

 <p><b>AURAGEN</b> AUvergne Rhône-Alpes Génomique</p>	<p><b>PROFIL DE POSTE</b> <b>Biostatisticien(ne)</b></p> <p><b>AURAGEN</b></p>	<p>Document : AURAGEN Pages : 1</p>
<p>Resp Rédaction : CSS / CS</p>	<p>Resp.Valid : CSS/ DS</p>	<p>Resp. diffusion :</p>
<p>Date Rédaction :</p>	<p>Date Validation :</p>	<p>Date prise d'effet :</p>

Dossier de candidature (CV et lettre de motivations)	
<p>✓ <b>Transmission par courrier à :</b></p>	<p>✓ <b>Transmission par mail à :</b></p> <p><a href="mailto:pascal.roy@chu-lyon.fr">pascal.roy@chu-lyon.fr</a>  <a href="mailto:christine.vinciguerra@univ-lyon1.fr">christine.vinciguerra@univ-lyon1.fr</a>  avec copie à  <a href="mailto:cristina.heddi@univ-lyon1.fr">cristina.heddi@univ-lyon1.fr</a>  <a href="mailto:sihem.kheddouci@lip-lyon1.fr">sihem.kheddouci@lip-lyon1.fr</a>  <a href="mailto:virginie.valentin@chu-lyon.fr">virginie.valentin@chu-lyon.fr</a></p> <p><b>Attention tout dossier incomplet ne sera pas examiné.</b></p>

## ÉTABLISSEMENT

**AURAGEN AUvergne Rhône-Alpes Génomique**  
PLATEFORME DE SEQUENÇAGE TRES HAUT DEBIT A VISEE SANITAIRE  
Groupement Hospitalier Centre  
Bâtiment 7  
5 Place d'ARSONVAL  
69003 LYON cedex 03

## CARACTÉRISTIQUES GÉNÉRALES DU SERVICE/LABORATOIRE

**Le projet AURAGEN** est un des deux projets retenus dans le cadre du **Plan France Médecine Génomique 2025** (PFMG2025). Il intègre la plateforme de séquençage à très haut débit à visée sanitaire localisée sur l'hôpital Edouard Herriot, dont l'objectif est de produire, à terme, 18 000 équivalents génomes par an. Cette plateforme a pour missions de recevoir les prélèvements, extraire l'ADN, préparer les bibliothèques et réaliser des exomes, transcriptomes et génomes entiers.

Le **Service de Biostatistique / Bioinformatique** des **Hospices Civils de Lyon (HCL)** [Dir Pr Delphine Maucort-Boulch] et l'**Equipe Biostatistique Santé** [Dir Dr Muriel Rabilloud] du Département Statistiques et Modélisation pour les Sciences de la Santé du **Laboratoire de Biométrie et Biologie Evolutive** [Dir Dr Fabrice Vavre] représentent les valences Hospitalière et Universitaire d'une même équipe de Biostatisticiens/Bioinformaticiens implantés sur le CHU de Lyon. Les modèles statistiques développés sont appliqués à recherche clinique, à la recherche épidémiologique, et à l'analyse des données de grande dimension (données massives). Cette équipe accompagne aujourd'hui différents programmes d'implémentation de modèles d'Intelligence Artificielle sur le CHU de Lyon. Elle a développé des méthodes d'inférence statistique pour l'analyse des données de grande dimension prenant en compte les déterminants biologiques et instrumentaux de la variabilité [1-5], proposé une adaptation du modèle log-linéaire d'accord par paires (Alan Agresti [6]) pour l'analyse des données

génomiques [7], étudié l'apport des méthodes d'analyse du génome intégrant des réplicats [8], analysé les propriétés de méthodes de correction du biais d'optimisme [9-10]. Cette équipe est en charge d'analyses statistiques du programme AURAGEN, en particulier pour l'analyse du contrôle de qualité des étapes humides du séquençage de l'ADN réalisées sur la plateforme de séquençage à très haut débit du laboratoire AURAGEN.

## DÉFINITION GÉNÉRALE DU POSTE

---

La contribution attendue concerne la mise en place d'études et le développement de modèles statistiques pour le suivi longitudinal du contrôle de qualité des étapes humides du séquençage de l'ADN. Il s'agit de fournir des indicateurs de qualité à l'étape de production, de développer des modèles pour le suivi longitudinal de ces indicateurs, et d'analyser leurs déterminants. Des modèles log-linéaires généralisés à effet mixtes intégrant des fonctions souples (GAM, splines cubiques restreints) seront utilisés. Le / la biostatisticien(ne) en charge de ces analyses travaillera en collaboration avec les acteurs du laboratoire humide et les bioinformaticiens(nes) de la plateforme AURAGEN.

## PRINCIPALES ACTIVITÉS DU POSTE

---

- Le/la biostatisticien(ne) :
  - Implémentera des études pour le suivi longitudinal du contrôle de qualité de la partie humide du séquençage de l'ADN (dessin d'étude, calibration, mise en place, suivi, programmation statistique, outil de suivi à long terme) en interaction avec le biologiste responsable
  - Effectuera les analyses statistiques des résultats de ces contrôles de qualité, rédigera les rapports d'analyses statistiques correspondants
  - Assurera une veille technologique sur les outils et les méthodes
  - Aidera les biologistes de la plateforme AURAGEN à l'interprétation des résultats des contrôles de qualité effectués dans le cadre de réunions pluridisciplinaires.
- Livrables
  - Plans d'analyses statistiques des études de contrôle de qualité réalisées (mensuels)
  - Rapports d'analyse des études de contrôle de qualité réalisées (trimestriels)
  - Contributions aux publications scientifiques (semestrielles)

## DIPLOME, CONNAISSANCE, COMPETENCES LIENS FONCTIONNELS

---

- **Diplôme :**
  - Diplôme d'ingénieur en Biostatistique / Bioinformatique, ou d'un master 2 de Biostatistique/Bioinformatique, ou d'un Doctorat de Biostatistique/Bioinformatique
- **Compétences requises**
  - Une expérience professionnelle d'un minimum de 3 ans est souhaitée
  - ⊖ Expérience en Analyse Biostatistique / Bioinformatique des données omiques
  - Expérience en développement de schémas d'études adaptés à l'analyse de données de séquençage de l'ADN
  - Expérience approfondie en programmation sous R
  - Notions de programmation Python
- **Capacités organisationnelles et interpersonnelles**
  - Motivation et dynamisme
  - Organisation et rigueur
  - Autonomie et travail en équipe
  - Aptitude à communiquer/ vulgariser son travail

## LIENS HIERARCHIQUES

---

- Le Directeur du Département Statistiques et Modélisation pour les Sciences de la Santé (Pr Pascal Roy)
- La Directrice du LBMMS AURAGEN (Pr Christine Vinciguerra)
- Le Directeur Scientifique AURAGEN (Pr Jean-Yves Blay)

## LIENS FONCTIONNELS

---

- Les responsables des différents axes du Programme AURAGEN
- Le biologiste médical et l'ingénieur de la plateforme de production
- Les bio-informaticiens impliqués dans la plateforme AURAGEN
- Les biologistes impliqués dans les différentes activités et les différents services hospitaliers
- Les cliniciens
- Les ingénieurs qualité et data manager
- Les techniciens de laboratoire

## CONDITIONS DE TRAVAIL - EVOLUTION DU POSTE

---

- Le poste est localisé dans le Service de Biostatistique-Bioinformatique des Hospices Civils de Lyon (HCL) avec une présence régulière sur le laboratoire humide
- Evolution du poste en fonction de l'employeur partenaire du GCS-AURAGEN

## REFERENCES

---

1. Klich A, Mercier C, Gerfault L, Grangeat P, Beaulieu C, Degout-Charrette E, Fortin T, Mahé P, Giovannelli JF, Charrier JP, Giremus A, Maucort-Boulch D, Roy P. Variance component analysis to assess protein quantification in biomarker validation: application to selected reaction monitoring-mass spectrometry. *BMC Bioinformatics*. 2018 Mar 1;19(1):73.
2. Mercier C, Klich A, Truntzer C, Picard V, Giovannelli JF, Ducoroy P, Grangeat P, Maucort-Boulch D, Roy P. Variance component analysis to assess protein quantification in biomarker discovery. Application to MALDI-TOF mass spectrometry *Biom J*. 2018 Mar;60(2):262-274.
3. Dridi N, Giremus A, Giovannelli JF, Truntzer C, Hadzagic M, Charrier JP, Gerfault L, Ducoroy P, Lacroix B, Grangeat P, Roy P. Bayesian inference for biomarker discovery in proteomics: an analytic solution. *EURASIP J Bioinform Syst Biol*. 2017 Dec;2017
4. Roy P, Truntzer C, Maucort-Boulch D, Jouve T, Molinari N. Protein mass spectra data analysis for clinical biomarker discovery: a global review. *Brief Bioinform*. 2011 Mar;12(2):176-86.
5. Mercier C, Truntzer C, Pecqueur D, Gimeno JP, Belz G, Roy P. Mixed-model of ANOVA for measurement reproducibility in proteomics. *Journal of proteomics*. 2009;72(6):974-981.
6. Agresti A. *Categorical Data Analysis*, 3rd edition. Wiley, 2013.
7. Elsensohn MH, Leblay N, Dimassi S, Campan-Fournier A, Labalme A, Roucher-Boulez F, Sanlaville D, Lesca G, Bardel C, Roy P. Statistical method to compare massive parallel sequencing pipelines. *BMC Bioinformatics* 2017;18:139
8. Zhai Y, Bardel C, Vallée M, Iwaz J, Roy P. Performance comparisons between clustering models for reconstructing NGS results from technical replicates. *Front Genet*. 2023 Mar 16;14:114814
9. Zhao Y, Dantony E, Roy P. Optimism bias correction in omic studies: assessment of penalized methods on simulated data *OMICS: A Journal of Integrative Biology*. 2019;23:207-213.
10. Truntzer C, Maucort-Boulch D, Roy P. Comparative optimism in models involving both classical clinical and gene expression information. *BMC Bioinformatics*. 2008 Oct