

#49798 – Le défi du suivi des contrôles dans un laboratoire de séquençage à très haut débit pour le diagnostic : mise en place des protocoles et outils de suivi à AURAGEN



Estelle Champion^{1,2,3}, Anne Thomas¹, Joanna Pautonnier^{2,3,4}, Nicolas Pons¹, Olivier Chenavier¹, Christine Vinciguerra^{1,5}, Alain Viari¹, Pascal Roy^{2,3,5}, Carole Ferraro-Peyret^{1,5}

¹GCS Auragen, ²Service de Biostatistique-Bioinformatique, ³Hospices Civils de Lyon, ⁴Laboratoire de Biométrie et Biologie Evolutive UMR5558, F-69003 Lyon, France, ⁵Université Lyon 1, F-69100 Villeurbanne, France.

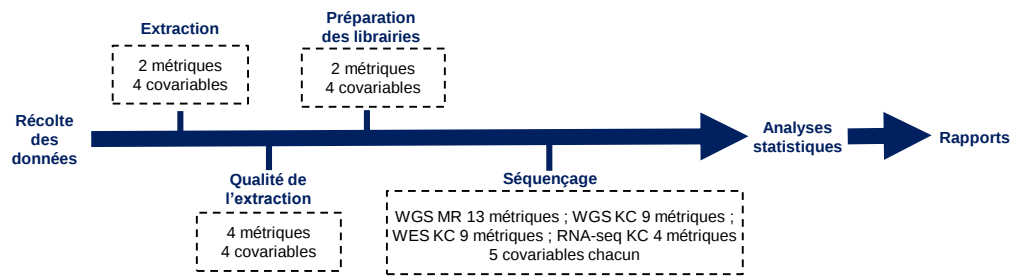
Introduction

AURAGEN, laboratoire de Biologie Médicale du Plan France Médecine Génomique 2025 a une activité dédiée au séquençage à très haut débit pour les patients atteints de maladies rares (MR) et de cancers (KC). Cette activité, nouvelle dans le domaine du diagnostic, a nécessité le développement de **nouveaux outils et protocoles de suivi de la qualité**, basé sur l'analyse conjointe de métriques de qualité et de covariables techniques afin de **détecter précocement les non-conformités et les dérives**.

Matériel et méthodes

Un dispositif de **contrôle qualité** a été mis en place au laboratoire afin de surveiller les **performances des examens**, de l'extraction de l'ADN/ARN au séquençage, et de détecter les dérives. Les données issues du Système de Gestion de Laboratoire (SGL), du post-séquençage Illumina et des pipelines bio-informatiques incluent :

- **des métriques de qualité** : indicateurs chiffrés décrivant la qualité des données ou d'une étape du processus, utilisés pour vérifier que les résultats sont fiables et conformes aux attentes (ex. profondeur, taux d'erreur).
- **des covariables techniques** : Variables décrivant les conditions techniques de production des données, susceptibles d'influencer les métriques de qualité sans lien avec le biologique (ex. Suivi des lots, robots, séquenceurs, opérateur, hygrométrie, température...).



1

- Analyses descriptives
 - Analyses temporelles
- ➡ positionnent les données par rapport à des seuils de contrôle internes (QC) guidant les actions correctives (Figure a).

2

L'algorithme **PELT** est utilisé pour détecter des ruptures temporelles révélant des dérives précoces (figure b).

Algorithme = **PELT (Pruned Exact Linear Time)**

$$\min_{k, i} \left(\sum_{t=1}^k C(x_{t-1} + 1; v) + j/K \right)$$

Fonction de coût = **RBF (réseaux à fonctions de base radiales)**

$$C_{RBF}(x_{a,b}) = \sum_{i=1}^n ||x_i - \mu_{a,b}||^2$$

Centre estimé du segment ab

3

Les performances des séquenceurs sont comparées à l'aide de **modèles additifs généralisés (GAM)** basés sur des splines pénalisées (figure c).

Formule 1 : $Metric \sim s(Date, k = k_{max})$

Formule 2 : $Metric \sim s(Date, k = k_{max}) + \lambda \beta$

Spline (fonction lisse composée de morceaux de polynômes)

Contrôle la complexité maximale de la spline

Covariable technique (Ex : Hygrométrie)

Résultats

Figure a

Détection des non-conformités de la profondeur médiane en WGS Maladies Rares – séquençés sur 2 Novaseq X Plus



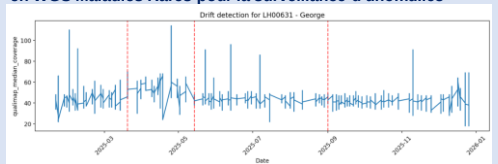
Les écarts aux seuils de contrôle (QC1, QC2; exemples d'écart cerclés en rouge) sont enregistrés automatiquement dans un rapport pour être investigués.

Impact de la montée de version du logiciel des séquenceurs Novaseq X Plus augmentant la quantité totale des données par flow cell.

Impact de l'augmentation du nombre d'échantillons par flow cell en conservant une profondeur médiane suffisante (> QC1).

Figure b

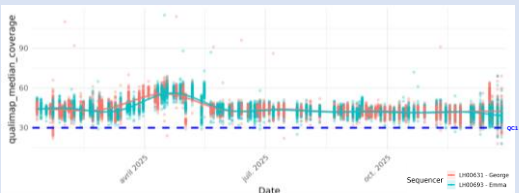
Détection de ruptures temporelles de la profondeur médiane en WGS Maladies Rares pour la surveillance d'anomalies



Les ruptures (traits hachés rouges) sont identifiées lors de l'analyse du contrôle : elles peuvent être la conséquence de maintenance, de changement dans les protocoles, de nouvelles versions de logiciel d'analyse, de réelles dérives... etc.

Figure c

Comparaison des performances des 2 séquenceurs Novaseq X Plus par modélisation de la profondeur médiane, en WGS Maladies Rares, à l'aide de splines pénalisées (GAM, formule 1)



Pas de différence observée sur l'année 2025 concernant la performance des deux séquenceurs X Plus.

Conclusion & Perspectives

- ✓ Surveillance continue des performances
- ✓ Détection précoce des dérives
- ✓ Améliore la traçabilité, la réactivité opérationnelle et la robustesse du processus analytique
- ➔ Intégration des covariables ($\lambda \beta$) dans les modèles GAM
- ➔ Développement d'approches prédictives pour consolider un contrôle qualité proactif en routine de biologie moléculaire à très haut



Ce travail est rendu possible grâce au plan France médecine génomique 2025.

Contacts :
estelle.champion@chu-lyon.fr;
anne.thomas@chu-lyon.fr;
carole.ferraro-peyret@chu-lyon.fr