



Le réseau régional d'ingénieurs en bioinformatique de Lille et le PPF bioinformatique vous convient à une conférence Mercredi 28 novembre 2012, 10h30-12h, Amphithéâtre de l'Institut de Biologie de Lille, 1, rue du Pr Calmette, LILLE.

## **Comment diversifier les données de séquençage à haut débit pour optimiser la bioinformatique ?**

### **Les stratégies hybrides PacBio & Roche 454/Illumina HiSeq**

*Dr Christophe Meynier et Sabrina Benoussaidh - GATC Biotech France & Belgique*

Les nouvelles technologies de séquençage (NGS) ont permis d'améliorer drastiquement la caractérisation des génomes avec la génération de séquence en haut débit et ce à faible coût. Malgré cela, l'assemblage d'un génome reste souvent problématique et diffère en fonction des technologies utilisées. En effet la qualité d'un assemblage de novo est influencée par différents facteurs inhérent à la structure intrinsèque du génome et en parallèle à la longueur des fragments, au nombre de séquences et à la qualité de ces dernières.

La technologie PacBio RS fournit les plus longues séquences jamais obtenues (3500 bases en moyenne) ce qui simplifie et améliore la création de contigs. Des études ont montré qu'en corrigeant les séquences générées par le PacBio RS avec des données issues de secondes générations de séquenceurs (Roche 454, HiSeq), nous optimisons alors les avantages de chacune de ces technologies (couverture, fiabilité des bases et longueurs des séquences). En diminuant le nombre de contigs tout en augmentant leurs tailles, les chercheurs ont désormais les outils pour obtenir des assemblages de génomes plus rapidement, à moindre coût ou pour obtenir des informations sur des génomes partiellement séquencés jusque-là disponibles uniquement manuellement (via PCR ou marche sur l'ADN).

Notre présentation montre des résultats d'assemblages obtenus avec différentes technologies, les compare et démontre l'intérêt d'association de technologies ensemble. En outre, nous discuterons également des technologies Illumina HiSeq 2000 et Roche 454 au travers d'applications spécifiques telles que le reséquencage de génomes entiers, le séquençage de transcriptomes (RNAseq) et le ChipSeq.

Contact : [sophie.gallina@univ-lille1.fr](mailto:sophie.gallina@univ-lille1.fr)

