



Le mot de la direction

En parcourant cette lettre, vous verrez que ce début d'année a été marqué par l'extension de nos capacités de séquençage : allongement des séquences sur Illumina NextSeq 2000 et plus de débit chez Oxford Nanopore Technologies grâce à l'acquisition du P2 Solo (photo ci-dessous), avec dans les deux cas des améliorations notables des qualités de séquençage.



Ces nouveautés nous poussent à optimiser toujours plus nos protocoles et la plus grande accessibilité de longues séquences de haute qualité promet de nouvelles applications.

Dans ce contexte, pouvoir compter sur une équipe solide est essentiel. Afin de pouvoir répondre efficacement aux besoins de production de données des projets tout en développant de nouvelles approches méthodologiques, nous recrutons un(e) Assistant(e) Ingénieur(e) en biologie moléculaire spécialisé(e) en séquençage Illumina, pour 12 mois renouvelables (plus d'information sur <https://pgtb.fr/>).

Enfin, nous félicitons Zoé Delporte qui a brillamment réussi le concours INRAE d'assistante ingénierie et qui nous rejoint donc durablement à la PGTB.

600 cycles sur NextSeq 2000 !

Illumina a récemment commercialisé sur NextSeq 2000 la chimie 600 cycles permettant de séquencer en 2x300pb, qui n'existeait jusque-là que sur les kits V3 du MiSeq et qui souffrait d'une baisse notable de qualité en fin de séquence.

Plus de longueur de séquençage facilite l'assemblage et le mapping et permet également de séquencer des régions d'intérêt plus longues qu'avec la chimie standard 300 cycles, comme par exemple certaines régions du 16S ou de l'ITS, pour les approches de métagénomique ciblée, ou encore des microsatellites plus longs, donc plus informatifs, qui sont caractérisés avec notre approche [SSRseq](#).



La PGTB a testé et validé cette nouvelle chimie et les résultats sont très bons : près de 90% des bases avec une qualité supérieure à Q30 sur l'ensemble des données produites (Illumina garantit 80%), soit 10 à 15% de mieux que des données équivalentes obtenues sur MiSeq V3. Pour un coût de run à peine plus élevé qu'un run MiSeq, on obtient donc 4 à 5 fois plus de données bien plus qualitatives, sur toute la longueur. La PGTB travaille actuellement sur une stratégie d'indexage UDI qui permettra à moyen terme de passer plus de 1000 échantillons dans un même run quand une profondeur élevée de séquençage n'est pas nécessaire.

Cette chimie 600 cycles est disponible sur les flow cell P1 (~100 millions de paires de reads) et sur les flow cell P2 (~300 millions de paires de reads). Ce sont désormais 11 kits de séquençage différents qui sont disponibles sur NextSeq 2000, de 1x50 à 2x300pb, sur 3 flow cell différentes.

Enfin, Illumina a confirmé le 30 mai 2023 le déploiement dès le début 2024 de la chimie XLEAP-SBS sur NextSeq 2000 (plus de détails dans notre [précédente Newsletter](#)), ainsi que la commercialisation des flow cell P4 produisant 500 Gb avec un coût de séquençage réduit.

Arrivée du P2 Solo et une qualité encore améliorée chez Oxford Nanopore

Nous mettons en service ce mois-ci le nouveau séquenceur PromethION 2 Solo (P2 Solo) d’Oxford Nanopore Technologies (ONT), acquis grâce à un financement de la Fédération Bordeaux Research Facilities.

Ce séquenceur produit ~100 Gb (giga bases nucléotidiques) en moyenne par flow cell PromethION, ouvrant ainsi la possibilité de travailler sur des génomes de taille importante ou sur davantage d’échantillons simultanément à moindre coût.

Le P2 Solo se présente comme une extension du séquenceur GridION présent à la PGTB depuis plus de 6 ans. Le P2 Solo ne possède pas d’informatique embarquée mais est connecté au séquenceur GridION qui le pilote et traite le signal issu des nanopores. Nous sommes ici donc loin de la logique d’obsolescence programmée fréquent dans le domaine hautement technologique de la génomique. La démarche d’ONT d’amélioration constante de la qualité et de la quantité de données produites par l’évolution des consommables, réactifs et algorithmes de traitement du signal, tout en limitant la nécessité de racheter de nouvelles machines, est à saluer car elle favorise la soutenabilité économique et environnementale de nos activités.

Les flow cell MinION que nous utilisons en routine sur le GridION possèdent 2 048 nanopores pour un débit moyen de 15 Gb par run, soit le génome humain séquencé avec une profondeur moyenne de 5X. Une flow cell PromethION possède 12 000 pores pour un débit moyen de 100 Gb par run (un génome humain séquencé en 33X). Le P2 Solo a la capacité de séquencer sur deux flow cell PromethION en parallèle, pour un débit maximal de ~200 Gb, soit l’équivalent de 13 flow cells MinION. Avec un débit et un coût au Gb similaire au séquenceur court-fragment Illumina NextSeq 2000, le P2 Solo nous permettra de proposer une offre de séquençage haut-débit compétitive pour du séquençage long fragments de haute qualité.



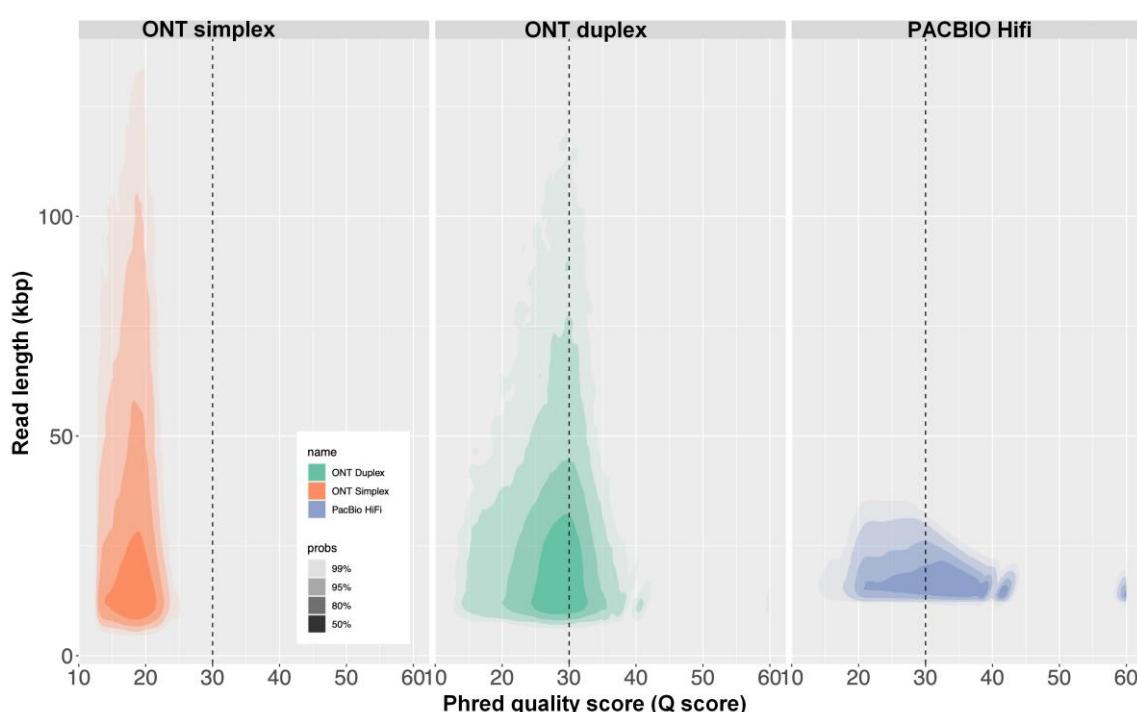
Outre ces nouvelles capacités en terme de débit, la technologie ONT a encore fait des progrès considérables sur la qualité de séquençage au cours des derniers mois. Lors de notre [Newsletter de novembre 2022](#), nous avions présenté la chimie Nanopore V112 (Q20+) que nous avions testée. Cette chimie transitoire proposait une qualité améliorée au dépend d’un débit plus faible. Elle a de nouveau évolué avec l’arrivée de la chimie V114, accompagnée de nouveaux pores R10.4.1. La qualité des données reste élevée tout en augmentant le débit pour atteindre le plus haut rendement de l’ancienne chimie V109.

Plus encore, l’un des grands axes de progression actuel est la capacité d’analyser les deux brins de l’ADN (séquences duplex) permettant de générer un consensus de haute qualité (Q30 soit 99.9% de précision) pour une partie des brins d’ADN. Aujourd’hui, cette chimie permet d’obtenir entre 5 à 15 % de reads duplex de haute qualité (Q30) sur l’ensemble des jeux de données.

La qualité des séquences est maintenant très similaire entre PacBio HiFi et ONT duplex : PacBio reste encore légèrement plus qualitatif mais ONT propose toujours des séquences plus longues. La capacité d’obtenir des séquences duplex très longues permet d’envisager des assemblages de génomes de très haute qualité à partir des seules séquences duplex ONT.

Par exemple, un assemblage telomere-to-telomere (14 contigs pour 12 chromosomes et génomes mitochondrial et chloroplastique) de la tomate (0.82Gb) a pu être réalisé à partir d'une profondeur de 46X de séquences duplex et de 42X de ultra-long (> 100Kb) séquences simplex (A. Wittenberg, https://lstu.fr/t2t_ontduplex).

Lors de la dernière conférence technique London Calling 2023 au mois de mai, Clive Brown (CEO ONT) a présenté des avancées très significatives issues des améliorations des flow cells PromethION, des tampons et du basecaller : séquences simplex Q22 et duplex Q32, augmentation du débit (> 150 Gb). Plus impressionnant, une toute dernière version de flowcell « HD », disponible en early access dès cet été, prévoit environ 80% de séquences duplex, ce qui devrait donner un rendement d'environ 50 Gb de séquences duplex (Q32) par flow cell PromethION. Cette avancée, qui sera disponible dans les prochains mois, rendra l'assemblage de génomes de haute qualité, notamment des petits génomes ou des métagénomes, nettement plus accessible.



Répartition de la qualité (score Phred en abscisse) des séquences simplex, duplex et PacBio HiFi en fonction de leur longueur (millier de nucléotides en ordonnée), d'après la présentation « Single platform telomere-to-telomere assemblies: examples from the tomato and maize genomes » d'Alexander Wittenberg (https://lstu.fr/t2t_ontduplex)

