses de séquençage haut débit par la plateforme bio-informatique d'analyse des données génomique (bigA). La plateforme déploie les chaînes de traitement pour traiter les données de séquençage et accompagne les chercheurs dans les analyses statistiques des données.

Dans le cadre du metaprogramme DIGIT-BIO chrocoNET (<https://digitbio.hub.inrae.fr/rubriques-verticales2/nos-actions/consortia/consortium-chroconet-2023-2025>), l'équipe sera accompagnée par Anne-Laure Valton et Sylvain Foissac chercheurs dans le laboratoire GenPhyse à l'INRAE Auzeville qui apporteront leur expertise dans le traitement et l'analyse des données issues des techniques de Chromosome Conformation Capture à haut débit (Hi-C) permettant d'étudier la conformation 3D de la chromatine dans les cellules et de mieux comprendre les liens entre la structure 3D du génome et son fonctionnement.

### **Mission et activités**

A l'aide du système cellulaire DiVA, nous réalisons des expériences de Hi-C pour étudier la conformation des chromosomes en réponse aux cassures double-brin. Par exemple, l'équipe a montré que suite à l'induction de cassures, un processus d'extrusion de boucle de la chromatine, médié par le complexe des cohésines, se produit au niveau des cassures afin de déposer une modification de la chromatine (la phosphorylation du variant d'histone H2AX) sur de larges domaines chromatiniens autour de la cassure (Arnould et al, Nature 2021). Plus récemment nous avons montré que suite aux cassures, un nouveau compartiment de chromatine se forme dans le noyau au sein duquel les cassures elles-mêmes, et les gènes activés suite aux dommages à l'ADN, se regroupent (Arnould, Rocher et al, Nature, 2023).

Un projet en cours vise à comprendre le rôle des R-loops (structures hybrides ARN:ADN), générées aux cassures doubles brins, dans les changements d'architecture du génome en réponse aux cassures. De façon intéressante nous avons accumulé des évidences indiquant qu'une modification sur la composante ARN de ces hybrides ARN:ADN (la méthylation m6A) contribue à ce processus. Nous avons généré des données de Hi-C en présence et en absence de cassures double brins dans des cellules déficientes pour un facteur clé de la régulation des R-loop (SETX) et un facteur clé du métabolisme du m6A (FTO). Le projet consiste à analyser ces données de Hi-C afin de comprendre la contribution des R-loops et de leur modification en m6A, dans l'architecture des chromosomes, en présence et en absence des cassures double-brin. Nous étudierons l'impact de ces déficiences sur l'extrusion de boucle, sur l'intégrité des domaines chromatiniens (TADs) et sur les compartiments nucléaires, en relation avec le transcriptome et l'épigénome générés dans les mêmes cellules avant et après induction des cassures (RNA-seq, ChIP-seq, ATAC-seq).

Le projet comportera donc une partie de traitement des données brutes issues des séquençages Hi-C à l'aide de chaînes de traitement et d'un serveur de calcul accompagné par la plateforme bigA, une partie analyse et visualisation des données en R et/ou Python en collaboration avec les experts du réseau chrocoNET. Le stage pourra se poursuivre par une partie d'intégration avec les autres données de séquençage disponibles dans l'équipe et d'interprétation des résultats avec les chercheurs de l'équipe.

### **Formation et compétences recherchées**

\* **Diplôme requis** : Master/Ingénieur (Bac + 5)

\* **Formation** : informatique, statistique ou bio-informatique

\* **Connaissances/compétences** :

- Bases de programmation spécifique à l'analyse de données (python/R)
- Bases en génomique
- Utilisation de Linux et serveur de calcul
- Maîtrise de l'Anglais écrit
- Communication, travail en équipe, travail avec différent interlocuteurs.

Envoyer le CV a [marion.aguirrebengoa@univ-tlse3.fr](mailto:marion.aguirrebengoa@univ-tlse3.fr) et [anne-laure.valton@inrae.fr](mailto:anne-laure.valton@inrae.fr)

**Contrat** : Stage

**Durée** : 6 mois

**Début du contrat** : début 2024

**Rémunération** : gratification

**Références** :

- Arnould C, Rocher V, Finoux AL, Clouaire T, Li K, Zhou F, Caron P, Mangeot PE, Ricci EP, Mourad R, Haber JE, Noordermeer D & Legube G (2021) Loop extrusion as a mechanism for formation of DNA damage repair foci, Nature
- Arnould C, Rocher V, Saur F, Bader AS, Muzzopappa F, Collins S, Lesage E, Le Bozec B, Puget N, Clouaire T, Mangeat T, Mourad R, Ahituv N, Noordermeer D, Erdel F, Bushell M, Marnef A & Legube G (2023) Chromatin compartmentalization regulates the response to DNA damage, Nature