

**Analyse transcriptomique 3' RNAseq**

**Lieu d'affectation** : Site TRiX, plateforme GeT, Centre INRAE Occitanie – Toulouse

**Adresse** : Toxalim UMR1331 INRAE/INP/UPS - 180 chemin de Tournefeuille - 31027 Toulouse

**Durée**: 6 mois à compter de janvier-mars 2023

**Contact** : Yannick Lippi

**Candidature** : Envoi de CV et lettre de candidature par courriel à [yannick.lippi@inrae.fr](mailto:yannick.lippi@inrae.fr) jusqu'au 15 novembre 2022

**Gratification** : environ 573€/mois selon le simulateur public :

<https://www.service-public.fr/simulateur/calcul/gratification-stagiaire>

## Contexte

---

Vous travaillerez au sein du site TRiX de la plateforme Génome et Transcriptome (GeT, <http://get.genotoul.fr>, Génopôle Toulouse Midi-Pyrénées) qui met à disposition des laboratoires publics ou privés une expertise en matière d'analyse de l'expression des gènes (étude du transcriptome) à l'aide de technologies sans a priori à haut-débit (microarrays, séquençage) et de PCR quantitative en temps réel. Vous serez affecté dans les locaux du laboratoire Toxalim (150 agents dont une centaine de permanents, [www6.toulouse.inrae.fr/toxalim](http://www6.toulouse.inrae.fr/toxalim)) qui héberge l'équipe TRiX composée d'un ingénieur analyste de données, d'une technicienne biologiste et une Ingénieure biologiste NGS. Les domaines d'expertises apportés par le plateau sont ceux de la biologie moléculaire et de la biostatistique. L'équipe dispose des compétences pour l'évaluation de la qualité des échantillons d'ARN, pour l'acquisition des signatures transcriptomiques par microarrays et RNAseq ainsi que pour l'analyse bioinformatique et biostatistique des données.

## Missions et activités confiées

---

L'accès aux technologies traditionnelles d'études transcriptomiques à haut débit reste encore difficilement abordable pour des dispositifs expérimentaux à grande échelle ou d'élevages du fait des coûts importants à engager, et de protocoles qui ne sont pas adaptés aux contraintes de terrain et des procédures d'analyses de données complexes à appréhender pour des équipes non expertes dans le domaine.

Dans le cadre d'un projet financé par l'institut Carnot France Futur Elevage, nous proposons de réaliser des tests de développement de méthodes afin de mettre en place au sein de la plateforme de transcriptomique GeT-TRiX un service intégré d'analyse transcriptomique à moindre coût à destination des équipes de recherche ou de R&D confrontées à des dispositifs expérimentaux à grande échelle.

Le projet reposera sur le test et la mise au point d'un protocole de séquençage RNA-seq limité aux extrémités 3' afin de réduire la quantité de lectures à produire par échantillon.

Vous participerez aux analyses bioinformatiques et biostatistiques à mettre en place pour répondre à trois grands objectifs :

- Identifier la profondeur de séquençage optimale par échantillon pour concilier pertinence de l'information biologique et coûts de séquençage, en comparaison avec des données RNAseq déjà produites sur les mêmes échantillons.
- Evaluer la performance et les conditions optimales de la méthode développée sur échantillon de sang total, adapté à la collecte semi-invasive applicable aux dispositifs d'élevage. Une analyse des données produites sera réalisée en comparaison avec des données déjà obtenues par microarray et MACE afin d'identifier les spécificités de chaque méthode et la performance du module de blocage

## **Analyse transcriptomique 3' RNAseq**

des transcrits de globuline, qui est une étape critique pour la réalisation d'études transcriptomiques par séquençage.

- Mettre au point les conditions de collectes et extractions de prélèvements sanguins à partir de diverses méthodes et identifier les spécificités de chaque méthode. Des librairies seront produites à partir de sang total, sérum et plasma afin de réaliser une identification et cartographie des profils d'expression de chacune des matrices.

Votre mission sera de réaliser les analyses bioinformatiques et biostatistiques des données générées dans le cadre du projet et déployer un pipeline d'analyse bioinformatique à l'aide d'un gestionnaire de workflow.

## **Compétences à mobiliser**

---

Le stage s'adresse à des étudiants de Master 2 ayant connaissance de l'analyses de données omiques, notamment transcriptomiques (si possible). Il/elle devra mobiliser les compétences suivantes :

- Etre à l'aise avec l'environnement Unix
- Utiliser un gestionnaire de workflow (nextflow, ...) pour créer des pipelines d'analyses
- Utiliser le cluster de calcul pour réaliser les traitements bioinformatiques
- Connaître et utiliser R pour la réalisation des analyses exploratoires et biostatistiques
- Connaître le domaine d'activité (biologie, transcriptomique)
- Rigueur et capacités de communication seront appréciées

## **Encadrement**

---

L'étudiant(e) sera encadré(e) par Yannick Lippi, Ingénieur en traitement de données et responsable du plateau Get-TRiX. Il aura à sa disposition les ressources informatiques disponibles sur le plateau et l'unité et sera amené à utiliser le cluster de calcul de la plateforme Genotoul-Bioinfo.