

# ILLUMINA GENOMICS ARCHITECTURE PERMET LA RÉALISATION D'ÉTUDES POPGEN AVEC ILLUMINA DNA PCR-FREE PREP

Des performances de  
préparation de bibliothèques  
uniformes grâce à  
l'automatisation des flux  
de travail de séquençage  
du génome entier



## Introduction

La technologie de séquençage de nouvelle génération (SNG) d'Illumina fournit des données précises et de haute qualité et permet un large éventail d'applications en génomique, transcriptomique et épigénomique. Le flux de travail de SNG va de la préparation des bibliothèques au séquençage, en passant par l'analyse et l'interprétation des données. La gamme de produits d'Illumina comprend une variété de composants tout au long du flux de travail qui traitent les nombreux domaines d'application possibles. En raison de la diversité des produits d'Illumina, certains nouveaux clients trouvent que l'intégration de composants individuels à un flux de travail unique (isolation du matériel génétique par l'entremise de la génération de rapports sur les variants) est laborieuse et prend du temps. De plus, bien que de nombreux flux de travail d'Illumina soient compatibles avec l'automatisation, certains clients sont découragés par l'expertise technique requise pour intégrer et optimiser les méthodes automatisées dans leurs pratiques existantes.

Illumina Genomics Architecture (IGA) répond à ces défis en offrant un **cadre flexible**, modulaire et normalisé pour l'adoption et la mise en œuvre rapides de **flux de travail du SNG de l'ADN à l'obtention des résultats d'analyse** compatibles avec l'automatisation pour les applications de recherche et de recherche clinique. IGA est le fruit de l'expérience acquise en aidant les clients à mettre en œuvre des flux de travail « de l'ADN aux données » pour le séquençage de l'exome entier (WES, Whole-Exome Sequencing), le séquençage du génome entier (WGS, Whole-Genome Sequencing) et les programmes de génomique des populations (PopGen) (figure 1).

Le **projet SG100K** est l'un des programmes tirant parti d'IGA. SG100K est un travail collaboratif entre Illumina et Precision Health Research Singapore (PRECISE) pour séquençer les génomes de 100 000 individus en bonne santé à Singapour afin de mieux comprendre la diversité génomique asiatique. Cette note d'application présente les données internes du WGS générées par des opérateurs indépendants dans différents sites au Royaume-Uni et à Singapour évaluant les scripts au sein d'IGA pour automatiser le flux de travail du SNG dans le cadre du projet SG100K.

## Méthodes

IGA comprend des scripts robotiques dédiés pour la préparation des bibliothèques et les intégrations logicielles afin d'automatiser et de simplifier le séquençage et l'analyse des données.

### Préparation de bibliothèques

Les bibliothèques de séquençage ont été préparées à l'aide d'Illumina DNA PCR-Free Prep (Illumina, référence n° 20041794) à partir de 400 ng d'ADN génomique (ADNg) de haute qualité extraits d'échantillons de sang prélevés auprès d'individus en bonne santé faisant partie du projet SG100K. La préparation des bibliothèques a été automatisée à l'aide de la plateforme de manipulation des liquides Hamilton STAR.

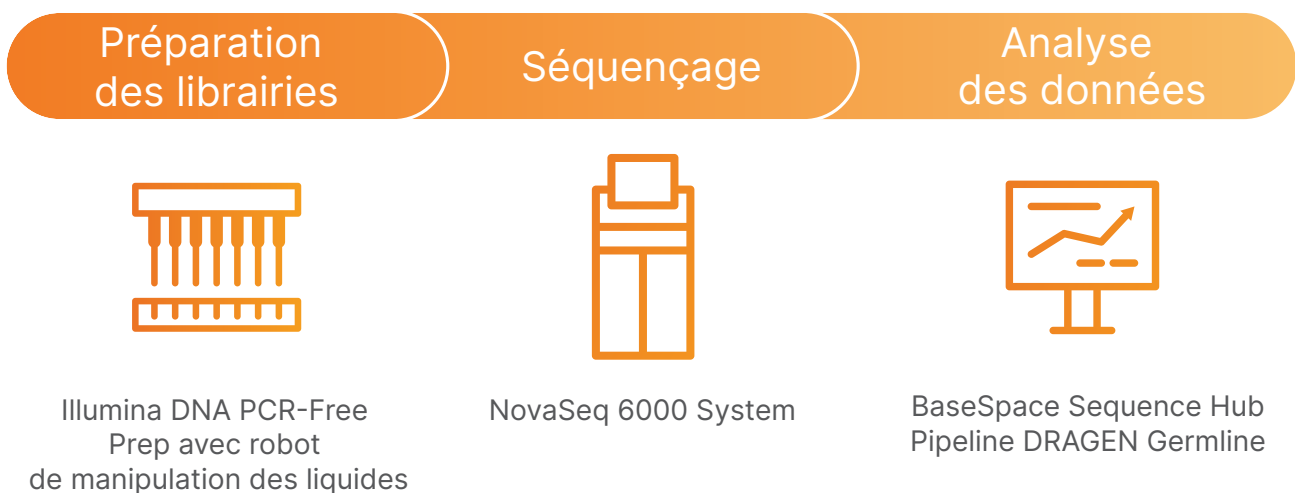


Figure 1 : Flux de travail de WGS d'Illumina Genomics Architecture – IGA prend en charge un flux de travail « de l'ADN aux données » pour le WGS qui intègre la préparation automatisée des bibliothèques avec Illumina DNA PCR-Free Prep, le séquençage sur le NovaSeq 6000 System et l'analyse avec le DRAGEN Germline pipeline.

## Séquençage

Les librairies préparées ont été séquençées sur le NovaSeq<sup>MC</sup> 6000 System (Illumina, référence n° 20012850) avec une configuration d'analyse de 2 × 151 pb. Vingt-quatre échantillons ont été analysés par Flow Cell S4 à une couverture de 30× (Illumina, référence n° 20028312). Dans le cadre du flux de travail automatisé avec IGA, le logiciel Clarity<sup>MC</sup> LIMS a envoyé l'instruction à la plateforme de manipulation des liquides d'effectuer un regroupement en masse, une dénaturation et un chargement des librairies sur le NovaSeq 6000 System et a envoyé les renseignements nécessaires pour démarrer automatiquement l'analyse de séquençage.

## Analyse des données

Une fois l'analyse de séquençage terminée, les données ont été transmises automatiquement dans BaseSpace<sup>MC</sup> Sequence Hub pour analyse avec DRAGEN<sup>MC</sup> Germline pipeline v3.7.8. Le logiciel JMP a été utilisé pour l'analyse statistique et le traçage de graphiques.

## Résultats

Pour évaluer les scripts au sein d'IGA afin d'automatiser et de simplifier un flux de travail de WGS avec Illumina DNA PCR-Free Prep, l'analyse de la variance (ANOVA, Analysis of Variance) a été utilisée pour comparer les données de séquençage obtenues par des opérateurs indépendants sur différents sites à Singapour et au Royaume-Uni. Les résultats ont démontré des performances fiables et une variabilité minimale entre les opérateurs indépendants qui mettent en œuvre des scripts de flux de travail IGA en utilisant des instruments distincts sur différents sites (figure 2 et tableau 1).

Tableau 1: Résumé de la cohérence des données avec IGA

Paramètre	Valeur moyenne ± écart-type
<b>Indicateur clé de performance (191 échantillons<sup>a</sup>)</b>	
Couverture autosomique moyenne	36,39 ± 5,9 %
Pourcentage de génomes avec couverture (≥ 15×)	94,72 ± 7,8 %
Bases Q30	102,1 ± 16,7 Gpb
Pourcentage de l'appelabilité autosomique	97,05 ± 7,1 %
Longueur médiane de l'insert	459,4 ± 20,2 pb
Estimation de la contamination des échantillons	0,001 ± 0,0002
<b>Analyse des SNV<sup>b</sup> (99 échantillons)</b>	
Rappel des SNV	99,84 ± 0,11 %
Précision des SNV	99,83 ± 0,03 %
Rappel des indels	99,58 ± 0,31 %
Précision des indels	99,66 ± 0,15 %

a. Extraction d'ADN de Coriell.  
b. Basée sur les données d'un seul échantillon représentatif de la réalité, NA12878.

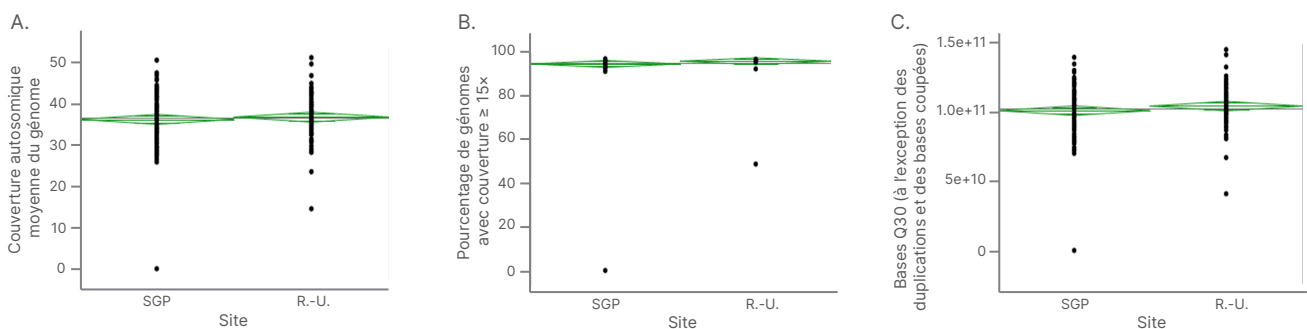


Figure 2 : Performance et cohérence des données avec des opérateurs indépendants sur différents sites – L'ANOVA des données de séquençage à Singapour (SGP) et au Royaume-Uni (R.-U.) a montré une performance hautement uniforme telle que mesurée par (A) la couverture autosomique moyenne, (B) le pourcentage de génomes avec une couverture ≥ 15× et (C) les bases Q30 (à l'exclusion des bases coupées et des duplications).

## Résumé

IGA offre un cadre flexible, modulaire et normalisé pour l'adoption rapide des flux de travail du SNG de l'ADN à l'obtention des résultats d'analyse compatibles avec l'automatisation. La mise en œuvre d'IGA avec Illumina DNA PCR-Free Prep pour le WGS permet la préparation de bibliothèques le jour même, le regroupement de 24 échantillons par Flow Cell S4 et le séquençage pour obtenir une couverture génomique de 30x. Dans le cadre du programme PopGen SG100K, IGA permet l'obtention de performances hautement uniformes et fiables entre les opérateurs indépendants sur différents sites. Ces résultats démontrent la capacité d'IGA à automatiser les flux de travail et fournir des performances fiables pour les méthodes de SNG.

## En savoir plus

[Illumina Genomics Architecture](#)

[Illumina DNA PCR-Free Prep](#)

[NovaSeq 6000 System](#)

[Analyse secondaire DRAGEN](#)



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

© 2023 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](http://www.illumina.com/company/legal.html).

M-GL-01991 FRA v1.0