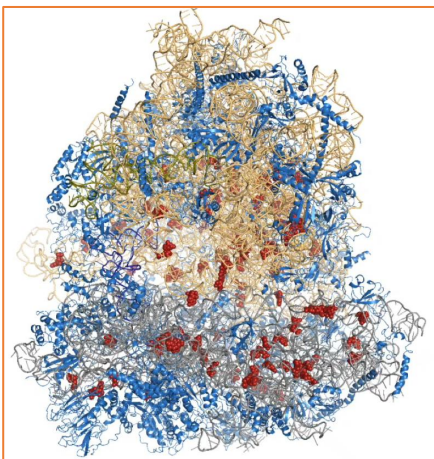


**Titre | Développement d'outils bio-informatiques dédiés à l'analyse de l'épitranscriptomique des ARN ribosomiques****Structure |** Centre de Recherche en Cancérologie de Lyon / Centre Léon Bérard, 69008 Lyon**Contact |** Virginie MARCEL  
Directrice de Recherche Inserm, CRCL  
[virginie.marcel@lyon.unicancer.fr](mailto:virginie.marcel@lyon.unicancer.fr)**Collaboration |**  
Emilie THOMAS et Anthony FERRARI (Plateforme Bio-informatique Gilles Thomas, CRCL)  
Fleur BOURDELAIS et Sébastien DURAND (Plateforme RibosOMICS, CRCL)**Public cible |** Projet de stage (4-6 mois) de fin d'étude pour un ingénieur ou M2 en bioinformatique.**Contexte général du projet**

Les récentes découvertes impliquant le ribosome ouvrent de nouvelles perspectives dans la compréhension de l'expression génique via le champ émergent de l'épitranscriptomique. Chez l'Homme, il apparaît que le ribosome peut présenter des compositions différentes, chaque variant de ribosomes favorisant la traduction de certains ARNm. Ces différences de composition impliquent les modifications chimiques des ARN ribosomiques (ARNr), comme la 2'O-ribose méthylation (2'Ome) ou la pseudouridylation ( $\Psi$ ) des ARNr.

L'étude de la 2'Ome des ARNr est possible grâce au développement récent d'une approche innovante basée sur le RNA-seq, le **RiboMethSeq** (Birkedal et al., 2015). Nous avons récemment démontré que la 2'Ome des ARNr est altérée dans les cancers (Marcel et al., 2021 ; Paraqindes et al., 2023). Pour ces études, nous avons développé un pipeline complet d'analyse des données de séquençages dédié à l'analyse données de RiboMethSeq

([github.com/RibosomeCRCL](https://github.com/RibosomeCRCL)). Théoriquement, ces outils sont aussi utilisables pour l'approche d'**HydraPsySeq** pour analyser la  $\Psi$  des ARNr (Marchand et al., NAR 2020). **Il reste aujourd'hui à adapter les outils bioinformatiques du RiboMethSeq à l'HydraPsySeq.**

Les résultats issus de ce projet permettront d'élargir le champ de nos outils d'analyse afin d'explorer de nouveaux phénomènes biologiques qui peuvent être impliqués dans la formation et le développement des cancers. Ces nouvelles technologies visent à identifier de potentiels nouveaux biomarqueurs pronostiques et prédictifs pour les patients atteints de cancers sur la base des modifications chimiques des ARNr.

**Objectifs du projet**

Les objectifs de ce stage seront donc **d'optimiser les outils de RiboMethSeq à l'HydraPsySeq**:

- (1) **ribomethseq-nf** : pipeline bioinformatique (nextflow) permettant de passer des données de séquençage (fastQ) aux matrices de comptage bruts base par base.
- (2) **rRMSAnalyzer** : package R permettant de déterminer le niveau de  $\Psi$  à partir des données de comptage produites par ribomethseq-nf ainsi que des analyses multivariées et différentielles.

**Jeux de données préexistants**

Afin de mener à bien ces missions, l'étudiant stagiaire disposera de jeux de données générés localement par la

plateforme de séquençage du CRCL ainsi que des données issues de la littérature (GSE202531).

### Méthodologie envisagée

Il s'agit de problèmes d'optimisation. On pourra considérer différentes approches :

- (1) rééchantillonnage des données pour déterminer une profondeur de séquençage optimale permettant une stabilité des résultats
- (2) test et comparaison de différents paramétrages (concernant notamment le calcul du taux de  $\Psi$ ) afin de maximiser les différences biologiques

### Analyses attendues

- (1) Tester les outils existants sur les données d'HydraPsiSeq
- (2) Ajuster les paramètres et méthodes
- (3) Déterminer les conditions de séquençage idéales
- (4) Intégrer les éventuels ajustements de méthodes dans les packages existants

### Environnement

Le stage sera réalisé dans le cadre d'une collaboration entre la plateforme nouvellement créée RibosOMICS, qui fournit une expertise complète des approches haut débit dédiées à l'analyse de la régulation traductionnelle composée d'une ingénieure biologiste et de deux chercheurs du CRCL, et la plateforme bioinformatique Gilles Thomas. L'étudiant stagiaire bénéficiera donc d'un co-encadrement par des membres des deux équipes. En terme d'environnement informatique, il aura accès au cluster HPC du centre.

*Birkedal, U., et al. 2015. Profiling of ribose methylations in RNA by high-throughput sequencing. Angew. Chem. Int. Ed. Engl. 54, 451–455.*

*Marcel, V., et al. 2020. Ribosomal RNA 2'O-methylation as a novel layer of inter-tumour heterogeneity in breast cancer. NAR Cancer., vol2.*

*Marchand, V., et al. 2020. HydraPsiSeq: a method for systematic and quantitative mapping of pseudouridines in RNA. NAR 48:e110.*

*Paraqindes, H, et al. 2023. IDHwt and IDHmut adult-type diffuse gliomas display distinct alterations in ribosome biogenesis and 2'O-methylation of ribosomal RNA. Neuro Oncol, noad140.*