

Illumina Human Comprehensive Panel

Couverture améliorée des
régions complexes dans les
gènes codant les protéines avec
Illumina Complete Long Read
Prep with Enrichment, Human



Introduction

Lors de la réalisation du séquençage du génome entier (WGS, Whole-genome Sequencing) humain, une petite fraction des régions géniques peuvent être difficiles à cartographier avec seulement des lectures courtes. Les régions complexes comprennent les régions hautement homologues ou répétitives, les variants structurels, pseudogènes et grandes insertions/suppression (indels) complexes. Dans ce type de situation, le séquençage à lectures longues peut compléter les données du WGS à lectures courtes standard pour fournir une couverture améliorée afin de surmonter les difficultés posées par les régions complexes et faciliter la mise en phase des variants et la définition des haplotypes.

La technologie Illumina Complete Long Reads utilise un flux de travail de séquençage de nouvelle génération (SNG) standard pour générer des séquences contiguës à lectures longues sur les systèmes de séquençage d'Illumina (figure 1). Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel offrent une approche rentable et évolutive pour aider les chercheurs à surmonter les difficultés posées par les régions complexes connues du génome. Les lectures longues ciblées sont analysées avec des données de WGS standard de haute précision pour permettre la cartographie de plus de régions.

Panel de sondes complet et optimisé

Illumina Human Comprehensive Panel est un panel de sondes de capture hybride optimisé conçu pour prendre en charge les régions géniques qui peuvent bénéficier d'une cartographiabilité accrue des lectures plus longues. Le panel cible les régions à faible cartographiabilité sur > 6 500 gènes codant les protéines (tableau 1)¹ pour permettre des lectures longues complémentaires qui améliorent la résolution des zones difficiles à cartographier à l'aide de méthodes de WGS à lectures courtes standard.

Tableau 1 : Paramètres du Human Comprehensive Panel¹

Système	Série NovaSeq X NovaSeq 6000 System
Taille de la région cible	> 95 Mb
Débit de séquençage par échantillon ^a	90 à 120 Gb
Nbre de sondes	~ 40 000
Nbre de gènes ciblés	> 6 500
Type d'échantillon	ADN génomique
Entrée d'ADN recommandée	50 ng
Durée totale de préparation des bibliothèques	~ 2 jours
Durée de manipulation	~ 6 h
Multiplexage	Jusqu'à 64 échantillons
N50	6,3 kb
Bloc de phase N50	15,6 kb
Uniformité	95 %
Enrichissement par lectures élargies ^c	83 %
% de SNV hétérozygotes mis en phase ^d	98 %

- Données générées à l'aide de 50 ng d'ADN génomique HG002 (Coriell, référence n° NA24385). Les performances peuvent varier en fonction de l'entrée d'ADN et de la qualité de l'échantillon.
- Requiert une analyse de séquençage de 2×150 pb, générant environ $30\times$ la couverture finale d'Illumina Complete Long Reads.
- Uniformité de la couverture calculée en % > $0,2\times$ la moyenne.
- Enrichissement par lectures élargies calculé comme $100 \times$ (lectures élargies alignées sur la cible/nombre total de lectures alignées).
- SNV : variants mononucléotidiques (Single Nucleotide Variant).

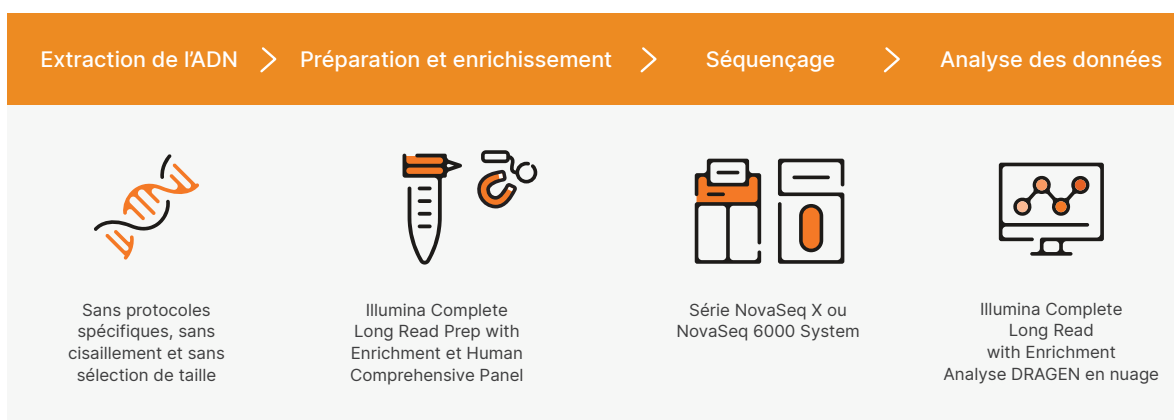


Figure 1 : Élément d'un flux de travail intégré : accédez à des données de WGS à lectures longues ciblées et rentables à l'aide d'un protocole de préparation de bibliothèques optimisé et évolutif, d'une chimie de séquençage Illumina éprouvée et d'une analyse secondaire DRAGEN. Requiert $\geq 30\times$ plus de données de WGS à lectures courtes standard du même échantillon pour l'analyse. Il est possible d'utiliser les fichiers FASTQ d'un échantillon d'analyse antérieur.

En développant ce panel, Illumina a pris en compte l'ensemble complet de plus de 20 000 gènes codant les protéines, notamment les introns et les régions non traduites (UTR, Untranslated Region), et a ciblé 70 % de toutes les paires de bases difficiles à cartographier de ces régions. Les gènes qui sont entièrement couverts par des lectures courtes seules ont été exclus. Ce panel a été testé en laboratoire humide et optimisé pour maximiser l'efficacité et les performances de conception et de séquençage.

Analyse et séquençage hautement évolutifs

Le flux de travail Illumina Complete Long Read Prep with Enrichment est hautement évolutif et facile à automatiser pour prendre en charge le WGS complet de plus d'échantillons ([tableau 2](#)). Sur NovaSeq^{MC} X Plus System, les utilisateurs peuvent générer jusqu'à 15 000 génomes de haute précision par an*. Pour réduire les exigences en matière de création de lots, séquencez moins d'échantillons sur des consommables à débit inférieur comme la Flow Cell 1.5B NovaSeq X. Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel peuvent être utilisés pour accroître le volume de données de WGS existant en tant qu'outil à réflexes pour parvenir à une meilleure détection des variants.

* Débit potentiel lorsqu'Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel sont utilisés avec NovaSeq X Plus System, l'analyse d'une double Flow Cell avec des Flow Cell 25B.

L'analyse des données pour Illumina Complete Long Read Prep with Enrichment, Human est disponible sous forme d'application dans BaseSpace^{MC} Sequence Hub ou via Illumina Connected Analytics. Les données sont transférées dans le nuage et analysées avec $\geq 30\times$ plus de données de WGS à lectures courtes standard provenant du même échantillon (générées précédemment ou en parallèle). Le pipeline DRAGEN^{MC} compile les résultats dans un seul ensemble de fichiers de sortie, notamment les paramètres d'appel ciblés par DRAGEN³.

Couverture améliorée et rentable

Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel améliorent la précision de l'appel des variants et la couverture entre les régions cibles, par rapport au WGS à lectures courtes standard ([figure 2](#), [figure 3](#), [tableau 3](#)). Les lectures longues ciblées à l'aide de Human Comprehensive Panel offrent une précision complète du génome entier avec un score F1 (SNV + indels) de 99,87 %, une amélioration considérable par rapport aux lectures courtes seules. Cela fournit des performances comparables à celles d'Illumina Complete Long Read Prep, Human et d'autres génomes entiers à lectures longues dans les gènes codant les protéines, tout en réduisant les coûts et en augmentant le débit ([tableau 3](#)).

Tableau 2 : Débit d'échantillons recommandé pour générer 30× la couverture finale pour Human Comprehensive Panel et Illumina Complete Long Read Prep with Enrichment, Human^{a,b,c,d}

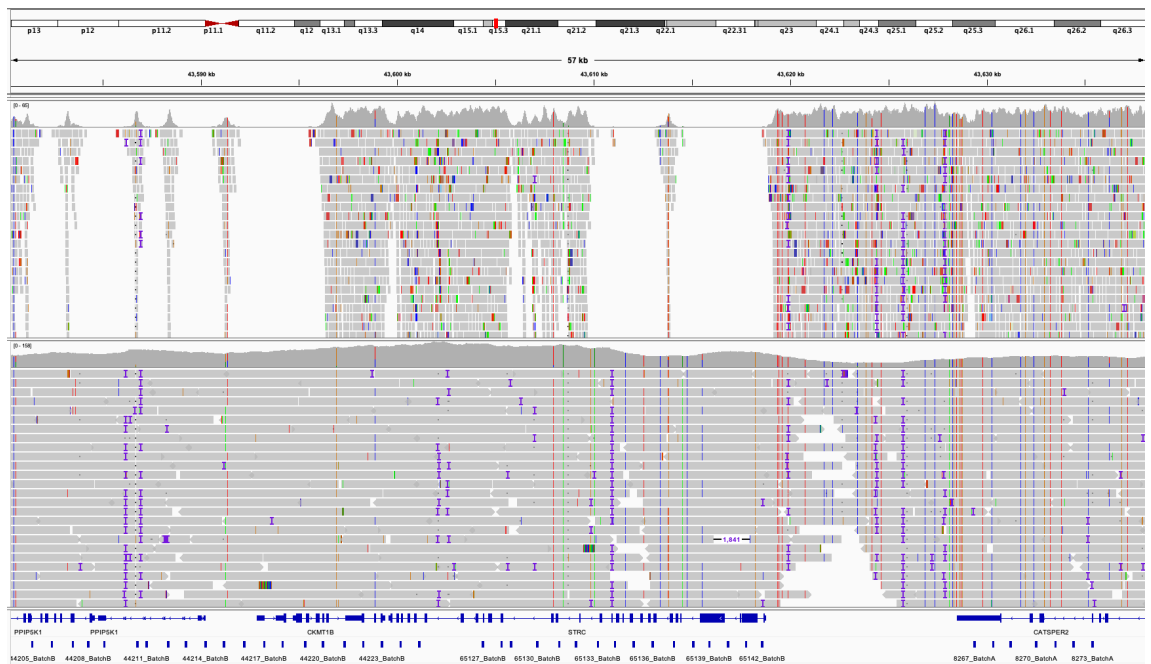
Trousse de réactifs de 300 cycles	NovaSeq 6000				NovaSeq X		
	SP	S1	S2	S4	1.5B	10B	25B
Échantillons par Flow Cell	2	4	10	24	4	24	64
Débit par Flow Cell	~ 250 Gb	~ 500 Gb	~ 1,25 Tb	~ 3 Tb	~ 500 Gb	~ 3 Tb	~ 8 Tb
Durée de l'analyse	~ 25 h	~ 25 h	~ 36 h	~ 44 h	~ 21 h	~ 25 h	~ 48 h

- a. La taille de la région cible du Human Comprehensive Panel est > 95 Mb et nécessite un débit de séquençage compris entre 90 et 120 Gb par échantillon.
 b. Requiert une analyse de séquençage de 2×150 pb, générant environ 30× la couverture finale d'Illumina Complete Long Reads.
 c. Requiert 30× plus de données du génome entier humain à lectures courtes standard du même échantillon pour l'analyse. [Illumina DNA PCR-Free Prep](#) est recommandé. Les trousse de séquençage du génome entier (WGS) tierces sont également compatibles. La librairie non marquée n'a pas besoin d'être préparée ou séquencée en parallèle; il est possible d'utiliser les fichiers FASTQ d'un échantillon d'analyse antérieur.
 d. Le séquençage des librairies Illumina Complete Long Read sur les plateformes NovaSeq pourrait faire passer le score Q30 rapporté d'une analyse en deçà de la valeur NovaSeq spécifiée. Il ne s'agit pas d'un problème de performance en lien avec l'analyse de séquençage ou la librairie.

A.

WGS à lectures courtes standard

Illumina Complete Long Read Prep with Enrichment



B.

WGS à lectures courtes standard

Illumina Complete Long Read Prep with Enrichment

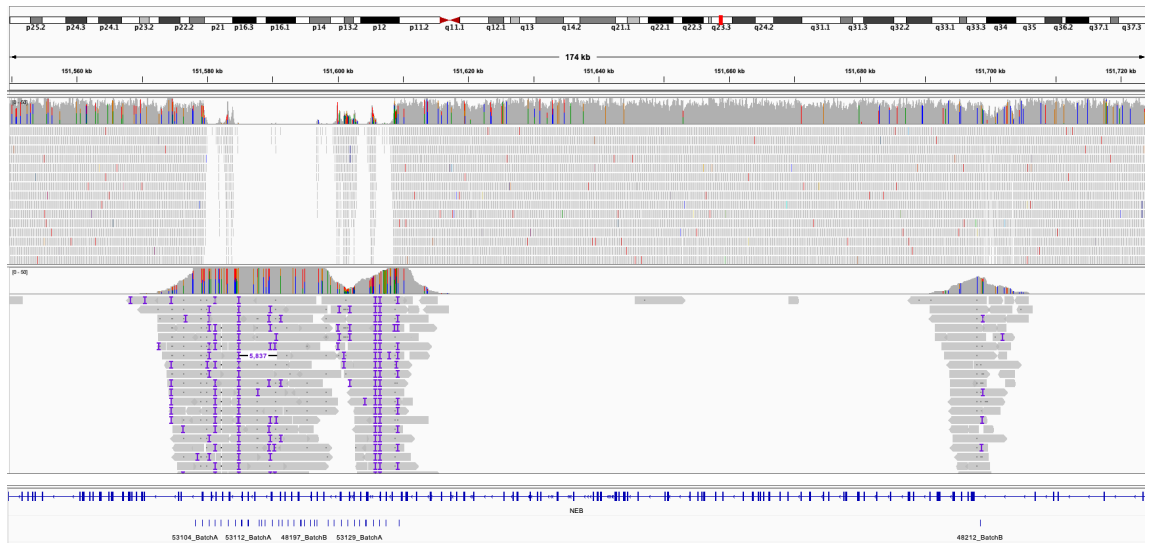


Figure 2 : Surmontez les difficultés posées par les régions difficiles à cartographier à l'aide de lectures longues ciblées avec Human Comprehensive Panel : Illumina Complete Long Read Prep with Enrichment, Human et Human Comprehensive Panel aident à améliorer la couverture dans les régions géniques complexes pour compléter le WGS humain à lectures courtes standard. Tracés d'Integrative Genomics Viewer (IGV) de (A) *STRC* et (B) *NEB* à l'aide du WGS à lectures courtes standard et d'Illumina Complete Long Reads with Enrichment.

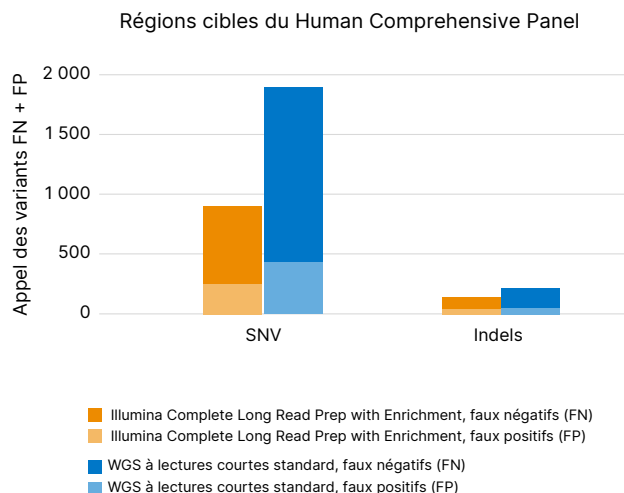


Figure 3 : Lectures longues ciblées pour améliorer la précision de l'appel des variants : appels des variants faux négatifs (FN) et faux positifs (FP) pour les SNV et les indels dans les régions géniques ciblées par Illumina Human Comprehensive Panel à l'aide d'Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel (orange) par rapport au WGS à lectures courtes standard (bleu).

Résumé

Illumina Human Comprehensive Panel permet une couverture des lectures longues de haute précision de la petite partie des régions géniques qui sont difficiles à cartographier avec des lectures courtes seules. Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel offrent un test de génome entier optimisé et rentable qui complète le WGS d'Illumina et se concentre sur les lectures longues là où elles offrent la plus grande valeur avec une solution de flux de travail complète.

En savoir plus

[Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel](#)

[Technologie de séquençage à lectures longues](#)

Tableau 3 : Performances du Human Comprehensive Panel

Précision pour les régions cibles du Human Comprehensive Panel ^a		
	Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel ^b	WGS à lectures courtes standard ^c
Score F1 (SNV)	99,54 %	99,04 %
Score F1 (indels)	99,47 %	99,15 %
Score F1 (VS)	80,43 %	60,18 %

Précision pour le génome entier, toutes les régions de référence ^d score F1 (SNV + indels)			
	Illumina Complete Long Read Prep with Enrichment et Human Comprehensive Panel ^b	Illumina Complete Long Read Prep, Human ^e	WGS à lectures courtes standard ^c
	99,87 %	99,90 %	99,84 %

a. Autosomes uniquement. SNV, variants mononucléotidiques; VS, variants structurels.

b. Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.

c. Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.

d. Comme mesuré par rapport à toutes les solutions proposées à l'aide des données de référence du PrecisionFDA Truth Challenge v2², données internes pour DRAGEN v4.

e. Illumina Complete Long Read Prep, Human + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.

Références

- Bekritsky MA, Bekritsky MA, Colombo C, Eberle MA. Identifying genomic regions with high quality single nucleotide variant calling. illumina.com/science/genomics-research/articles/identifying-genomic-regions-with-high-quality-single-nucleotide-.html. Publié en 2021. Consulté le 30 août 2023.
- PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Consulté le 2 octobre 2023.
- Roessler K. Illumina Complete Long Reads software analysis workflow for human WGS. illumina.com/science/genomics-research/articles/complete-long-read-software-analysis.html. Publié en 2023. Consulté le 22 septembre 2023.

Renseignements relatifs à la commande

Produit	N° de référence
Illumina Complete Long Read Prep with Enrichment, Human (24 échantillons)	20113832
Illumina Complete Long Read Prep with Enrichment, Human (96 échantillons)	20113833
Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel (24 échantillons)	20113834
Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel (96 échantillons)	20113835
Illumina Human Comprehensive Panel (24 échantillons)	20113836
Illumina Human Comprehensive Panel (96 échantillons)	20113837
Illumina Unique Dual Indexes, LT (48 index, 48 échantillons)	20098166
Illumina DNA/RNA UD Indexes Set A, Tagmentation (96 index, 96 échantillons)	20091654
Illumina DNA/RNA UD Indexes Set B, Tagmentation (96 index, 96 échantillons)	20091656
Illumina DNA/RNA UD Indexes Set C, Tagmentation (96 index, 96 échantillons)	20091658
Illumina DNA/RNA UD Indexes Set D, Tagmentation (96 index, 96 échantillons)	20091660
Illumina Analytics – 1 iCredit	20042038



Numéro sans frais aux États-Unis : + (1) 800 809-4566 | Téléphone : + (1) 858 202-4566
techsupport@illumina.com | www.illumina.com

© 2024 Illumina, Inc. Tous droits réservés. Toutes les marques de commerce sont la propriété d'Illumina, Inc. ou de leurs détenteurs respectifs. Pour obtenir des renseignements sur les marques de commerce, consultez la page www.illumina.com/company/legal.html.
M-GL-02191 FRA v1.0