

## Notice

DESTINÉ AU DIAGNOSTIC IN VITRO UNIQUEMENT. POUR L'EXPORTATION UNIQUEMENT.

**Catalogue n° 20005715**

## Utilisation prévue

Le Instrument NextSeq 550Dx est conçu pour le séquençage de bibliothèques d'ADN lorsqu'il est utilisé dans le cadre de tests de diagnostic *in vitro*. Le Instrument NextSeq 550Dx doit être utilisé avec des réactifs de diagnostic *in vitro* et un logiciel d'analytique spécifiques, enregistrés, certifiés ou approuvés.

## Principes de procédures

Le Illumina Instrument NextSeq 550Dx est destiné au séquençage des bibliothèques d'ADN avec des tests de diagnostic *in vitro* et est destiné à être utilisé par du personnel de laboratoire clinique qualifié et formé à l'utilisation des procédures de diagnostic *in vitro* effectuées dans un laboratoire clinique. Au départ, le NextSeq 550Dx utilise des bibliothèques préparées à partir de l'ADN où les index d'échantillons et les séquences de capture sont ajoutés aux cibles amplifiées. Les bibliothèques d'échantillons sont emprisonnées sur une Flow Cell et séquencées avec l'instrument en utilisant la chimie de séquençage par synthèse (SBS). La chimie SBS utilise une méthode basée sur des terminateurs réversibles pour détecter les bases à simple nucléotide à marqueur fluorescent, à mesure qu'elles sont intégrées aux brins d'ADN croissants. Le logiciel Real-Time Analysis (RTA) effectue des analyses d'images et des définitions de bases et affecte un score de qualité à chaque base pour chaque cycle de séquençage. Une fois l'analyse primaire terminée, une analyse secondaire peut être faite sur l'instrument pour le traitement des définition des bases. Le NextSeq 550Dx utilise différents modules d'analyse secondaire en fonction du flux de travail. Dans le cas des modules Germline Variant ou Somatic Variant, le traitement comprend le démultiplexage, la génération de fichiers FASTQ, l'alignement, la définition des variants et la génération de fichiers de définition de variant (VCF et gVCF). Les fichiers VCF et gVCF contiennent des renseignements sur les variants que l'on trouve à des positions spécifiques dans un génome de référence.

## Configuration à double amorçage

Le NextSeq 550Dx comprend une configuration à double amorçage pour permettre l'utilisation de l'instrument en mode diagnostic (Dx) ou en mode recherche uniquement (RUO). Les tests de séquençage de diagnostic *in vitro*, y compris les modules de Germline Variant et Somatic Variant, sont exécutés en mode diagnostic. Seuls les réactifs de séquençage IVD peuvent être utilisés en mode diagnostic. Les caractéristiques de performance et les limites de la procédure pour l'instrument NextSeq 550Dx ont été établies à l'aide des modules de Germline Variant et Somatic Variant en mode diagnostic.

# Limites de la procédure

1. Destiné au diagnostic *in vitro* uniquement.
2. Les modules Germline Variant et Somatic Variant, lorsqu'ils sont utilisés avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) ou le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), sont capables de fournir :
  - Sortie de séquençage  $\geq 90$  gigabases (Gb)
  - Longueur de lecture (en analyse à lecture appariée) de 2 x 150 paires de base (pb)
  - Bases égales ou supérieures à Q30  $\geq 75$  % à la longueur de lecture de 2 x 150 pb  
Au moins 75 % des bases ont un score de qualité sur l'échelle Phred  $\geq 30$ , indiquant que la précision des définitions des bases est supérieure à 99,9 %.
3. Les lectures avec indels (insertions, délétions ou combinaisons) dont la longueur du contenu est  $> 25$  pb ne sont pas alignées par le logiciel de test. Par conséquent, le logiciel analytique ne détecte pas les indels d'une longueur  $> 25$  pb.
4. Il est possible que le logiciel de test n'aligne pas les lectures d'amplicon présentant du contenu de variant extrême, ce qui fait que la région est signalée comme de type sauvage. Un tel contenu extrême comprend :
  - Lectures contenant plus de trois indels
  - Lectures d'une longueur d'au moins 30 pb avec une teneur en variant à un seul nucléotide (SNV)  $> 4$  % de la longueur cible totale de l'amplicon (à l'exclusion des régions sondées)
  - Lectures d'une longueur  $< 30$  pb avec une teneur en SNV  $> 10$  % de la longueur totale de l'amplicon (y compris les régions sondées)
5. Les variants importants, y compris les variants à multiples nucléotides (MNV) et les indels importants, peuvent être désignés comme plusieurs variants distincts de plus petite taille dans le fichier VCF de sortie.
6. Les variants de délétion peuvent être filtrés ou manqués lorsqu'ils s'étendent sur deux amplicons en mosaïque, si la longueur de délétion est supérieure ou égale au chevauchement entre les amplicons en mosaïque.
7. Le système est incapable de détecter les indels si celles-ci sont directement adjacentes à un primer et si aucun amplicon ne se chevauche. Pour les zones où les amplicons se chevauchent, l'analyse ne peut pas détecter les délétions lorsque la zone de chevauchement est plus petite que la taille de la délétion à détecter. Par exemple, si la région de chevauchement entre deux amplicons adjacents est de deux bases, le test ne peut détecter aucune délétion incluant ces deux bases. Une délétion à base unique, sur l'une de ces bases peut être détectée.
8. Comme pour toute préparation de bibliothèque basée sur l'hybridation, les polymorphismes sous-jacents, les mutations, les insertions ou les délétions dans les régions de liaison aux oligonucléotides peuvent affecter les allèles sondés et les appels effectués lors du séquençage. Par exemple :
  - Un variant en phase avec un variant dans la zone de primer peut ne pas être amplifié et peut résulter en un faux négatif.

- Les variants dans la zone de primer peuvent empêcher l'amplification de l'allèle de référence et cela peut causer une définition de variant homozygote incorrecte.
  - Les variants de type indel dans la zone de primer peuvent causer une définition fautive positive à la fin de la lecture adjacente au primer.
9. Les indels peuvent être filtrées en raison d'une distorsion du brin si elles surviennent à proximité de la fin d'une lecture et sont ignorées lors de l'alignement.
10. Les petits MNV n'ont pas été validés et ne sont signalés que dans le module Somatic Variant.
11. Les délétions sont signalées dans le VCF à la coordonnée de la base précédente conformément au format VCF. Par conséquent, considérez les variants adjacents avant de signaler qu'une définition des bases individuelle est une référence homozygote.
12. Limites relatives aux variants germinaux :
- Le Instrument NextSeq 550Dx, qui utilise le Module Germline Variant du Local Run Manager pour le NextSeq 550Dx, est conçu pour fournir des résultats qualitatifs pour la définition de variants germinaux (par exemple, homozygote, hétérozygote, de type sauvage).
  - Lorsqu'il est utilisé avec le module Germline Variant, la couverture minimale par amplicon nécessaire pour un définition de variant précis est de 150x. En conséquence, 150 fragments d'ADN de support sont nécessaires, ce qui équivaut à 300 lectures appariées qui se chevauchent. Le nombre d'échantillons et le nombre total de bases ciblées affectent la couverture. Le contenu GC et d'autres taux génomiques peuvent affecter la couverture.
  - La variation du nombre de copies peut affecter l'identification d'un variant comme homozygote ou hétérozygote.
  - Les variants dans un certain contexte répétitif ne sont pas inclus dans les fichiers VCF. Le filtre de répétition RMxN est utilisé pour filtrer les variants si toute, ou une partie de, la séquence du variant est présente de façon répétée dans le génome de référence adjacent à la position du variant. Pour la définition de variants germinaux, au moins neuf répétitions dans la référence sont nécessaires pour qu'un variant soit filtré. Seules les répétitions d'une longueur allant jusqu'à 5 pb sont prises en compte (R5x9).
  - Une indel et un SNV à un seul locus peuvent entraîner le signalement d'un seul variant.
13. Limites relatives aux variants somatiques :
- Le Instrument NextSeq 550Dx, qui utilise le module Somatic Variant du Local Run Manager pour le NextSeq 550Dx, est conçu pour fournir des résultats qualitatifs pour la définition de variants somatiques (par exemple, la présence d'un variant somatique avec une fréquence de variant supérieure ou égale à 0,026 avec une limite de détection de 0,05).
  - Lorsqu'il est utilisé avec le module Somatic Variant, la couverture minimale par amplicon nécessaire pour un définition précis des variants est de 450x par pool d'oligonucléotides. En conséquence, 450 fragments d'ADN de support sont nécessaires par pool d'oligonucléotides, ce qui équivaut à

900 lectures appariées qui se chevauchent. Le nombre d'échantillons et le nombre total de bases ciblées affectent la couverture. Le contenu GC et d'autres taux génomiques peuvent affecter la couverture.

- Pour la définition de variants somatiques, au moins six répétitions dans la référence sont nécessaires pour que le variant soit filtré, et seules les répétitions d'une longueur de 3 pb maximum sont prises en compte (R3x6).
- Le module Somatic Variant est incapable de différencier les variants germinaux et somatiques. Le module est conçu pour détecter les variants à travers une plage de fréquences de variant, mais la fréquence de variant ne peut pas être utilisée pour différencier les variants somatiques et les variants germinaux.
- Les tissus normaux dans le spécimen affectent la détection des variants. La limite de détection signalée est basée sur la fréquence de variant par rapport à l'ADN total extrait du tissu tumoral et normal.

## Composants du produit

Le Illumina NextSeq 550Dx se compose des éléments suivants :

1. Instrument NextSeq 550Dx (Catalogue n° 20005715)
2. Les composants logiciels pour le Instrument NextSeq 550Dx, y compris les éléments suivants :

Application logicielle	Fonction	Description
NextSeq 550Dx Operating Software (NOS)	Contrôle le fonctionnement de l'instrument	L'application logicielle NOS prend en charge le fonctionnement de l'instrument lors du séquençage et génère des images à utiliser avec le logiciel Real-Time Analysis (RTA).
Logiciel Real-time Analysis (RTA)	Réalise des analyses primaires	L'application logicielle RTA convertit les images générées par NOS pour chaque plaque par cycle de séries de séquençage en fichiers de définition des bases, qui sont des entrées pour les modules d'analyse Local Run Manager. L'application logicielle RTA n'est pas dotée d'une interface utilisateur.
Local Run Manager	Interface pour la sélection du module	Le logiciel Local Run Manager est une solution intégrée à l'instrument pour la gestion des utilisateurs, la sélection du module d'analyse approprié et la vérification de l'état.
Module Somatic Variant	Réalise des analyses secondaires	Ce logiciel de module d'analyse Local Run Manager traite les définitions des bases par le biais d'une analyse secondaire. Le traitement comprend le démultiplexage, la génération de fichiers FASTQ, l'alignement, la définition des variantes et le reporting. Le paramètre de définition des variantes (Pisces) génère des fichiers VCF qui contiennent des informations sur les variantes trouvées à des positions spécifiques dans un génome de référence et inclut la fréquence de variante mesurée.

Application logicielle	Fonction	Description
Module Germline Variant	Réalise des analyses secondaires	Ce logiciel de module d'analyse Local Run Manager traite les définitions des bases par le biais d'une analyse secondaire. Le traitement comprend le démultiplexage, la génération de fichiers FASTQ, l'alignement, la définition des variantes et le reporting. Le paramètre de définition des variantes (Pisces) génère des fichiers VCF qui contiennent des informations sur les variantes trouvées à des positions spécifiques dans un génome de référence et identifie chaque variante comme hétérozygote ou homozygote.

3. **En option** Serveur Illumina DRAGEN pour NextSeq 550Dx (Catalogue n° 20086130), y compris le composant logiciel suivant :

Application logicielle	Fonction	Description
Illumina Run Manager	Interface pour la sélection du module d'application	Le logiciel Illumina Run Manager est installé sur le serveur DRAGEN hors instrument en option. Illumina Run Manager permet la gestion des utilisateurs, la sélection du module d'analyse et la vérification de l'état de l'analyse et de la série de séquençage.

Le serveur Illumina DRAGEN en option pour NextSeq 550Dx n'est disponible que dans certains pays. Contactez un représentant Illumina pour connaître la disponibilité régionale.

## Conditions de fonctionnement

Pour en savoir plus sur les conditions de fonctionnement, consultez la section Considérations environnementales du *Guide de préparation du site de l'instrument NextSeq 550Dx (document n° 1000000009869)*.

Élément	Spécification
Température	Maintenez la température du laboratoire entre 19 °C et 25 °C (22 °C ± 3 °C). Cette température est la température de fonctionnement de l'instrument. Au cours d'une analyse, empêchez toute variation de la température ambiante de plus de ± 2 °C.
Humidité	Maintenez une humidité relative sans condensation comprise entre 20 et 80 %.

## Équipements et matériaux

### Équipement et matériel requis, vendus séparément

NextSeq 550Dx High Output Reagent Kit v2.5 (75 cycles), Catalogue n° 20028870

NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), Catalogue n° 20028871

## Équipement et matériaux nécessaires, non fournis

### Consommables fournis par l'utilisateur pour le séquençage des analyses

Consommable	Fournisseur	Utilisation
Lingettes imbibées d'alcool, alcool isopropylique à 70 % ou Éthanol à 70 %	VWR, catalogue n° 95041-714 (ou équivalent) Fournisseur de laboratoire général	Nettoyage de la Flow Cell et usage général
Serviette de laboratoire, non-pelucheuse	VWR, catalogue n° 21905-026 (ou équivalent)	Nettoyage de la Flow Cell et usage général

### Consommables fournis par l'utilisateur pour la maintenance des instruments

Consommable	Fournisseur	Utilisation
NaOCl, 5 % (hypochlorite de sodium)	Sigma-Aldrich, n° de référence 239305 (ou équivalent destiné à un usage en laboratoire)	Lavage de l'instrument au moyen du lavage à la main après séquençage ; dilué à 0,12 %
Tween 20	Sigma-Aldrich, n° de référence P7949	Lavage de l'instrument au moyen des options de lavage à la main ; dilué à 0,05 %
Eau, usage en laboratoire	Fournisseur de laboratoire général	Lavage de l'instrument (à la main)
Filtre à air	Illumina, catalogue n° 20063988	Purification de l'air que l'instrument prend pour le refroidissement

### Directives relatives à l'eau destinée à un usage en laboratoire

Toujours utiliser de l'eau destinée à un usage en laboratoire ou de l'eau désionisée pour effectuer les procédures relatives à l'instrument. Ne jamais utiliser l'eau du robinet. Utiliser uniquement les catégories d'eau suivantes ou leurs équivalents :

- Eau désionisée
- Illumina PW1
- Eau 18 mégohms (M $\Omega$ )
- Eau Milli-Q
- Eau Super-Q
- Eau destinée à un usage en biologie moléculaire

# Avertissements et Précautions



## ATTENTION

La loi fédérale réserve la vente de cet appareil par ou sur l'ordre d'un médecin ou d'un autre praticien autorisé par la loi de l'État dans lequel il exerce, à utiliser ou à ordonner l'utilisation de l'appareil.

- Certains composants des réactifs fournis par Illumina pour être utilisés avec le Instrument NextSeq 550Dx contiennent des produits chimiques potentiellement dangereux. Des dommages corporels peuvent survenir en cas d'inhalation, d'ingestion, de contact avec la peau ou les yeux. Porter un équipement de protection, y compris des lunettes de protection, des gants et une blouse de laboratoire adaptés au risque d'exposition. Manipuler les réactifs usagés comme des déchets chimiques et les mettre au rebut conformément aux lois et règles régionales, nationales et locales en vigueur. Pour de plus amples renseignements relatifs à l'environnement, à la santé et à la sécurité, consulter la Fiche de données de sécurité (FDS) à l'adresse [support.illumina.com/sds.html](https://support.illumina.com/sds.html).**
- Signaler immédiatement tout incident grave relatif à ce produit à Illumina et aux Autorités compétentes des États membres dans lequel l'utilisateur et le patient sont établis.
- Manipuler tous les spécimens de sang comme s'ils étaient contaminés par le virus de l'immunodéficience humaine (VIH), le virus de l'hépatite B (VHB) ou et d'autres agents pathogènes transmissibles par le sang (précautions universelles).
- Le défaut de suivre les procédures, telles que spécifiées, peut causer des résultats erronés ou une diminution significative de la qualité de l'échantillon.
- Prendre les précautions de routine en laboratoire. Ne pas injecter avec la bouche. Ne pas manger, boire ou fumer dans les zones de travail désignées. Porter des gants à usage unique et une blouse de laboratoire lors de la manipulation des spécimens et des réactifs du kit. Se laver correctement les mains après avoir manipulé les spécimens et les réactifs du kit.
- Des pratiques de laboratoire adéquates et une bonne hygiène de laboratoire sont nécessaires pour empêcher les produits PCR de contaminer les réactifs, l'instrument et les échantillons d'ADN génomiques. La contamination par PCR peut causer des résultats inexacts et peu fiables.
- Pour empêcher la contamination, s'assurer que les zones de pré- et post-amplification sont équipées de l'équipement et des consommables nécessaires (p. ex., les pipettes, les blocs chauffant, les vortex et les centrifugeuses).
- Le jumelage de l'index et de l'échantillon doit correspondre exactement à la disposition de plaque imprimée. Local Run Manager remplit automatiquement les primers d'indexation associés aux noms d'échantillons, lorsqu'ils sont saisis dans le module. Il est recommandé de vérifier les primers d'indexation associés aux échantillons avant de commencer le séquençage. Le mauvais placement de l'échantillon par rapport à la disposition du plateau cause une diminution d'identification des échantillons positifs et un signalement incorrect des résultats.

9. L'installation d'un logiciel antivirus fourni par l'utilisateur est fortement recommandée pour protéger l'ordinateur des virus. Consulter le manuel de l'utilisateur pour obtenir des instructions relatives à l'installation.
10. Ne pas utiliser le NextSeq 550Dx si l'un des panneaux est retiré. L'utilisation de l'instrument avec l'un des panneaux retiré crée une exposition potentielle aux tensions secteur et CC.
11. Ne pas toucher la platine de Flow Cell dans le compartiment de la Flow Cell. Le chauffage de ce compartiment fonctionne à des températures comprises en 22 °C et 95 °C et peut causer des brûlures.
12. L'instrument pèse environ 83,9 kg (185 lbs) et peut causer de graves blessures s'il tombe et n'est pas manipulé correctement.

## Mode d'emploi

Les instructions suivantes concernant l'utilisation du Instrument NextSeq 550Dx nécessitent les réactifs fournis dans le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) ou le NextSeq 550Dx High Output Reagent Kit v2.5 (75 cycles).

## Créer une série

Créez une série de séquençage à l'aide de Local Run Manager ou Illumina Run Manager. Les instructions d'utilisation de Local Run Manager sont incluses ci-dessous et dans le Guide de référence de l'instrument NextSeq 550Dx (document n° 1000000009513). Pour obtenir des instructions sur la création d'une série à l'aide de Illumina Run Manager, reportez-vous au Guide du logiciel Illumina Run Manager pour NextSeq 550Dx (document n° 200025239).

Pour des instructions sur la sélection entre Local Run Manager ou Illumina Run Manager, reportez-vous au Guide du logiciel Illumina Run Manager pour NextSeq 550Dx (document n° 200025239). Pour des instructions détaillées sur des applications spécifiques, se reporter au module ou au guide d'application du test en question.

Les instructions suivantes concernent l'utilisation des modules de Local Run Manager Germline Variant et Somatic Variant.

## Définir les paramètres

1. Connectez-vous à Local Run Manager.
2. Sélectionnez **Create Run** (Créer une série), et sélectionnez **Somatic Variant** (Variante somatique) ou **Germline Variant** (Variante germinale).
3. Saisissez un nom de la série qui identifie la série allant du séquençage jusqu'à l'analyse. Utilisez des caractères alphanumériques, des espaces, des traits de soulignement ou des tirets.
4. [Facultatif] Saisissez une description de série pour identifier davantage la série. Utilisez des caractères alphanumériques, des espaces, des traits de soulignement ou des tirets.
5. Sélectionnez le nombre d'échantillons et l'index défini dans la liste déroulante. Tenez compte des informations suivantes lorsque vous faites une sélection.

- La liste déroulante contient le nombre d'échantillons avec un ensemble d'index. Par exemple, 24-Set 1 indique 24 échantillons à tester, avec des index provenant de l'index 1.
- Les numéros d'index se réfèrent à différents ensembles de paires d'index i5 et i7. L'ensemble 1 et l'ensemble 2 fournissent tous deux une diversité d'index. Deux ensembles d'index sont proposés pour éviter l'épuisement d'un seul.
- Choisissez le nombre d'échantillons le plus proche du nombre d'échantillons que vous analysez. Si le nombre exact d'échantillons ne figure pas dans la liste, sélectionnez le nombre le plus proche, mais inférieur au nombre que vous testez. Par exemple, si vous souhaitez tester 18 échantillons, sélectionnez 16 échantillons.
- Les puits d'échantillon suggérés et les combinaisons d'indices qui répondent aux exigences de diversité d'indices sont surlignés en vert.

## Importer les fichiers de manifeste pour l'analyse

1. Assurez-vous que les manifestes que vous souhaitez importer sont disponibles dans un emplacement réseau accessible ou sur une clé USB.
2. Sélectionnez **Import Manifests** (Importer les manifestes).
3. Accédez au fichier de manifeste et sélectionnez les manifestes que vous souhaitez ajouter.

**REMARQUE** Pour rendre les fichiers manifestes disponibles pour toutes les analyses à l'aide du module d'analyse de Germline Variant ou Somatic Variant, ajoutez des manifestes à l'aide de la fonctionnalité Module Settings (Paramètres du module). Cette fonctionnalité nécessite des autorisations de niveau utilisateur administrateur. Pour en savoir plus, consultez *Guide de référence de l'instrument NextSeq 550Dx (document n° 100000009513)*.

## Spécifier les échantillons pour la série

Spécifiez les échantillons pour la série à l'aide de l'une des options et les instructions qui suivent.

**Saisir les échantillons manuellement**—Utilisez le tableau vide sur l'écran Create Run (Créer une série).

**Importer des échantillons**—Accédez à un fichier externe au format de valeurs séparées par des virgules (\*.csv). Un modèle peut être téléchargé sur l'écran Create Run (Créer une série).

### Saisir les échantillons manuellement

1. Saisissez un nom d'échantillon unique (*module d'analyse Somatic Variant*) ou un identifiant d'échantillon (*module d'analyse Germline Variant*).  
Utilisez des caractères alphanumériques, des tirets ou des traits de soulignement.
2. [Facultatif] Pour les échantillons de contrôle positif ou négatif, cliquez avec le bouton droit et sélectionnez le type de contrôle.

Le contrôle dans un puits d'échantillon remplit automatiquement le puits correspondant dans l'autre pool avec le même contrôle.

3. [Facultatif] Saisissez une description d'échantillon dans le champ Sample Description (Description d'échantillon).  
Utilisez des caractères alphanumériques, des tirets ou des traits de soulignement.
4. Sélectionnez un adaptateur Index 1 dans la liste déroulante Index 1 (i7).  
Lorsque vous utilisez des puits d'échantillon suggérés, le logiciel remplit automatiquement les adaptateurs d'index i7 et i5 qui répondent aux exigences d'index de diversité. Si le nombre exact d'échantillons que vous testez ne figure pas dans la liste, assurez-vous de sélectionner des adaptateurs d'index pour les puits supplémentaires.
5. Sélectionnez un adaptateur Index 2 dans la liste déroulante Index 2 (i5).
6. Sélectionnez un fichier manifeste dans la liste déroulante Manifeste.  
Les échantillons du pool A nécessitent un manifeste différent des échantillons du pool B.
7. Choisissez une option pour afficher, imprimer ou enregistrer la disposition de la plaque comme référence pour la préparation des bibliothèques :
  - Sélectionnez l'icône  **Print** (Imprimer) pour afficher la disposition de la plaque. Sélectionnez **Print** (Imprimer) pour imprimer la disposition de la plaque.
  - Sélectionnez **Export** (Exporter) pour exporter les informations d'échantillon vers un fichier externe.
8. Sélectionnez **Save Run** (Sauvegarder la série).

## Importer des échantillons

1. Sélectionnez **Import Samples** (Importer des échantillons) et accédez à l'emplacement du fichier d'informations sur l'échantillon. Il existe deux types de fichiers que vous pouvez importer.
  - Sélectionnez **Template** (Modèle) sur l'écran Create Run (Créer une série) pour créer une nouvelle disposition de plaque. Le fichier du modèle contient les en-têtes de colonne corrects pour l'importation. Saisissez les informations sur les échantillons dans chaque colonne pour les échantillons de la série. Supprimez les informations d'exemple dans les cellules inutilisées, puis enregistrez le fichier.
  - Utilisez un fichier d'informations sur l'échantillon exporté à partir du module Germline Variant (Variante germinale) ou Somatic Variant (Variante somatique) à l'aide de la fonction Export (Exporter).
2. Sélectionnez l'icône  **Print** (Imprimer) pour afficher la disposition de la plaque.
3. Sélectionnez **Print** (Imprimer) pour imprimer la disposition de la plaque comme référence pour la préparation des bibliothèques.
4. Sélectionnez **Save Run** (Sauvegarder la série).

## Préparer la cartouche de réactifs

Assurez-vous de suivre attentivement les instructions de la cartouche de réactif pour un séquençage réussi.

1. Retirez la cartouche de réactif de son stockage à -25 °C ou -15 °C.

2. Choisissez l'une des méthodes suivantes pour décongeler les réactifs. Ne pas immerger la cartouche. Une fois la cartouche décongelée, séchez-la avant de passer à l'étape suivante.

Température	Temps de décongélation	Limite de stabilité
Bain-marie de 15 °C à 30 °C	60 minutes	Ne pas dépasser 6 heures
2 °C à 8 °C	7 heures	Ne pas dépasser 5 jours

**REMARQUE** Si plus d'une cartouche est décongelée dans le même bain-marie, prévoyez un temps de décongélation supplémentaire.

3. Retournez la cartouche cinq fois pour mélanger les réactifs.
4. Inspectez le fond de la cartouche pour vous assurer que les réactifs sont décongelés et exempts de précipités. Confirmez que les positions 29, 30, 31 et 32 sont décongelées, car elles sont les plus grandes et prennent le plus de temps à décongeler.
5. Tapotez doucement sur la paillasse pour éliminer les bulles d'air.  
Pour de meilleurs résultats, chargez directement les échantillons et lancez le séquençage.

## Préparer la Flow Cell

1. Sortez une nouvelle boîte de Flow Cell du lieu de stockage à une température maintenue entre 2 °C et 8 °C.
2. Retirez l'emballage en aluminium de la boîte et mettez-la de côté à température ambiante pendant 30 minutes.

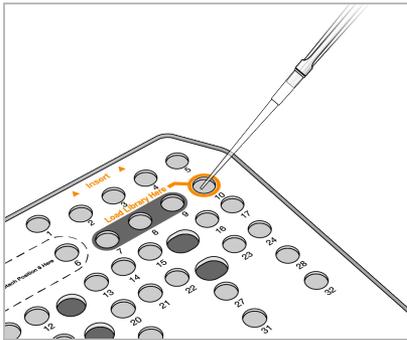
## Préparation des bibliothèques pour le séquençage

Dénaturez et diluez vos bibliothèques à un volume de chargement de 1,3 ml. Dans la pratique, la concentration de charge peut varier en fonction des méthodes de préparation et de quantification de la bibliothèque. La dilution des bibliothèques d'échantillons dépend de la complexité des groupes d'oligonucléotides. Pour savoir comment préparer des bibliothèques d'échantillons pour le séquençage, y compris la dilution et le regroupement des bibliothèques, consultez la rubrique Mode d'emploi de la notice de la trousse de préparation de bibliothèques. L'optimisation de la densité de l'amplifiat sur le NextSeq 550Dx est requise.

## Chargez les bibliothèques sur la cartouche de réactif

1. Nettoyez à l'aide d'un tissu non pelucheux le joint en aluminium recouvrant le réservoir n° 10 marqué **Load Library Here** (Charger la bibliothèque ici).
2. Percez le joint avec un embout de pipette propre de 1 ml.
3. Chargez 1,3 ml de bibliothèques préparées dans le réservoir n° 10 marqué **Load Library Here** (Charger la bibliothèque ici). Évitez de toucher le joint en aluminium lorsque vous distribuez les bibliothèques.

Figure 1 Charger les bibliothèques



## Configurez une analyse de séquençage

Reportez-vous à la section Guide de référence de l'instrument NextSeq 550Dx (document n° 1000000009513) pour obtenir des instructions complètes sur la configuration du cycle.

1. Connectez-vous à NextSeq 550Dx avec votre mot de passe du logiciel Local Run Manager ou Illumina Run Manager.
2. À partir de l'écran d'accueil du logiciel NOS, sélectionnez **Sequence** (Séquence).
3. Sélectionnez un séquençage dans la liste, puis sélectionnez **Next** (Suivant).

Une série d'écrans de configuration de l'exécution s'ouvre dans l'ordre suivant : Load Flow Cell (Chargement de la cellule de débit), Load Buffer Cartridge (Chargement de la cartouche de tampon), Load Reagent Cartridge (Chargement de la cartouche de réactif) et Pre-run Check (Contrôle préalable).

**REMARQUE** Les séries ne sont accessibles qu'en utilisant le même Run Manager que celui utilisé lors de la planification de séquençage. Pour des instructions sur la configuration du logiciel Run Manager, reportez-vous à la section Guide du logiciel Illumina Run Manager pour NextSeq 550Dx (document n° 200025239).

4. Lorsque l'écran Load Flow Cell (Charger la cellule de débit) apparaît, nettoyez et chargez la Flow Cell.
  - Sortez la Flow Cell de son emballage.
  - Ouvrez l'emballage en plastique transparent et retirez la Flow Cell
  - Nettoyez la surface en verre de la Flow Cell avec une lingette non pelucheuse imbibée d'alcool. Séchez le verre à l'aide d'un tissu non pelucheux
  - Assurez-vous que la surface en verre de la Flow Cell est propre. Si nécessaire, répétez l'étape de nettoyage.
  - Retirez la Flow Cell utilisée lors d'une analyse précédente.
  - Alignez la Flow Cell sur les broches d'alignement et placez-la sur la platine.
5. Sélectionnez **Load** (Charger).

La porte se ferme automatiquement, l'identifiant de la Flow Cell s'affiche à l'écran et les capteurs sont vérifiés.

6. Suivez les instructions du logiciel pour vider le récipient de réactifs usagé, charger la cartouche de tampon NextSeq 550Dx et charger la cartouche de réactif NextSeq 550Dx.  
Lorsque le tampon NextSeq 550Dx et les cartouches de réactifs sont chargés, le logiciel lit et enregistre la RFID. Les identifiants de tampon et de cartouche de réactif apparaissent à l'écran et les capteurs sont vérifiés.
7. Une fois la vérification préalable automatisée est terminée, sélectionnez **Start** (Démarrer). (Non requise si configurée pour démarrer automatiquement.)
8. L'écran Sequencing (Séquençage) s'ouvre au début du séquençage. Cet écran fournit une représentation visuelle du séquençage en cours d'exécution, y compris les intensités et les scores de qualité (Q-scores).

## Résultats

Real-Time Analysis (RTA) est un logiciel intégré qui effectue des analyses d'images et des définitions de bases et affecte un score de qualité à chaque base pour chaque cycle de séquençage. Lorsque l'analyse principale est terminée, le module d'application sélectionné commence automatiquement l'analyse secondaire. Les processus d'analyse secondaire décrits ici concernent les modules Local Run Manager Germline Variant et Somatic Variant sur le Instrument NextSeq 550Dx.

## Démultiplexage

Le démultiplexage compare chaque séquence de lecture d'index aux séquences d'indexage définies pour la série. Aucune valeur de qualité n'est prise en compte lors de cette étape.

Les lectures d'index sont identifiées en suivant les étapes ci-dessous :

- Les échantillons sont numérotés en commençant par 1, selon l'ordre dans lequel ils sont classés pour la série.
- Le numéro d'échantillon 0 est réservé aux amplifiats qui n'ont pas été assignés à un échantillon.
- Les amplifiats sont assignés à un échantillon lorsque la séquence d'indexage est identique ou lorsqu'il y a une seule non-correspondance par lecture d'index.

## Génération de fichier FASTQ

Après le démultiplexage, le logiciel génère des fichiers d'analyse intermédiaires au format FASTQ, qui est un format texte utilisé pour représenter les séquences. Les fichiers FASTQ contiennent des lectures pour chaque échantillon et les scores de qualité associés. Les amplifiats qui n'ont pas passé le filtre sont exclus.

Chaque fichier FASTQ contient des lectures pour un seul échantillon, et le nom de cet échantillon est inclus dans le nom du fichier FASTQ. Dans les modules de Germline Variant et Somatic Variant, huit fichiers FASTQ sont générés par échantillon et par pool d'oligonucléotides, soit quatre dans la lecture 1 et quatre dans la lecture 2. Ce résultat donne un total de 8 et 16 fichiers FASTQ par échantillon pour Germline et Somatic, respectivement. Les fichiers FASTQ constituent les principales données d'entrée pour l'alignement.

## Alignement

Pendant l'étape d'alignement, l'algorithme Smith-Waterman à bandes aligne les clusters de chaque échantillon avec les séquences d'amplicons spécifiées dans le fichier manifeste.

L'algorithme de Smith-Waterman à bandes effectue des alignements semi-globaux de séquences pour déterminer des régions similaires entre deux séquences. Au lieu de comparer la séquence totale, l'algorithme Smith-Waterman compare des segments de toutes les longueurs possibles.

Chaque lecture d'extrémité appariée est évaluée en termes d'alignement avec les séquences de sonde pertinentes pour cette lecture.

- La lecture 1 est évaluée par rapport au complément inverse des oligos spécifiques au locus en aval (DLSO).
- La lecture 2 est évaluée par rapport aux oligos spécifiques au locus en amont (ULSO).
- Si le début d'une lecture correspond à une séquence de sonde sans plus d'une non-concordance, la longueur totale de la lecture est alignée par rapport à la cible d'amplicon pour cette séquence.
- Si le début d'une lecture correspond à une séquence de sonde sans plus de trois différences (non-correspondances ou décalages dus à des indels de tête), la longueur totale de la lecture est alignée sur la cible d'amplicon pour cette séquence.
- Les indels dans le DLSO et l'ULSO ne sont pas observés compte tenu de la chimie du test.

Les alignements sont filtrés à partir des résultats d'alignement en fonction des taux de discordance sur la région d'intérêt ou l'amplicon complet, en fonction de la longueur de l'amplicon. Les alignements filtrés sont écrits dans les fichiers d'alignement comme non alignés et ne sont pas utilisés dans la recherche de variante.

## Définition de variant

La définition des variants Pisces est conçue pour effectuer des appels de variants SNV et indel à partir de bibliothèques préparées pour l'instrument.

## Rapports et fichiers de sortie supplémentaires

Les modules d'analyse des variants produisent des rapports PDF et délimités par des tabulations (\*.txt) qui affichent des mesures telles que la profondeur de séquençage et le nombre de variants. Les modules produisent également des fichiers de sortie tels que les fichiers VCF et gVCF (genome Variant Call Format) pour les applications de définition de variants.

# Procédures de contrôle de la qualité

Le logiciel NextSeq 550Dx évalue chaque séquençage, chaque échantillon et chaque définition de bases par rapport à des métriques de contrôle de la qualité. Les contrôles positifs et négatifs sont aussi recommandés dans la préparation des bibliothèques et doivent faire l'objet d'une évaluation. Évaluez les contrôles comme suit :

- **Contrôle négatif (sans modèle) ou autre contrôle négatif**— Doit générer le résultat attendu. Si le contrôle négatif génère un résultat différent de celui attendu, il peut s'agir d'une erreur dans le suivi de l'échantillon, un enregistrement incorrect des primers d'indexation ou une contamination.
- **Échantillon de contrôle positif**— Doit générer le résultat attendu. Si le contrôle positif génère un résultat différent de celui attendu, alors il peut s'agir d'une erreur dans le suivi de l'échantillon ou un enregistrement incorrect des primers d'indexation.

## Caractéristiques de performances

Les caractéristiques de performance pour le Instrument NextSeq 550Dx ont été établies à l'aide des modules Germline Variant et Somatic Variant avec le TruSeq Custom Amplicon Kit Dx et le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) et confirmées à l'aide du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Les études comprenaient l'indexage des échantillons, la contamination par transfert entre échantillons, l'entrée d'ADN, la sensibilité analytique (limite de blanc et limite de détection), l'exactitude, la précision, la comparaison des méthodes et la reproductibilité.

Les études analytiques utilisant le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) ont été conçues pour évaluer les revendications de performance précédemment établies avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Les résultats démontrent que les kits de réactifs (v2 et v2.5) ont des performances comparables à l'aide du TruSeq Custom Amplicon Kit Dx. Consultez la *notice du TruSeq Custom Amplicon Kit Dx* afin de savoir les caractéristiques de performance liées aux facteurs préanalytiques, comme les méthodes d'extraction ou les substances interférentes.

## Définitions des calculs utilisés dans les caractéristiques de performances

1. Le pourcentage de concordance positive (PCP) est calculé comme le pourcentage de loci classés comme variants par une méthode de référence que l'analyse signale correctement.
  - $(\text{nombre de loci avec variant signalés par l'analyse}) / (\text{nombre total de loci avec variant})$Les loci avec variant signalés par le séquençage qui correspondent à la méthode de référence sont des vrais positifs (VP). Les loci avec variant signalés comme définitions de référence ou définitions de variants différents sont des faux négatifs (FN)

2. Le pourcentage de concordance négative (PCN) est calculé comme le pourcentage de loci classés comme sauvages par une méthode de référence que l'analyse le signale correctement.
  - $(\text{nombre de loci sauvages signalés par l'analyse}) / (\text{nombre total de loci sauvages})$   
Les loci sauvages signalés par l'analyse qui correspondent à la méthode de référence sont des vrais négatifs (VN). Les loci sauvages signalés comme variants par l'analyse sont des faux positifs (FP).
3. Le pourcentage de concordance globale (PCG) est calculé comme le pourcentage de loci correctement signalés par l'analyse par rapport à une méthode de référence.
  - $((\text{nombre de loci avec variant signalés par l'analyse}) + (\text{nombre de loci sauvages correctement signalés par l'analyse})) / ((\text{nombre total de loci avec variant}) + (\text{nombre total de loci sauvages}))$
4. Les calculs du PCP, PCN et PCG ne comprennent pas les absence de définitions (loci avec variant ou de référence ne répondant pas à un ou plusieurs filtres de qualité).
5. Le taux de définitions autosomiques est calculé comme le nombre total de loci passant les filtres divisé par le nombre total de positions séquencées pour les chromosomes 1 à 22 ; les chromosomes x et Y étant exclus. Cette métrique ne tient pas compte de la concordance des définitions avec la méthode de référence.

## NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) Performance

### Indexation des échantillons

Les primers d'index de l'échantillon, ajoutés pendant la préparation des bibliothèques, attribuent une séquence unique à chaque échantillon d'ADN. Ces séquences uniques permettent le regroupement de multiples échantillons dans une seule analyse de séquençage. L'indexation des échantillons a été testée pour les flux de travail germinaux et somatiques. L'objectif de cette étude était d'établir le nombre d'échantillons minimal (8) et maximal (96) pouvant être traités lors d'une seule série de séquençage par le Instrument NextSeq 550Dx. Huit échantillons Platinum Genome uniques ont été testés avec 12 combinaisons de primers d'indexation différentes par échantillon. Les résultats des échantillons générés de quatre séries de séquençage par le module Germline Variant ont été comparés au Platinum Genomes version 2016-1.0.

Pour la première série d'analyses, 96 bibliothèques d'échantillons indexés de manière unique ont été testées avec un test représentatif conçu pour interroger une variété de gènes couvrant 12 588 bases par brin sur l'ensemble des 23 chromosomes humains afin de vérifier la capacité du test à effectuer une définition de génotypage de manière cohérente pour un échantillon donné dans différentes combinaisons primer d'indexation. Pour la deuxième série d'analyses, huit bibliothèques d'échantillons à index unique ont été séquencées en deux séries de séquençage afin de vérifier le nombre minimal d'index pris en charge.

Dans le cas des séries de 96 index, la PCP pour les SNV allait de 98,7 % à 100 %, la PCP pour les insertions était de 100 % et la NPA était de 100 % pour chacune des 96 combinaisons d'index. Les analyses de 8 indices présentaient des valeurs PPA de 100 % (SNV, insertions et délétions) et une NPA de 100 % pour chacune des huit combinaisons d'indices.

## Contamination par transfert entre échantillons

L'instrument NextSeq 550Dx permet le séquençage de multiples échantillons et contrôles dans une seule analyse de séquençage. Une étude a été menée pour évaluer l'étendue de la contamination par transfert entre échantillons au cours d'une analyse de séquençage et entre les analyses de séquençage. Deux échantillons de Platinum Genome, un masculin et un féminin, ont été testés au moyen d'un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases (150 amplicons) sur 23 chromosomes différents, y compris les deux chromosomes sexuels. Les bibliothèques ont été séquencées sur l'instrument NextSeq 550Dx à l'aide du module Germline Variant. La contamination par transfert des échantillons masculins aux échantillons féminins a été observée par la présence de lectures de amplicon du chromosome Y dans les échantillons féminins.

La contamination par transfert entre échantillons réalisé au cours d'une analyse de séquençage peut survenir au cours de la génération d'amplifiats, de la définition des bases du cycle de l'index et du démultiplexage des échantillons. Pour tester le transfert d'échantillons au cours d'une analyse de séquençage, un pool de bibliothèques composé de 46 répliquats d'échantillons masculins et féminins plus quatre contrôles sans modèle a été séquencé une fois sur l'instrument NextSeq 550Dx. La contamination par transfert entre échantillons au cours d'une série de séquençage a été évaluée par la comparaison de la couverture d'amplicon du chromosome Y de chaque répliquat femelle avec la couverture moyenne d'amplicon du chromosome Y de tous les répliquats Masculins dans le regroupement. La médiane observée pendant le report était de 0,084 %.

Pour tester le transfert d'échantillons d'une analyse à l'autre, deux regroupements de bibliothèques ont été préparés et séquencés consécutivement sur un instrument NextSeq 550Dx. Le premier regroupement contenait 46 répliquats d'échantillon féminin plus deux contrôles sans modèle. Le deuxième regroupement contenait 46 répliquats d'échantillon masculin plus deux contrôles sans modèle. Les deux regroupements ont utilisé le même ensemble d'adaptateurs d'index. Le regroupement des échantillons femelles a été séquencé en premier, suivi de l'analyse de séquençage du regroupement des échantillons mâles, puis d'une autre analyse de séquençage du regroupement des échantillons femelles. La contamination par transfert entre échantillons a été évaluée en comparant la couverture d'amplicon du chromosome Y entre les répliquats correspondant de la deuxième analyse de séquençage du regroupement des échantillons femelles et de l'analyse de séquençage du regroupement des échantillons mâles. La médiane observée pendant le report était de 0,0076 %

## Entrée d'ADN

### Sang (variant germinal)

La plage d'entrée d'ADN sanguin pour la préparation de la bibliothèque TruSeq Custom Amplicon Kit Dx à l'aide du flux de travail du module Germline Variant a été établie pour le Instrument NextSeq 550Dx. Cette plage a été évaluée en effectuant une étude de dilution en série au moyen de 13 échantillons de Platinum Genome avec un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases sur les 23 chromosomes humains. La bibliothèque a été séquencée sur deux instruments NextSeq 550Dx à l'aide d'un lot de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Cinq échantillons ont été testés en double exemplaire à cinq niveaux d'entrée d'ADN allant de 250 ng à 12 ng (250 ng, 100 ng, 50 ng, 25 ng, et 12 ng). Huit échantillons ont été testés en tant que répliquat unique à chacun des cinq niveaux d'entrée d'ADN. Pour la détermination de la précision, les génotypes des échantillons ont été

comparés aux données de Platinum Genomes version 2016-1.0. Les résultats ont été déterminés pour chaque niveau d'entrée. La PCP pour chaque type de variant (NVS, insertions et délétions) est présentée dans le [Tableau 1](#) ; la PCN est présentée dans le [Tableau 2](#). Tous les niveaux d'entrée ont affiché une précision similaire. Le niveau d'entrée d'ADN recommandé pour le TruSeq Custom Amplicon Kit Dx est 50 ng, avec 25 ng, et 100 ng représentant les limites inférieures et supérieures permettant de respecter les caractéristiques de performance.

Tableau 1 Résultats de la PCP pour chaque niveau d'entrée d'ADN par type de variant

Entrée d'ADN (ng)	Type de variant	Variants prévus	TP	FN	Absence de définition de variant	PCP (%)
12	SNV	2 412	2 381	31	0	98,7
25			2 404	8	0	99,7
50			2 403	9	0	99,6
100			2 412	0	0	100
250			2 412	0	0	100
12	Insertion	808	784	3	21	99,6
25			781	5	22	99,4
50			786	2	20	99,8
100			786	0	22	100
250			786	0	22	100
12	Délétion	758	732	12	14	98,4
25			737	7	14	99,1
50			742	2	14	99,7
100			744	0	14	100
250			744	0	14	100

Tableau 2 PCN pour chaque niveau d'entrée d'ADN

Entrée d'ADN (ng)	TN	FP	Absence de définition dans la référence	PCN (%)
12	430 940	4	26	> 99,9
25	430 936	0	34	100
50	430 936	2	32	> 99,9
100	430 942	0	28	100
250	430 942	0	28	100

## FFPE (variant somatique)

La plage d'entrée d'ADN fixé au formol et inclus en paraffine (FFPE) pour la préparation de la bibliothèque TruSeq Custom Amplicon Kit Dx à l'aide du flux de travail du module Somatic Variant a été établie pour l'instrument NextSeq 550Dx. La plage d'entrée de l'ADN a été évaluée en effectuant une étude de dilution en série au moyen de trois échantillons Platinum Genome avec un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases sur les 23 chromosomes humains. Les lignées cellulaires GM12878 et GM12877

de Platinum Genome ont été fixées au formol et imprégnées à la paraffine avant d'en extraire l'ADN. GM12878 a été dilué avec GM12877 de façon à ce que les fréquences alléliques des variants (VAF) de 79 variants (55 SNV, 9 insertions et 15 délétions) soient proches de 0,025, 0,05 et 0,10. En outre, chaque échantillon présentait 91 variants avec des fréquences de variants plus élevées allant jusqu'à 1,0 VAF. Les échantillons ont été traités en double à cinq niveaux d'entrée d'ADN avec un cycle quantitatif delta moyen (dCq) de 2,1, 3,6, 4,6, 6,0 et 7,8, mesuré par le TruSeq Custom Amplicon Dx - FFPE QC Kit. Chaque bibliothèque a été séquencée sur deux instruments NextSeq 550Dx à l'aide de deux lots de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Pour la détermination de la précision, les définitions des variants des échantillons ont été comparées aux données de Platinum Genomes version 2016-1.0. La PCP pour chaque type de variant (NVS, insertions et délétions) est présentée dans le [Tableau 3](#) ; la PCN est présentée dans le [Tableau 4](#). L'entrée d'ADN recommandée pour les variants à une VAF de 0,05 ou plus est dCq ≤ 4 avec 4,6, ce qui fournit une limite inférieure pour répondre aux caractéristiques de performance.

Tableau 3 Résultats du PCP pour chaque niveau d'entrée d'ADN par type de variant

dCq Moyenne	Type de variant	Variants prévus	Absence de définition prévue	Dilution cible - VAF					
				0,025		0,05		0,10	
				Absence de définition de variant	PCP (%)	Absence de définition de variant	PCP (%)	Absence de définition de variant	PCP (%)
2,1	SNV	808	Non applicable	196	100	0	100	0	100
3,6				250	99,3	4	100	0	100
4,6				251	94,6	51	99,2	5	100
6,0				257	65,3	213	91,4	100	100
7,8				254	69,3	185	90,7	100	100
2,1	Insertion	264	8	66	96,5	8	100	8	100
3,6				62	97,0	8	100	8	100
4,6				48	96,3	21	100	8	100
6,0				40	80,4	47	98,2	24	95,8
7,8				57	87,0	56	96,2	31	100
2,1	Délétion	304	16	58	100	16	100	16	100
3,6				80	100	16	100	16	100
4,6				65	95,4	28	100	16	100
6,0				78	74,8	105	94,0	36	100
7,8				76	75,0	79	95,1	57	98,8

Tableau 4 PCN pour chaque niveau d'entrée d'ADN

dCq Moyenne	Type sauvage attendu	Dilution cible - VAF					
		0,025		0,05		0,10	
		Absence de définition dans la référence	PCN (%)	Absence de définition dans la référence	PCN (%)	Absence de définition dans la référence	PCN (%)
2,1	93 688	344	100	260	100	324	100
3,6		400	100	332	100	380	100
4,6		1 308	100	1 336	100	784	100
6,0		3900	> 99,9	3 296	> 99,9	2 996	100
7,8		3 020	> 99,9	2 880	> 99,9	2448	> 99,9

### Sensibilité analytique (Limite du blanc [LB] et Limite de détection [LD])

Cette étude a été menée pour évaluer la limite de blanc (LB) et la limite de détection (LD) pour le module Somatic Variant sur l'instrument NextSeq 550Dx. Cette étude a utilisé un test représentatif conçu pour demander plusieurs gènes couvrant 12 588 bases à travers 23 chromosomes différents. Les lignées cellulaires GM12878 et GM12877 de Platinum Genome ont été fixées au formol et imprégnées à la paraffine avant d'en extraire l'ADN. Le GM12878 a été dilué avec le GM12877 de façon à ce que les fréquences de variant de 74 variants (53 SNV, 7 insertions et 14 délétions) soient proches de  $0,05 \pm 0,02$ . Le GM12877 et le GM12878 dilué (GM12878-D) ont été testés sur six jours consécutifs de démarrage avec un seul instrument, en alternance entre deux lots du NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), pour un total de six séries de séquençage. Ce test a entraîné un résultat de 60 répliques de chaque variant dans GM12878-D et de 72 répliques de chaque coordonnée de type sauvage correspondante dans GM12877 pour chaque lot de réactifs. Les LB et LD ont été calculées grâce à l'approche classique prévue dans CLSI EP17-A2 à l'aide de l'option non-paramétrique. Les LB et LD ont été calculées pour les SNV, les insertions et les délétions séparément en regroupant les fréquences de variant pour un type de variant donné. L'erreur de Type I a été définie à 0,01 et l'erreur de Type II a été définie à 0,05.

Pour la LB, les fréquences de variant regroupées ont été triées de la plus faible à la plus élevée, et la 99 e position pour chaque lot de réactifs pour chaque type de variant a été calculée [Tableau 5](#). Le module Somatic Variant utilise un seuil (LB réelle) d'une VAF de 0,026 afin de déterminer la détection qualitative des variants. Les LB calculées ont confirmé que ce seuil entraîne une erreur de Type I inférieure à 0,01.

Tableau 5 Limite du blanc

Type de variant	Observations totales	LB Lot de réactifs 1 (%)	LB Lot de réactifs 2 (%)
SNV	3 816	0,77	0,77
Insertion	504	0,56	0,56
Délétion	1 008	1,20	1,20

Pour la LD, le pourcentage de fréquence de mutation individuelle pour chaque lot de réactifs pour chaque type de variant se trouvant en-dessous du seuil de 0,026 a été calculé [Tableau 6](#). Parce que les pourcentages étaient inférieurs à l'erreur de Type II de 5 % (0,05), la médiane des fréquences de variant a été calculée comme la LD [Tableau 6](#). La plus grande valeur des deux valeurs calculées a été choisie comme LD pour les deux lots de réactifs – 4,97 % pour les SNV, 5,12 % pour les insertions et 5,26 % pour les délétions.

Tableau 6 Limite de détection

Lot de réactifs	Type de variant	Observations totales	Nombre de mesures VAF < 2,6 %	% de mesures VAF < 2,6 %	Limite de détection (%)
1	SNV	3 180	53	1,7	4,94
	Insertion	420	6	1,4	5,08
	Délétion	840	7	0,8	5,22
2	SNV	3 180	51	1,6	4,97
	Insertion	420	5	1,2	5,12
	Délétion	840	7	0,80	5,26

## Précision

### Variant germinale

L'étude suivante a été menée pour évaluer la précision de la définition des variants du module Germline Variant sur le Instrument NextSeq 550Dx à l'aide du NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). 13 échantillons Platinum Genome uniques ont été testés à l'aide d'un test représentatif conçu pour interroger une variété de gènes couvrant 12 588 bases (150 amplicons) sur 23 chromosomes différents. Un total de neuf analyses ont été effectuées au moyen de trois instruments de séquençage, trois lots de réactifs et trois opérateurs au cours d'une période de cinq jours de démarrage. La précision a été déterminée pour les SNV, les insertions et les délétions en comparant les résultats à la méthode de référence composite bien caractérisée, Platinum Genomes version 2016-1.0. Sauf indication contraire, les zones génomiques sûres ont été définies en fonction de cette méthode de référence.

Tableau 7 Résumé de la concordance des variants germinaux

Critère	Observations totales <sup>1</sup>	Résultat en fonction de l'observation <sup>2</sup>	Résultat en fonction de l'analyse <sup>3</sup>
PCP pour les SNV	819	98,7	> 99,9
PCP pour les insertions	819	95,0	98,9
PCP pour les délétions	819	100	100
NPA	819	100	100
PCG	819	> 99,9	> 99,9

<sup>1</sup>Calculées en fonction du nombre d'échantillons par analyse (91) x nombre d'analyses (9) = 819.

<sup>2</sup>Valeur la plus faible observée par réplicat d'échantillon lors des 9 analyses.

<sup>3</sup>Valeur la plus faible lorsque les données provenant de chacune des analyses sont analysées de façon regroupée.

Le [Tableau 8](#) contient les données de l'étude présentées avec le pourcentage de concordance positive et négative par échantillon, où les résultats des variants sont comparés aux données de Platinum Genomes version 2016-1.0 pour le calcul du PCP. Les trois types de variants (SNV, insertions et délétions) sont combinés. Parce que la méthode de référence fournit uniquement les résultats pour les variants mononucléotidiques et les insertions/délétions, les résultats de la base non-variante sont comparés à la séquence du génome humain de référence hg19 pour les calculs du PCN.

Tableau 8 Concordance germinale par échantillon

Échantillon	Taux moyen de définitions	Variants prévus <sup>1</sup>	TP	FN	Absence de définition de variant	TN	FP	PPA	NPA	PCG
NA12877	> 99,9	4 788	4 788	0	0	756 762	0	100	100	100
NA12878	> 99,9	8 505	8 379	1	125	751 464	0	> 99,9	100	> 99,9
NA12879	> 99,9	6048	5 985	5	58	757 701	0	99,9	100	> 99,9
NA12880	> 99,9	6 993	6 930	0	63	757 638	0	100	100	100
NA12881	> 99,9	7 875	7 811	3	61	751 653	0	> 99,9	100	> 99,9
NA12882	> 99,9	6 300	6 174	3	123	754 803	0	> 99,9	100	> 99,9
NA12883	> 99,9	7 119	7 056	0	63	751 905	0	100	100	100
NA12884	> 99,9	7 182	7 119	6	57	754 146	0	99,9	100	> 99,9
NA12885	> 99,9	7 686	7 560	2	124	754 173	0	> 99,9	100	> 99,9
NA12886	> 99,9	7 245	7 182	7	56	752 469	0	99,9	100	> 99,9
NA12887	> 99,9	7 119	7 119	0	0	750 645	0	100	100	100
NA12888	> 99,9	6 804	6 804	0	0	756 065	0	100	100	100
NA12893	> 99,9	7 434	7 371	1	62	750 015	0	> 99,9	100	> 99,9

<sup>1</sup> Nombre total de variants dans tous les réplicats d'échantillons des 9 analyses.

Le [Tableau 9](#) contient les données de l'étude présentée sur une base « par échantillon », où les résultats de variants sont comparés à la méthode composite bien définie de référence. La détection est évaluée pour chaque type de variant : les SNV, les insertions et les délétions séparément. Les positions de référence sont exclues.

Tableau 9 Concordance des variants germinaux, par échantillon, par type de variant

Échantillon	SNV			Insertions			Délétions		
	Prévu	TP	FN	Prévu	TP	FN	Prévu	TP	FN
NA12877	2 331	2 331	0	1 323	1 323	0	1 134	1 134	0
NA12878	5 733	5 733	0	1 260	1 197	1	1 512	1 449	0
NA12879	3 591	3 591	0	1 323	1 260	5	1 134	1 134	0

Échantillon	SNV			Insertions			Délétions		
	Prévu	TP	FN	Prévu	TP	FN	Prévu	TP	FN
NA12880	4 221	4 221	0	1 512	1 512	0	1 260	1 197	0
NA12881	4 914	4 913	1	1 512	1 449	2	1 449	1 449	0
NA12882	3 717	3 717	0	1 386	1 323	3	1 197	1 134	0
NA12883	4 284	4 284	0	1 449	1 449	0	1 386	1 323	0
NA12884	4 284	4 284	0	1 575	1 512	6	1 323	1 323	0
NA12885	4 725	4 725	0	1 575	1 512	2	1 386	1 323	0
NA12886	4 347	4 347	0	1 449	1 386	7	1 449	1 449	0
NA12887	4 284	4 284	0	1 323	1 323	0	1 512	1 512	0
NA12888	4 158	4 158	0	1 449	1 449	0	1 197	1 197	0
NA12893	4 599	4 599	0	1 386	1 323	1	1 449	1 449	0

Les échantillons ont également été analysés pour définir des petites insertions et délétions (indels). Un résumé global est présenté dans le [Tableau 10](#). Il y avait un total de 71 indels dont la taille variait de 1 à 24 pb pour les insertions et de 1 à 25 pb pour les délétions.

Tableau 10 Résumé de la détection des indels germinaux

Variant Type	Variants prévus	TP	FN	Absence de définition de variant	PPA
Insertion	18 522	18 018	27	477	99,9
Délétion	17 388	17 073	0	315	100

L'analyse représentative portait sur 150 amplicons conçus pour couvrir un contenu génomique varié. Le contenu GC des amplicons était compris entre 0,19 et 0,87. Les amplicons comprenaient également des répétitions de mononucléotides (p. ex., PolyA, PolyT), de dinucléotides et de trinucléotides. Les données ont été compilées par amplicon (Tableau 11) pour déterminer l'effet du contenu génomique sur le pourcentage de définitions corrects. Le pourcentage de définitions exactes se rapporte aux définitions de variants et aux définitions de référence et est inférieur à 100 % en cas de définitions inexactes ou d'absences de définition.

Tableau 11 Précision au niveau de l'amplicon germlinal

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
1	1	36 450 499	36 450 591	93	93	Indel	0,22	76 167	0	0	100
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	64 701	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	74 529	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	75 348	0	0	100
5	1	228526602	228526682	81	81	Poly G (5)	0,69	66 339	0	0	100
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	57 330	0	0	100
7	1	247812041	247812128	88	88	Poly A (5), CT (3), TAA (3), indel	0,27	72 072	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	73 710	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	65 520	0	0	100
10	2	177016721	177016805	85	81	S.O.	0,65	66 339	0	0	100
11	2	186625727	186625801	75	75	Poly A (8)	0,35	61 425	0	0	100
12	2	190323504	190323591	88	88	Poly T (5)	0,42	72 072	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	71 253	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	74 529	0	0	100
15	2	228147052	228147144	93	93	Indel	0,43	76 167	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	59 787	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	74 823	0	1 344	98,2
18	3	46620561	46620643	83	83	S.O.	0,43	67 977	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	57 330	0	0	100
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG(3)	0,41	72 072	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	60 543	0	63	99,9
22	4	2233667	2233744	78	78	Poly A (6)	0,26	63 882	0	0	100
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	79 443	0	0	100
24	4	15688604	15688681	78	78	S.O.	0,29	63 882	0	0	100
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	50 778	0	0	100
26	4	102839244	102839314	71	69	Poly A (5)	0,46	56 511	0	0	100
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	50 778	0	0	100
28	5	1882081	1882158	78	75	S.O.	0,78	61 425	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	68 796	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
30	5	41069808	41069871	64	64	S.O.	0,39	52 416	0	0	100
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	67 977	0	0	100
32	5	147475343	147475409	67	67	Poly T (5)	0,37	54 873	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	74 529	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	61 425	0	0	100
35	6	6318713	6318814	102	102	PolyG (6)	0,68	83 538	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	75 348	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	76 608	0	378	99,5
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	80 262	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	77 805	0	0	100
40	6	33408498	33408583	86	86	Poly C (6)	0,7	70 434	0	0	100
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	76 986	0	0	100
42	6	112435865	112435955	91	91	Poly A (5)	0,44	74 529	0	0	100
43	7	22202076	22202148	73	73	S.O.	0,44	59 787	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	72 072	0	0	100
45	7	77365735	77365821	87	87	Poly A (7), AG(4)	0,26	71 253	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	69 615	0	0	100
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	73 710	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	74 529	0	0	100
49	7	154404519	154404599	81	66	S.O.	0,31	54 054	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	76 167	0	0	100
51	8	1817312	1817394	83	83	S.O.	0,42	67 977	0	0	100
52	8	24811020	24811109	90	89	Poly G (7), CTC(4), indel	0,61	72 171	0	720	99,0
53	8	76518625	76518691	67	67	Indel	0,3	54 873	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	80 262	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	53 235	0	0	100
56	9	107620823	107620918	96	96	S.O.	0,49	78 624	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	67 977	0	0	100
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	79 443	0	0	100
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	63 882	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	74 529	0	0	100
61	10	27317777	27317855	79	79	Poly T (5)	0,3	64 701	0	0	100
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	73 710	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	77 805	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	71 747	0	325	99,5
65	10	101611250	101611329	80	80	S.O.	0,49	65 520	0	0	100
66	10	118351373	118351453	81	81	S.O.	0,51	66 339	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
67	11	8159816	8159912	97	96	S.O.	0,45	78 624	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	57 330	0	0	100
69	11	47470345	47470444	100	100	S.O.	0,65	81 900	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	50 778	0	0	100
71	11	64418856	64418957	102	102	S.O.	0,59	83 538	0	0	100
72	11	93529612	93529684	73	73	Poly A (5)	0,4	59 787	0	0	100
73	11	101347052	101347136	85	85	S.O.	0,42	69 615	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	74 529	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	69 615	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	69 615	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	69 615	0	0	100
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	68 796	0	0	100
79	12	26811004	26811096	93	93	Poly A (7), AC(4)	0,33	76 167	0	0	100
80	12	30881766	30881846	81	81	S.O.	0,49	66 339	0	0	100
81	12	88474105	88474175	71	71	Poly A (6)	0,35	58 149	0	0	100
82	12	120966872	120966966	95	95	Poly G (5)	0,68	77 805	0	0	100
83	13	24167504	24167576	73	73	S.O.	0,52	59 787	0	0	100
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	72 072	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	72 891	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	63 063	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	54 873	0	0	100
88	14	39517884	39517966	83	83	S.O.	0,25	67 977	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	58 642	0	326	99,4
90	14	58050030	58050110	81	81	Indel	0,38	66 339	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	74 529	0	0	100
92	14	92549544	92549609	66	66	Poly A (5)	0,41	54 054	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	76 986	0	0	100
94	15	43170751	43170848	98	96	Poly C (5)	0,45	78 624	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	55 692	0	0	100
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	76 167	0	0	100
97	15	81625334	81625428	95	95	Poly T (6)	0,43	77 805	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	58 149	0	0	100
99	15	89817413	89817503	91	91	S.O.	0,36	74 529	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	57 330	0	0	100
101	16	1894910	1894972	63	63	S.O.	0,27	51 597	0	0	100
102	16	28997904	28997998	95	95	Poly C (5)	0,67	77 805	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	71 253	0	0	100
104	16	57954406	57954509	104	104	Poly C (5)	0,67	85 176	0	0	100
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	74 529	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	72 891	0	0	100
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	71 247	0	6	100
108	17	3970090	3970180	91	91	Indel	0,46	74 529	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
109	17	16084945	16085037	93	93	Indel	0,26	76 167	0	0	100
110	17	33998759	33998849	91	89	Poly T (5)	0,54	72 891	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	66 343	27	788	98,8
112	17	41244394	41244484	91	91	Poly A (5)	0,34	74 529	0	0	100
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT (4), AT(4), indel	0,26	75 348	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	64 413	0	288	99,6
115	17	64023582	64023667	86	86	Poly T (7)	0,22	70 434	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	68 796	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	54 873	0	0	100
118	18	6980478	6980568	91	91	S.O.	0,37	74 529	0	0	100
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	56 511	0	0	100
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	61 425	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	66 339	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	69 615	0	0	100
123	18	59773996	59774060	65	65	S.O.	0,48	53 235	0	0	100
124	19	625143	625241	99	99	S.O.	0,59	81 081	0	0	100
125	19	18121418	18121491	74	74	S.O.	0,68	60 605	1	0	100
126	19	18186574	18186643	70	70	S.O.	0,64	57 330	0	0	100
127	20	746056	746149	94	94	S.O.	0,61	76 986	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
128	20	10633195	10633276	82	82	AC(3)	0,59	67 158	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	62 244	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG (4), indel	0,46	57 330	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	82 719	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	54 873	0	0	100
133	20	62331904	62331994	91	88	PolyG (6)	0,73	72 072	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	71 253	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	54 054	0	0	100
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	80 262	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	71 253	0	0	100
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	56 439	0	72	99,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	73 710	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	81 900	0	0	100
141	22	32439233	32439329	97	97	S.O.	0,68	79 443	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	79 443	0	0	100
143	22	37637596	37637694	99	99	S.O.	0,6	81 081	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	75 348	0	0	100
145	X	15870424	15870492	69	69	Poly T (5)	0,26	56 511	0	0	100
146	X	135288543	135288611	69	69	Poly C (5)	0,62	56 511	0	0	100
147	X	135290777	135290847	71	71	S.O.	0,52	58 149	0	0	100
148	Y	2655397	2655461	65	0	S.O.	0,55	0	0	0	S.O.
149	Y	2655519	2655609	91	0	S.O.	0,48	0	0	0	S.O.

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
150	Y	2655609	2655679	71	0	Poly A (5)	0,37	0	0	0	S.O.

Les résultats du séquençage pour l'échantillon NA12878 ont été comparés à un génotype très sûr pour NA12878, établi par le National Institute of Standards and Technology (NIST) (v.2.19). Sur les 150 amplicons, 92 amplicons se trouvaient entièrement dans les zones génomiques très sûres, 41 amplicons se chevauchaient en partie, et 17 amplicons ne se chevauchaient pas dans la séquence NIST. Cela a résulté en 10 000 coordonnées par réplicat pour la comparaison. Les définitions de bases non-variants ont été comparés à la séquence de référence du génome humain hg19. Les résultats de précision sont indiqués dans le [Tableau 12](#).

Tableau 12 Concordance des variants germinaux pour l'échantillon NA12878 avec la base de données NIST

Échantillon	Nombre d'amplicons	Taux moyen de définitions	TP	FN	TN	FP	PPA	NPA	PCG
NA12878	133	> 99,9	6 552	1	610 470	0	> 99,9	100	> 99,9

D'après les données fournies par les neuf analyses de cette étude germinale, le Instrument NextSeq 550Dx peut séquencer ce qui suit de façon uniforme :

- Le contenu GC  $\geq 19$  % (toutes les bases définies dans 819 amplicons séquencés avec un contenu de 19 % de GC définis correctement avec un taux sans définition de 0,6 %)
- Le contenu GC  $\leq 87$  % (toutes les bases définies dans 819 amplicons séquencés avec un contenu de 87 % de GC définis correctement avec zéro sans définitions)
- Longueurs de polyA  $\leq 9$  (toutes les bases définies dans 819 amplicons séquencés contenant une répétition de polyA de neuf nucléotides définis correctement avec zéro absence de définitions)
- Longueurs de PolyT  $\leq 10$  (toutes les bases définies dans 819 amplicons séquencés contenant une répétition de PolyT de dix nucléotides définis correctement avec zéro absence de définitions)
- Longueurs de PolyG  $\leq 7$  (toutes les bases définies dans 819 amplicons séquencés contenant une répétition PolyG de sept nucléotides définis correctement avec un taux sans définitions de 1,0 %)
- Longueurs de PolyC  $\leq 6$  (toutes les bases définies dans 2 457 amplicons séquencés contenant une répétition PolyC de six nucléotides définis correctement avec zéro absence de définitions)
- Longueurs de répétitions de dinucléotides  $\leq 11x$  (toutes les bases définies dans 819 amplicons séquencés avec 11 répétitions de dinucléotides définis correctement avec un taux sans définitions de 0,5 %).
- Longueurs de répétitions de trinucleotides  $\leq 5x$  (toutes les bases définies dans 819 amplicons séquencés avec 5 répétitions de trinucleotides définis correctement avec un taux sans définitions de 0,5 %)
- Longueurs d'insertion  $\leq 24$  (66 343 sur 66 370 bases définies dans 819 amplicons séquencés contenant une insertion de 24 nucléotides définis correctement avec un taux sans définitions de 1,2 % ; aucune définition incorrecte n'est survenue dans la région contenant l'insertion de 24 nucléotides)
- Longueurs de délétion  $\leq 25$  (toutes les bases définies dans 2 457 amplicons séquencés avec une délétion de 25 nucléotides définis correctement avec zéro sans définitions)

## Variant somatique

L'étude décrite ici a été utilisée pour évaluer la précision de la définition des variants du module Somatic Variant sur le Instrument NextSeq 550Dx en utilisant le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Cette étude a utilisé une analyse représentative conçue pour demander une variété de gènes couvrant 12 588 bases (150 amplicons) à travers 23 chromosomes. L'ADN de Platinum Genome a été extrait de blocs traités au FFPE pour générer six échantillons uniques aux fins de l'évaluation dans le cadre de l'étude.

L'échantillon d'ADN GM12877 a été dilué avec l'échantillon d'ADN GM12878 afin de créer les échantillons GM12877-D5 et GM12877-D7 en tant qu'un ensemble de variants hétérozygotes uniques dont les fréquences des variants est d'environ 5 % et 7 %. L'ADN de l'échantillon GM12878 a été dilué de la même manière avec l'ADN de l'échantillon GM12877 pour créer GM12878-D5 et GM12878-D7. Chacun des échantillons a été testé en triple, à l'exception des échantillons dilués, qui ont été testés en six réplicats. Un total de neuf analyses ont été effectuées au moyen de trois instruments de séquençage, trois lots de réactifs et trois opérateurs au cours d'une période de cinq jours de démarrage. La précision a été déterminée pour les SNV, les insertions et les délétions en comparant les résultats à la méthode de référence composite bien caractérisée, Platinum Genomes version 2016- 1.0. Sauf indication contraire, les zones génomiques sûres ont été définies en fonction de cette méthode de référence.

Tableau 13 Résumé de la concordance somatique

Critère	Observations totales <sup>1</sup>	Résultat en fonction de l'observation <sup>2</sup>	Résultat en fonction de l'analyse <sup>3</sup>
PCP pour les SNV	378	98,9	99,9
PCP pour les insertions	378	96,9	99,9
PCP pour les délétions	378	97,1	99,9
NPA	378	> 99,9	> 99,9
PCG	378	> 99,9	> 99,9

<sup>1</sup>Calculées en fonction du nombre d'échantillons par analyse (42) x nombre d'analyses (9) = 378.

<sup>2</sup>Valeur la plus faible observée par réplicat d'échantillon lors des 9 analyses.

<sup>3</sup>Valeur la plus faible lorsque les données provenant de chacune des analyses sont analysées de façon regroupée.

Le [Tableau 14](#) contient les données de l'étude présentée avec les pourcentages de concordance positive et négative sur une base « par échantillon », où les résultats de variants sont comparés à la méthode composite de référence bien définie pour les calculs du PCP. Les trois types de variants (SNV, insertions et délétions) sont combinés. Parce que la méthode de référence fournit uniquement les résultats pour les variants mononucléotidiques et les insertions/délétions, les résultats de la base non-variante sont comparés à la séquence du génome humain de référence hg19 pour les calculs du PCN.

Tableau 14 Concordance des variants somatiques par échantillon

Échantillon	Taux moyen de définitions	Prévu	TP	FN	Absence de définition de variant	TN	FP	PPA	NPA	PCG
GM12877	98,7	2 052	2 025	0	27	318 682	15	100	> 99,9	> 99,9
GM12878	98,8	3 645	3 564	0	81	317 645	0	100	100	100
GM12879	99,8	2 592	2 538	0	54	323 614	2	100	> 99,9	> 99,9
GM12884	99,8	3 078	3 024	0	54	322 038	5	100	> 99,9	> 99,9
GM12885	99,8	3 294	3 213	0	81	322 121	0	100	100	100
GM12888	99,8	2 916	2 889	0	27	323 048	2	100	> 99,9	> 99,9
GM12877-D5	99,8	9 288	8 930	0	358	630 621	0	100	100	100
GM12877-D7	99,7	9 288	9 032	0	256	629 719	0	100	100	100
GM12878-D5	99,5	9 288	8 699	42	547	628 582	0	99,5	100	> 99,9
GM12878-D7	99,7	9 288	9 108	0	180	629 803	0	100	100	100

Le [Tableau 15](#) contient les données de l'étude présentée sur une base « par échantillon », où les résultats de variants sont comparés à la méthode composite bien définie de référence. La détection est évaluée pour chaque type de variant : les SNV, les insertions et les délétions séparément. Les positions de référence sont exclues.

Tableau 15 Concordance des variants somatiques, par échantillon, par type de variant

Échantillon	SNV			Insertions			Délétions		
	Prévu	TP	FN	Prévu	TP	FN	Prévu	TP	FN
GM12877	999	999	0	567	567	0	486	459	0
GM12878	2 457	2 457	0	540	513	0	648	594	0
GM12879	1 539	1 539	0	567	540	0	486	459	0
GM12884	1 836	1 836	0	675	648	0	567	540	0
GM12885	2 025	2 025	0	675	648	0	594	540	0
GM12888	1 782	1 782	0	621	621	0	513	486	0
GM12877-D5	5 454	5 392	0	1 782	1 647	0	2 052	1 891	0
GM12877-D7	5 454	5 406	0	1 782	1 728	0	2 052	1 898	0
GM12878-D5	5 454	5 192	28	1 782	1 651	9	2 052	1 856	5
GM12878-D7	5 454	5 445	0	1 782	1 719	0	2 052	1 944	0

Les dix échantillons ont également été analysés pour définir des petites insertions et délétions (indels) ([Tableau 16](#)). Il y avait un total de 71 indels dont la taille variait de 1 à 24 pb pour les insertions et de 1 à 25 pb pour les délétions.

Tableau 16 Résumé de la détection des indels somatiques

Type de variant	Variants prévus	TP	FN	Absence de définition de variant	PPA
Insertion	10 773	10 282	9	482	99,2
Délétion	11 502	10 667	5	830	> 99,9

Les 150 amplicons ont été conçus pour couvrir une variété de contenu génomique varié. Le contenu GC des amplicons était compris entre 0,19 et 0,87 %. Les amplicons comprenaient également des répétitions de mononucléotides (p. ex., PolyA, PolyT), de dinucléotides et de trinucléotides. Les données ont été compilées par amplicon ([Tableau 17](#)) pour déterminer l'effet du contenu génomique sur le pourcentage de définitions corrects. Le pourcentage de définitions exactes se rapporte aux définitions de variants et aux définitions de référence et est inférieur à 100 % en cas de définitions inexactes ou d'absences de définition.

Tableau 17 Précision somatique au niveau de l'amplicon

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
1	1	36 450 499	36 450 591	93	93	Indel	0,22	35 066	0	88	99,7
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	29 827	0	35	99,9
3	1	218353867	218353957	91	91	Indel	0,4	34 202	0	283	99,2
4	1	223906657	223906748	92	92	Indel	0,49	34 613	0	163	99,5
5	1	228526602	228526682	81	81	Poly G (5)	0,69	30 571	0	47	99,8
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	26 452	0	8	100,0
7	1	247812041	247812128	88	88	Poly A (5), CT (3), TAA (3), indel	0,27	33 148	0	116	99,7
8	2	55862774	55862863	90	90	Indel	0,28	33 928	0	92	99,7
9	2	87003930	87004009	80	80	Indel	0,38	30 218	0	22	99,9
10	2	177016721	177016805	85	81	S.O.	0,65	30 616	0	2	> 99,9
11	2	186625727	186625801	75	75	Poly A (8)	0,35	28 017	0	499	98,3
12	2	190323504	190323591	88	88	Poly T (5)	0,42	33 207	0	57	99,8
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	32 524	9	718	97,8

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	33 972	0	456	98,7
15	2	228147052	228147144	93	93	S.O.	0,43	35 051	0	103	99,7
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	27 459	0	136	99,5
17	3	4466229	4466321	93	93	AT(3), indel	0,27	34 534	0	620	98,2
18	3	46620561	46620643	83	83	S.O.	0,43	31 339	0	44	99,9
19	3	49851331	49851400	70	70	CT(3), indel	0,49	26 373	0	87	99,7
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG(3)	0,41	32 829	0	857	97,5
21	3	190106030	190106104	75	74	Indel	0,57	27 925	0	47	99,8
22	4	2233667	2233744	78	78	Poly A (6)	0,26	29 327	4	162	99,4
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	36 585	0	117	99,7
24	4	15688604	15688681	78	78	S.O.	0,29	29 427	0	57	99,8
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	23 356	5	75	99,7
26	4	102839244	102839314	71	69	Poly A (5)	0,46	25 942	0	140	99,5
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	22 944	0	560	97,6
28	5	1882081	1882158	78	75	S.O.	0,78	28 299	0	53	99,8
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	31 658	0	94	99,7

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
30	5	41069808	41069871	64	64	S.O.	0,39	24 120	0	72	99,7
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	31 297	0	77	99,8
32	5	147475343	147475409	67	67	Poly T (5)	0,37	25 277	0	55	99,8
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	34 308	0	90	99,7
34	5	155662213	155662287	75	75	Indel	0,43	28 266	0	163	99,4
35	6	6318713	6318814	102	102	PolyG (6)	0,68	38 489	0	67	99,8
36	6	24949983	24950074	92	92	Indel	0,63	34 730	0	46	99,9
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	35 057	0	483	98,6
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	36 647	0	406	98,9
39	6	32986864	32986958	95	95	Indel	0,53	35 681	0	238	99,3
40	6	33408498	33408583	86	86	Poly C (6)	0,7	32 438	0	70	99,8
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	35 441	0	91	99,7
42	6	112435865	112435955	91	91	Poly A (5)	0,44	34 354	0	44	99,9
43	7	22202076	22202148	73	73	S.O.	0,44	27 575	0	28	99,9
44	7	66276100	66276187	88	88	Indel	0,35	33 060	0	213	99,4
45	7	77365735	77365821	87	87	Poly A (7), AG(4)	0,26	32 423	0	489	98,5
46	7	110939946	110940030	85	85	Indel	0,38	32 074	0	56	99,8
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	33 791	0	281	99,2

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	34 316	0	82	99,8
49	7	154404519	154404599	81	66	S.O.	0,31	24 901	0	47	99,8
50	7	156476507	156476599	93	93	Indel	0,35	35 067	0	87	99,8
51	8	1817312	1817394	83	83	S.O.	0,42	31 365	0	9	> 99,9
52	8	24811020	24811109	90	89	Poly G (7), CTC(4), indel	0,61	32 781	0	890	97,4
53	8	76518625	76518691	67	67	Indel	0,3	25 228	0	146	99,4
54	9	103054909	103055006	98	98	PolyG (6)	0,67	36 968	0	76	99,8
55	9	105586150	105586214	65	65	Indel	0,32	24 472	0	100	99,6
56	9	107620823	107620918	96	96	S.O.	0,49	36 203	0	85	99,8
57	9	123769149	123769231	83	83	AT(3)	0,37	31 329	0	45	99,9
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	36 472	0	201	99,5
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	29 473	0	11	> 99,9
60	10	11784629	11784726	98	91	GC(3)	0,87	34 188	0	213	99,4
61	10	27317777	27317855	79	79	Poly T (5)	0,3	29 843	0	19	99,9
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	33 968	0	68	99,8
63	10	45084159	45084253	95	95	Indel	0,35	35 829	0	81	99,8
64	10	55892599	55892687	89	88	AC(11), indel	0,42	32 098	88	2 048	93,8
65	10	101611250	101611329	80	80	S.O.	0,49	30 217	0	28	99,9
66	10	118351373	118351453	81	81	S.O.	0,51	30 531	0	96	99,7

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
67	11	8159816	8159912	97	96	S.O.	0,45	36 105	0	192	99,5
68	11	30177648	30177717	70	70	Indel	0,46	26 318	0	153	99,4
69	11	47470345	47470444	100	100	S.O.	0,65	37 785	0	24	99,9
70	11	59837679	59837740	62	62	Indel	0,37	23 368	0	68	99,7
71	11	64418856	64418957	102	102	S.O.	0,59	38 546	0	10	> 99,9
72	11	93529612	93529684	73	73	Poly A (5)	0,4	27 516	0	78	99,7
73	11	101347052	101347136	85	85	S.O.	0,42	32 083	0	48	99,9
74	11	102477336	102477426	91	91	PolyG (6)	0,55	34 047	0	369	98,9
75	11	118406285	118406369	85	85	Indel	0,53	32 065	0	74	99,8
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	32 083	0	47	99,9
77	11	125769313	125769397	85	85	GA(3)	0,52	32 103	0	27	99,9
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	31 645	16	525	98,3
79	12	26811004	26811096	93	93	Poly A (7), AC(4)	0,33	34 824	0	330	99,1
80	12	30881766	30881846	81	81	S.O.	0,49	30 497	0	121	99,6
81	12	88474105	88474175	71	71	Poly A (6)	0,35	26 773	0	65	99,8
82	12	120966872	120966966	95	95	Poly G (5)	0,68	35 830	9	72	99,8
83	13	24167504	24167576	73	73	S.O.	0,52	27 498	0	114	99,6
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	32 824	0	566	98,3
85	13	44880112	44880200	89	89	Indel	0,49	33 574	0	77	99,8
86	13	77665218	77665294	77	77	Indel	0,39	29 075	0	31	99,9

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	25 313	0	13	99,9
88	14	39517884	39517966	83	83	S.O.	0,25	31 360	0	22	99,9
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	26 499	0	717	97,4
90	14	58050030	58050110	81	81	Indel	0,38	30 494	0	133	99,6
91	14	82390559	82390649	91	91	Indel	0,35	34 313	0	86	99,7
92	14	92549544	92549609	66	66	Poly A (5)	0,41	24 555	0	1 527	94,1
93	14	102808496	102808589	94	94	Indel	0,62	35 472	0	69	99,8
94	15	43170751	43170848	98	96	Poly C (5)	0,45	36 264	0	24	99,9
95	15	63446149	63446216	68	68	Indel	0,25	25 667	0	37	99,9
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	34 745	0	432	98,8
97	15	81625334	81625428	95	95	Poly T (6)	0,43	35 870	0	40	99,9
98	15	85438263	85438334	72	71	Indel	0,65	26 762	0	76	99,7
99	15	89817413	89817503	91	91	S.O.	0,36	34 286	0	112	99,7
100	15	89864274	89864343	70	70	Indel	0,56	26 449	0	11	> 99,9
101	16	1894910	1894972	63	63	S.O.	0,27	23 809	0	5	> 99,9
102	16	28997904	28997998	95	95	Poly C (5)	0,67	35 860	0	50	99,9
103	16	53682908	53682994	87	87	TA(3)	0,41	32 835	0	60	99,8
104	16	57954406	57954509	104	104	Poly C (5)	0,67	39 177	0	144	99,6
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	34 075	0	323	99,1
106	17	3563920	3564008	89	89	GC(3)	0,64	33 632	0	11	> 99,9
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	32 752	0	134	99,6
108	17	3970090	3970180	91	91	Indel	0,46	34 343	0	82	99,8

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
109	17	16084945	16085037	93	93	Indel	0,26	35 077	0	78	99,8
110	17	33998759	33998849	91	89	Poly T (5)	0,54	33 553	0	89	99,7
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	30 554	53	2 296	92,9
112	17	41244394	41244484	91	91	Poly A (5)	0,34	34 360	0	38	99,9
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT (4), AT(4), indel	0,26	34 367	0	418	98,8
114	17	61502432	61502510	79	79	Indel	0,41	29 751	0	119	99,6
115	17	64023582	64023667	86	86	Poly T (7)	0,22	32 176	0	340	99,0
116	17	72308237	72308320	84	84	GAG(3)	0,62	31 604	7	141	99,5
117	18	2616456	2616522	67	67	GA(3)	0,31	25 273	8	45	99,8
118	18	6980478	6980568	91	91	S.O.	0,37	34 386	0	12	> 99,9
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	25 692	0	399	98,5
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	27 923	0	893	96,9
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	30 598	0	20	99,9
122	18	54815665	54815749	85	85	CT(3), indel	0,45	31 969	0	161	99,5
123	18	59773996	59774060	65	65	S.O.	0,48	24 531	0	48	99,8
124	19	625143	625241	99	99	S.O.	0,59	37 298	0	124	99,7
125	19	18121418	18121491	74	74	S.O.	0,68	27 881	0	109	99,6
126	19	18186574	18186643	70	70	S.O.	0,64	26 442	0	26	99,9
127	20	746056	746149	94	94	S.O.	0,61	35 501	0	31	99,9

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
128	20	10633195	10633276	82	82	AC(3)	0,59	30 951	0	72	99,8
129	20	17705633	17705708	76	76	CT(3)	0,58	28 686	0	42	99,9
130	20	21766821	21766890	70	70	GT(3), TG (4), indel	0,46	26 372	0	88	99,7
131	20	25278421	25278521	101	101	Indel	0,63	38 159	0	20	99,9
132	20	50897302	50897368	67	67	Indel	0,36	25 188	0	544	97,9
133	20	62331904	62331994	91	88	PolyG (6)	0,73	32 969	0	309	99,1
134	20	62690860	62690946	87	87	Indel	0,57	32 818	0	77	99,8
135	21	30300823	30300888	66	66	Indel	0,35	24 758	9	181	99,2
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	36 902	0	160	99,6
137	21	36710706	36710792	87	87	GT(3), indel	0,39	32 841	0	48	99,9
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	25 939	0	280	98,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	33 942	0	78	99,8
140	22	25750774	25750873	100	100	Indel	0,63	37 733	0	86	99,8
141	22	32439233	32439329	97	97	S.O.	0,68	36 617	0	49	99,9
142	22	37409844	37409940	97	97	Indel	0,46	36 525	0	162	99,6
143	22	37637596	37637694	99	99	S.O.	0,6	37 398	0	24	99,9
144	22	47081347	47081438	92	92	Indel	0,66	34 754	0	22	99,9
145	X	15870424	15870492	69	69	Poly T (5)	0,26	26 046	0	36	99,9
146	X	135288543	135288611	69	69	Poly C (5)	0,62	26 019	0	63	99,8
147	X	135290777	135290847	71	71	S.O.	0,52	26 780	0	58	99,8
148	Y	2655397	2655461	65	0	S.O.	0,55	0	0	0	S.O.
149	Y	2655519	2655609	91	0	S.O.	0,48	0	0	0	S.O.

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les zones sûres	Contenu génomique de l'amplicon	Contenu GC	Définitions correctes	Définitions incorrectes	Absence de définitions	Pourcentage de définitions correctes
150	Y	2655609	2655679	71	0	Poly A (5)	0,37	0	0	0	S.O.

Les résultats du séquençage pour l'échantillon GM12878 ont été comparés à un génotype très sûr pour NA12878, établi par les Instituts nationaux de normalisation et de technologie (National Institute of Standards and Technology, NIST) (v.2.19). Sur les 150 amplicons, 92 amplicons se trouvaient entièrement dans les zones génomiques très sûres, 41 amplicons se chevauchaient en partie, et 17 amplicons ne se chevauchaient pas dans la séquence NIST. Cela a résulté en 10 000 coordonnées par réplicat pour la comparaison. Les définitions de bases non-variants ont été comparés à la séquence de référence du génome humain hg19. Les résultats de précision sont indiqués dans le [Tableau 18](#).

Tableau 18 Concordance somatique de l'échantillon GM12878 avec la base de données du NIST

Échantillon	Nombre d'amplicons	Taux moyen de définitions	TP	FN	TN	FP	PPA	NPA	PCG
GM12878	133	98,8	2 808	0	258 488	0	100	100	100

Sur la base des données fournies par cette étude somatique de neuf analyses, le Instrument NextSeq 550Dx peut séquencer ce qui suit de façon uniforme :

- Le contenu GC  $\geq$  19 % (toutes les bases définies dans 378 amplicons séquencés avec un contenu de 19 % de GC définis correctement avec un taux de sans définitions de 2,6 %)
- Le contenu GC  $\leq$  87 % (toutes les bases définies dans 378 amplicons séquencés avec un contenu de 87 % de GC définis correctement avec un taux de sans définitions de 0,6 %)
- Longueurs de PolyA  $\leq$  9 (toutes les bases définies dans 378 amplicons séquencés avec une répétition de PolyA de neuf nucléotides définis correctement avec un taux de sans définitions de 2,5 %)
- Longueurs de PolyT  $\leq$  10 (toutes les bases définies dans 378 amplicons séquencés contenant une répétition PolyT de dix nucléotides définis correctement avec un taux de sans définitions inférieur à 0,1 %)
- Longueurs de PolyG  $\leq$  6 (toutes les bases définies dans 2 268 amplicons séquencés contenant une répétition PolyG de six nucléotides définis correctement avec un taux de sans définitions inférieur à 0,5 %)
- Longueurs de PolyC  $\leq$  6 (toutes les bases définies dans 756 amplicons séquencés contenant une répétition PolyC de six nucléotides définis correctement avec un taux de sans définitions de 0,4 %)
- Longueurs de répétitions de dinucléotides  $\leq$  4x (toutes les bases définies dans 1 890 amplicons séquencés contenant 4 répétitions de dinucléotides définis correctement avec un taux de sans définitions de 0,9 %)
- Longueurs de répétitions de trinuécléotides  $\leq$  5x (toutes les bases définies dans 378 amplicons séquencés avec 5 répétitions de trinuécléotides définis correctement avec un taux de sans définitions de 1,4 %)
- Longueurs de PolyC  $\leq$  23 (toutes les bases définies dans 378 amplicons séquencés contenant une insertion de 23 nucléotides définis correctement avec un taux de sans définitions de 0,8 %)

- Longueurs de délétion  $\leq 25$  (toutes les bases définies dans 1 134 amplicons séquencés contenant une délétion de 25 nucléotides définis correctement avec un taux de sans définitions de 0,7 %)

## Précision

La précision de l'instrument NextSeq 550Dx a été déterminée en testant 13 échantillons de Platinum Genome à l'aide de trois instruments, trois lots de réactifs et trois opérateurs pour générer neuf analyses de séquençage sur cinq jours de démarrage. L'analyse représentative, les échantillons et la méthode de référence sont les mêmes que ceux décrits pour l'étude de précision germinale. Les contributions de précision ont été déterminées par l'analyse des composantes de variance en utilisant la VAF comme variable de réponse et en calculant les écarts-types au niveau des composantes pour l'instrument, le lot de réactifs, l'opérateur et le jour de démarrage (Tableau 19). Le nombre total d'observations utilisées dans l'analyse pour chaque composant de la variabilité de l'instrument, de l'opérateur ou du lot de réactifs était de 699, 176 et 235 pour les SNV, les insertions et les délétions, respectivement.

Tableau 19 Résultats de précision pour l'instrument NextSeq 550Dx (Écart-type)

Composant	Type de variant	Composant SD		SD total	
		Max	Médiane	Max	Médiane
Lot	SNV	0,0076	0,0002	0,0833	0,0154
	Insertion	0,0104	0,0000	0,0410	0,0157
	Délétion	0,0046	0,0005	0,0560	0,0187
Instrument	SNV	0,0114	0,0003	0,0840	0,0153
	Insertion	0,0138	0,0009	0,0407	0,0161
	Délétion	0,0079	0,0008	0,0549	0,0187
Opérateur	SNV	0,0226	0,0008	0,0841	0,0155
	Insertion	0,0344	0,0010	0,0417	0,0164
	Délétion	0,0083	0,0013	0,0547	0,0187
Jour	SNV	0,0277	0,0012	0,0825	0,0160
	Insertion	0,0235	0,0012	0,0409	0,0169
	Délétion	0,0271	0,0014	0,0548	0,0188

## Comparaison des méthodes (plateforme de séquençage)

Les échantillons de sang complet et FFPE ont été évalués sur le Instrument NextSeq 550Dx et l'instrument MiSeqDx à l'aide des flux de travail germinaux et somatiques du TruSeq Custom Amplicon Kit Dx. L'accord de fréquence des variants pour les échantillons de sang et FFPE a été évalué à l'aide de plusieurs tests représentatifs. Figure 2 représente la corrélation VAF entre les deux instruments pour un dosage représentatif et le Tableau 20 résume cette corrélation par panel de test. Sur la base de la forte corrélation entre l'instrument MiSeqDx et Instrument NextSeq 550Dx, les caractéristiques de performance liées aux facteurs pré-analytiques (par exemple, les méthodes d'extraction ou les substances interférentes) sont déterminées comme étant applicables aux deux instruments. Voir la notice du TruSeq Custom Amplicon Kit Dx pour plus de détails.

Figure 2 Corrélation de la VAF des instruments MiSeqDx et NextSeq 550Dx pour les échantillons FFPE (gauche) et sanguins (droite) à l'aide du test 1

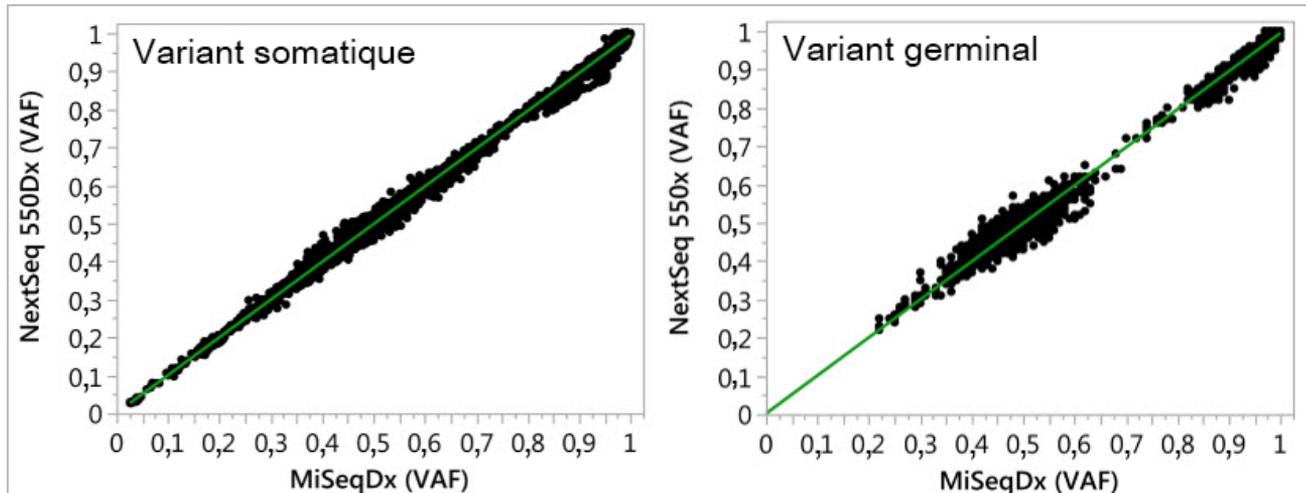


Tableau 20 Résultats de comparaison des méthodes utilisant des échantillons de sang et FFPE uniques

Source de l'ADNg	Test (panneau d'Oligo)	Réplicats biologiques (échantillons)	Réplicats techniques (par échantillon)	Observations (nombre de variants)	Pente	Intercepter	Corrélation (R <sup>2</sup> )
Sang	Test 1	45	2	8 369 <sup>1</sup>	0,992	0,002	0,995 <sup>2</sup>
Sang	Test 2	45	2	5 457	0,995	0,005	0,981
FFPE	Test 1	46	2	8 319	0,993	0,000	0,997 <sup>2</sup>
FFPE	Test 3	40	1	280	0,969	0,015	0,978

<sup>1</sup>Deux points de données ont été supprimés en fonction de la limitation indiquée pour le module Germline Variant.

<sup>2</sup>Coefficient de détermination pour les tracés VAF comme illustré sur la Figure 2.

## Reproductibilité

La reproductibilité du Instrument NextSeq 550Dx a été évaluée au moyen d'échantillons de Platinum Genome et d'un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases sur 23 chromosomes différents utilisant 150 amplicons. Les tests germinaux comprenaient sept réplicats de 13 échantillons ; les tests somatiques comprenaient six réplicats de sept échantillons à différents niveaux de VAF. Les échantillons ont été préparés à l'aide du TruSeq Custom Amplicon Kit Dx.

Les tests ont été effectués sur trois sites externes à l'aide d'un lot de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Un seul instrument NextSeq 550Dx a été utilisé sur chaque site. Deux opérateurs ont effectué les tests sur chaque site. Chaque opérateur a réalisé les tests lors de trois jours non consécutifs pour chacun des types d'échantillons, pour un total de 36 analyses sur les trois sites. Il en a résulté 18 analyses pour chacun des flux de travail germinaux et somatiques.

## Variant germinal

Les variants germinaux avec un taux de VAF  $\geq 0,2$  sont rapportés comme positifs (variant). Pour les variants germinaux positifs attendus, les données ont été évaluées pour le taux d'absence de définition et la définition positive correcte (PPC) dans chaque type de variant (SNV, insertion, délétion). Le [Tableau 21](#) résume les taux observés, ainsi que les niveaux de confiance inférieurs (NCI) et supérieurs (NCS) à 95 % calculés selon la méthode de Wilson, pour chacun des types de variant.

Tableau 21 Observations des définitions de variants germinaux pour les résultats positifs prévus selon le type de variant

Type de variant	Absence de définition			Corriger la définition positive				
	Observée	Total	Pourcentage	Observée	Total	Pourcentage	LCL 95 %	NCS 95 %
SNV	16	110 376	0,014	110 349	110 360	99,99	99,98	99,99
Insertions	1 026	37 044	2,77	36 018	36 018	100	99,99	100,00
Délétions	648	34 776	1,86	34 128	34 128	100	99,99	100,00

Les variants germinaux avec un taux de VAF  $< 0,2$  sont rapportés comme négatifs (type sauvage). Dans le cas des emplacements germinaux négatifs prévus, les données ont été évaluées selon les taux d'absence de définition et la définition correcte de type sauvage. Le [Tableau 22](#) résume les taux observés, ainsi que les niveaux de confiance inférieurs (NCI) et supérieurs (NCS) à 95 % calculés selon la méthode de Wilson Score.

Tableau 22 Observations des définitions de variants germinaux pour les résultats négatifs prévus

Type de variant	Absence de définition			Définition négative correcte				
	Observée	Total	Pourcentage	Observée	Total	Pourcentage	LCL 95 %	NCS 95 %
Type sauvage	4 883	19 600 182	0,025	19 595 299	19 595 299	100	100,00	100,00

Les variants germinaux avec un taux de VAF  $\geq 0,2$  et  $< 0,7$  sont appelés positifs hétérozygotes pour le variant, et les variants avec un taux de VAF  $\geq 0,7$  sont appelés positifs homozygotes pour le variant. Des échantillons germinaux avec des variants hétérozygotes ont été utilisés pour déterminer si la variabilité inhérente au test affecterait la définition du génotype. Le Cx a été déterminé pour les deux seuils (0,2 pour les génotypes hétérozygotes et 0,7 pour les génotypes homozygotes), où x est la proportion de tests répétés qui dépassent le seuil. Pour le seuil inférieur de 0,2 VAF, le Cx était  $\geq 99,999$  %, ce qui indique que  $\geq 99,999$  % des variants hétérozygotes seraient appelés hétérozygotes. En ce qui concerne le seuil supérieur de 0,7 VAF, le Cx était  $\leq 0,001$  %, ce qui indique que  $\leq 0,001$  % des variants hétérozygotes seraient appelés homozygotes. Le [Tableau 23](#) résume les résultats par type de variant.

Les variants germinaux avec un taux de VAF  $\geq 0,2$  et  $< 0,7$  sont appelés positifs hétérozygotes pour le variant, et les variants avec un taux de VAF  $\geq 0,7$  sont appelés positifs homozygotes pour le variant. Des échantillons germinaux avec des variants hétérozygotes ont été utilisés pour déterminer si la variabilité inhérente au test affecterait la définition du génotype. Le Cx a été déterminé pour les deux seuils (0,2 pour les génotypes hétérozygotes et 0,7 pour les génotypes homozygotes), où x est la proportion de tests répétés qui dépassent le seuil. En ce qui concerne le seuil inférieur de 0,2 VAF, le Cx était  $\geq 99,999$  %, indiquant ainsi que  $\geq 99,999$  % des

variants hétérozygotes seraient appelés hétérozygotes. Pour le seuil supérieur de 0,7 VAF, le Cx était  $\leq 0,001\%$ , ce qui indique que  $\leq 0,001\%$  des variants hétérozygotes seraient appelés homozygotes. Le [Tableau 23](#) résume les résultats par type de variant.

Tableau 23 Valeurs Cx germinales pour les variants hétérozygotes

Type de variant	Seuil à 0,2 VAF	Seuil à 0,7 VAF
	$\geq C99,999\%$	$\leq C0,001\%$
SNV	94/94	94/94
Insertions	24/24	24/24
Délétions	35/35	35/35
Total	153	153

## Variant somatique

Les variants somatiques avec des taux de VAF  $\geq 0,026$  sont rapportés comme positifs (variant). Les observations avec des taux de VAF  $\geq 0,01$  et  $< 0,026$  ont été considérées comme équivoques aux fins de cette analyse (ni positives ni négatives, marquées comme une faible fréquence de variants). Pour évaluer la performance, les résultats ont été calculés de trois manières :

- Meilleur cas : Tout résultat équivoque a été considéré comme un définition positif correct (accord avec les résultats attendus)
- Pire cas : Tout résultat équivoque a été considéré comme un définition incorrect (désaccord avec les résultats attendus)
- Cas d'exclusion : Tout résultat équivoque a été exclu de l'analyse

Trois tableaux, le [Tableau 24](#), le [Tableau 25](#) et le [Tableau 26](#), résument les résultats de la définition pour le meilleur cas, le pire cas et le cas d'exclusion, respectivement, ainsi que les niveaux de confiance à 95 % inférieurs et supérieurs (LCL/UCL) calculés à l'aide de la méthode Wilson Score.

Tableau 24 Observations des définitions de variants somatiques pour les résultats positifs prévus selon le type de variant (Meilleur cas)

Type de variant	Corriger la définition positive				
	Observé	Total	Pourcentage	LCL 95 %	NCS 95 %
SNV	54 346	54 346	100	99,99	100,00
Insertions	18 036	18 036	100	99,98	100,00
Délétions	18 381	18 381	100	99,98	100,00

Tableau 25 Observations des définitions de variants somatiques pour les résultats positifs prévus selon le type de variant (Pire cas)

Type de variant	Corriger la définition positive				
	Observée	Total	Pourcentage	LCL 95 %	NCS 95 %
SNV	54 346	54 346	100	99,99	100,00
Insertions	18 000	18 036	99,8	99,72	99,86
Délétions	18 381	18 381	100	99,98	100,00

Tableau 26 Observations somatiques des appels pour les résultats positifs attendus par type de variant (appels équivoques supprimés)

Type de variant	Corriger la définition positive				
	Observée	Total	Pourcentage	LCL 95 %	NCS 95 %
SNV	54 346	54 346	100	99,99	100,00
Insertions	18 000	18 000	100	99,98	100,00
Délétions	18 381	18 381	100	99,98	100,00

Les variants somatiques avec un taux de VAF < 0,01 sont rapportés comme des appels négatifs (de type sauvage). Dans le cas des emplacements somatiques négatifs prévus, les données ont été évaluées selon les taux d'absence de définition et la définition correcte de type sauvage. Les définitions de type sauvage corrects ont été déterminées en excluant les absences de définition et en soustrayant les définitions observées qui sont tombées dans la zone équivoque (niveaux de VAF  $\geq 0,01$  et  $< 0,026$ ) ainsi que les définitions incorrectes qui étaient au-dessus du seuil (niveaux de VAF  $\geq 0,026$ ) du total. Le [Tableau 27](#) résume les résultats observés, totaux et en pourcentage pour les emplacements somatiques négatifs pour le taux d'absence de définition et le taux de définition de type sauvage correct, ainsi que les niveaux de confiance à 95 % inférieurs et supérieurs (LCL/UCL) calculés à l'aide de la méthode score de Wilson.

Tableau 27 Observations des définitions de variants somatiques pour les résultats négatifs prévus

Variant Type	Absence de définition			Définition correcte						
	Observée	Total	Pourcentage	Équivoque	Incorrect	Correct	Total	Pourcentage	LCL 95 %	NCS 95 %
Type sauvage	36 326	8 909 676	0,408	2254	121	8 870 975	8 873 350	99,97	99,972	99,974

Des échantillons somatiques à différents niveaux de VAF pour le même variant ont été évalués pour déterminer le C95 du test (dans chaque type de variant). Pour évaluer la variabilité près de la limite du test, des échantillons dont les taux de VAF attendus étaient compris entre 0,02 et 0,07 ont été utilisés. Le C95 a été déterminé pour chaque variant, avec le C95 le plus élevé pour chaque type de variant rapporté dans le [Tableau 28](#).

Tableau 28 Résumé du C95 Somatique

Type de variant	N	C95
SNV	74	0,0613

Insertion	24	0,0573
Délétion	33	0,0575

## NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) Performance

### Présentation

Le NextSeq 550Dx est pris en charge par deux kits de réactifs : le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) et le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Pour démontrer que le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) peut répondre aux exigences de performance analytique vérifiées et validées avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), des études ont été menées avec le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Deux préparations de bibliothèque utilisant le TruSeq Custom Amplicon Kit Dx ont été effectuées, l'une avec le flux de travail Germinal et l'autre avec le flux de travail Somatique. Les bibliothèques de chaque flux de travail ont été testées avec trois lots de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) à l'aide de trois instruments NextSeq 550Dx. En outre, les tests pour chaque flux de travail comprenaient une série unique avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

### Sensibilité analytique (Limite du blanc [LB] et Limite de détection [LD])

La vérification avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) a démontré que le Instrument NextSeq 550Dx pouvait détecter des variants à une VAF de 0,05 avec une erreur de type II  $\leq 0,05$  et que le seuil d'une VAF de 0,026 utilisé par le module Somatic Variant (LB réelle) prend en charge une erreur de type I  $\leq 0,01$ . Sur la base de ces revendications, on s'attend à ce qu'un variant à une VAF de 0,05 soit supérieur ou égal à une VAF de 0,026, 95 % du temps, et qu'une position de type sauvage soit inférieure à une VAF de 0,026, 99 % du temps. Pour s'assurer que ces revendications ont été satisfaites avec le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), des mesures répétées ont été effectuées sur l'instrument NextSeq 550Dx avec des échantillons de type sauvage (échantillons LB) et avec des échantillons contenant des variants à une VAF de 0,05 (échantillons LD) à l'aide du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). La proportion de définitions au-dessus et en dessous du seuil de 0,026 a ensuite été comparée aux revendications établies avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Les tests comprenaient deux échantillons de la LD, chacun avec un ensemble unique de variants ciblés à une VAF de 0,05 et des échantillons de LB correspondants de type sauvage pour les variants ciblés. Pour la préparation de la bibliothèque, les échantillons LD et LB ont été traités en réplicats de huit et sept, respectivement, à l'aide du TruSeq Custom Amplicon Kit Dx. Les bibliothèques ont été initialement séquencées à l'aide du NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) pour identifier les variants/coordonnées génomiques pour l'évaluation de la LB/LD avec le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Tous les variants avec une VAF moyenne comprise entre 0,045 et 0,055 (variants de la LD) d'après les résultats du NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) ont été utilisés pour l'analyse de la LD (N = 51 variants). Pour l'analyse de la LB, les 51 coordonnées génomiques correspondantes ont été évaluées.

Pour l'évaluation du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), les bibliothèques ont été séquencées en trois analyses sur trois jours consécutifs en utilisant le même lot de kits d'instruments et de réactifs. Ce test s'élevait à 24 réplicats pour chacun des 51 variants de la LD et 21 réplicats pour chacune des positions de type sauvage correspondantes. La proportion de définitions de type sauvage avec une VAF < 0,026 est présentée dans le [Tableau 29](#). La proportion de définitions de variants de la LD avec une VAF supérieure ou égale à 0,026 est présentée dans le [Tableau 30](#).

Tableau 29 Proportion de définitions < 0,026 pour les positions de type sauvage (évaluation de la réclamation LB)

Variant Type	Positions évaluées	Observations totales	Nombre de mesures VAF ≥ 2,6 %	Proportion < 2,6 %	Proportion 95 % Intervalle de confiance
SNV	32	672	0	1	0,994 – 1
Insertion	11	231	0	1	0,984 – 1
Délétion	8	168	0	1	0,978 – 1

Tableau 30 Proportion de définitions ≥ 0,026 VAF pour les variants de la LD (évaluation de la revendication de la LD)

Variant Type	Positions évaluées	Observations totales	Nombre de mesures VAF < 2,6 %	Nombre de mesures VAF ≥ 2,6 %	Proportion ≥ 2,6 %	Proportion 95 % Intervalle de confiance
SNV	32	768	1	767	0,999	0,993 – 1
Insertion	11	264	3	261	0,989	0,967 – 0,996
Délétion	8	192	2	190	0,99	0,963 – 0,997

## Précision

### Variant germinale

L'étude suivante a été menée pour évaluer la précision de définition des variants du module Germline Variant sur l'instrument NextSeq 550Dx à l'aide du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Douze échantillons Platinum Genome uniques ont été testés à l'aide d'un test représentatif. Au total, 11 analyses ont été réalisées à l'aide de trois instruments NextSeq 550Dx et de trois lots du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

La précision a été déterminée pour les SNV, les insertions et les délétions en comparant les résultats à la méthode de référence composite bien caractérisée, Platinum Genomes version 2016-1.0. Les résultats de précision d'une seule analyse de séquençage avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) sont fournis à titre de référence. Un résumé des résultats est fourni dans le [Tableau 31](#).

Tableau 31 Résumé de la concordance des variants germinaux

Critère	Observations totales (v2.5) <sup>1</sup>	Résultat par observation (v2.5) <sup>2</sup>	Résultat par observation (v2) <sup>3</sup>	Résultat par exécution (v2.5) <sup>4</sup>	Résultat par exécution (v2) <sup>4</sup>
PCP pour les SNV	1 056	98,7	98,7	> 99,9	> 99,9
PCP pour les insertions	1 056	100	100	100	100
PCP pour les délétions	1 056	95,2	95,2	> 99,9	> 99,9
NPA	1 056	100	100	100	100
PCG	1 056	> 99,9	> 99,9	> 99,9	> 99,9

<sup>1</sup>Calculée comme le nombre d'échantillons par analyse x le nombre d'analyses (96 échantillons par série x 11 analyses = 1 056 observations).

<sup>2</sup>Valeur observée la plus faible par réplicat d'échantillon sur toutes les séries (sur la base de 11 séries pour le NextSeq 550Dx High Output Reagent Kit v2.5).

<sup>3</sup>Valeur la plus faible observée par réplicat d'échantillon lors d'une série (total de 96 observations).

<sup>4</sup> Valeur la plus faible lorsque les données provenant de chacune des analyses sont analysées de façon regroupée.

## Variant somatique

L'étude suivante a été menée pour évaluer la précision de définition des variants du module SOmatic Variant sur l'instrument NextSeq 550Dx à l'aide du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Dix échantillons FFPE du Platinum Genome (deux avec des variants dilués à 0,05 VAF) ont été testés à l'aide d'un test représentatif. Au total, 11 analyses ont été réalisées à l'aide de trois instruments NextSeq 550Dx et de trois lots du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

La précision a été déterminée pour les SNV, les insertions et les délétions en comparant les résultats à la méthode de référence composite bien caractérisée, Platinum Genomes version 2016-1.0. Les résultats de précision d'une seule analyse de séquençage avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) sont fournis à titre de référence. Un résumé des résultats est fourni dans le [Tableau 32](#).

Tableau 32 Résumé de la concordance somatique

Critère	Observations totales (v2.5) <sup>1</sup>	Résultat en fonction de l'observation (v2.5) <sup>2</sup>	Résultat en fonction de l'observation (v2) <sup>3</sup>	Résultat en fonction de la série (v2.5) <sup>4</sup>	Résultat en fonction de la série (v2) <sup>4</sup>
PCP pour les SNV	528	100	100	100	100
PCP pour les insertions	528	96,9	96,9	> 99,9	> 99,9
PCP pour les délétions	528	100	100	100	100

NPA	528	> 99,9	> 99,9	> 99,9	> 99,9
PCG	528	> 99,9	> 99,9	> 99,9	> 99,9

<sup>1</sup>Calculé comme le nombre d'échantillons par série x nombre de séries (48 échantillons par série x 11 séries = 528 observations).

<sup>2</sup>Valeur observée la plus faible par réplicat d'échantillon sur toutes les séries (sur la base de 11 séries pour le NextSeq 550Dx High Output Reagent Kit v2.5).

<sup>3</sup>Valeur la plus faible observée par réplicat d'échantillon lors d'une série (total de 96 observations).

<sup>4</sup>Valeur la plus faible lorsque les données provenant de chacune des séries sont analysées de façon regroupée.

## Précision

### Variant germlinal

La précision du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) avec le module Germline Variant a été évaluée à l'aide d'échantillons Platinum Genome et d'un test représentatif. Les tests consistaient en une seule préparation de bibliothèque à l'aide du TruSeq Custom Amplicon Kit Dx et comprenaient 12 échantillons traités avec huit réplicats chacun. Les bibliothèques ont été séquencées avec trois lots de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) et trois instruments NextSeq 550Dx pour un total de neuf analyses de séquençage.

Des échantillons avec des variants hétérozygotes ont été utilisés pour déterminer si la variabilité inhérente au test affecterait la définition du génotype (N = 153 variants hétérozygotes uniques). Le Cx a été déterminé pour les deux seuils du module de Germline Variant (0,2 pour les génotypes hétérozygotes et 0,7 pour les génotypes homozygotes), où x est la proportion de tests répétés qui dépassent le seuil. Pour le seuil inférieur d'une VAF de 0,2, le variant avec le Cx minimum pour le NextSeq 550Dx Reagent Kit v2.5 (300 cycles) était > 99,9 %, ce qui indique que > 99,9 % des variants hétérozygotes seraient appelés hétérozygotes. Pour le seuil supérieur d'une VAF de 0,7, le variant avec le Cx maximum pour le NextSeq 550Dx Reagent Kit v2.5 (300 cycles) était < 1,5 %, ce qui indique que ≤ 1,5 % des variants hétérozygotes seraient appelés homozygotes. Le [Tableau 33](#) résume les résultats par type de variant. Les valeurs Cx d'une seule analyse de séquençage avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) sont fournis à titre de référence.

Tableau 33 Valeurs Cx germinales pour les variants hétérozygotes

Type de variant	N	Seuil à 0,2 VAF		Seuil à 0,7 VAF	
		Cx min. (v2.5) <sup>1</sup>	Cx min. (v2) <sup>2</sup>	Cx max. (v2.5) <sup>1</sup>	Cx max. (v2) <sup>2</sup>
SNV	94	> 99,9 %	> 99,9 %	1,5 %	1,0 %
Insertions	24	100 %	100 %	0 %	< 0,1 %
Délétions	35	100 %	> 99,9 %	< 0,1 %	< 0,1 %

<sup>1</sup>Valeurs Cx basées sur les estimations de l'écart-type total issues de l'analyse des composantes de variance.

<sup>2</sup>Valeurs Cx basées sur des écarts types d'échantillon.

### Variant somatique

La précision du NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) avec le module Somatic Variant a été évaluée à l'aide d'échantillons FFPE Platinum Genome et d'un test représentatif. Les tests consistaient en une seule préparation de bibliothèque à l'aide du TruSeq Custom Amplicon Kit Dx et comprenaient deux échantillons avec huit réplicats chacun. Les bibliothèques ont été séquencées à l'aide de trois lots de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) et de trois instruments NextSeq 550Dx pour un total de neuf analyses de séquençage.

Des variants somatiques avec des taux de VAF attendus  $\leq 0,10$  VAF (N = 131 variants uniques) ont été utilisés pour évaluer la variabilité de l'instrument près du seuil de VAF du module Somatic Variant (les variants somatiques avec un taux de VAF  $\geq 0,026$  sont appelés positifs pour le variant). Les valeurs C95 ont été déterminées pour chacun des variants somatiques. Les valeurs C95 représentent le VAF auquel la probabilité d'être supérieure au seuil du VAF du module Somatic Variant est de 95 %. Les valeurs C95 les plus élevées par type de variant sont indiquées dans le [Tableau 34](#). Les résultats C95 d'une seule analyse de séquençage avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) sont fournis à titre de référence.

Tableau 34 Résumé C95 somatique

Type de variant	Nombre de variants évalués	C95 (v2.5) <sup>1</sup>	C95 (v2) <sup>2</sup>
SNV	74	0,064	0,063
Insertions	24	0,062	0,061
Délétions	33	0,060	0,060

<sup>1</sup>Valeurs C95 basées sur les estimations de l'écart-type total issues de l'analyse des composantes de variance.

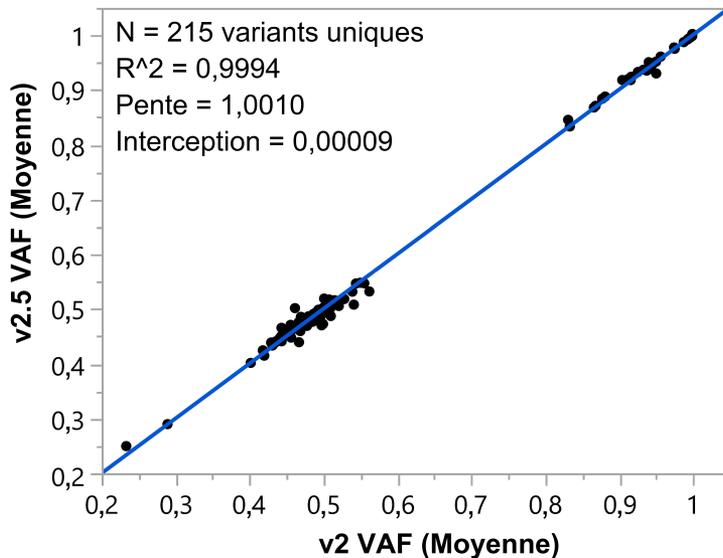
<sup>2</sup>Valeurs C95 basées sur des écarts types d'échantillon.

## Comparaison des méthodes (kit réactif)

### Variant germlinal

Les VAF moyennes provenant de 215 variants uniques ont été évaluées dans le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) et le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) à l'aide des résultats générés par le module Germline Variant. Les VAF moyennes ont été calculées à partir de 11 analyses de séquençage (v2.5) et d'une analyse de séquençage (v2). Au moins huit réplicats ont été utilisés pour calculer la moyenne pour chaque variant. [Figure 3](#) représente la corrélation de VAF entre les deux kits de réactifs. D'après la forte corrélation linéaire de la VAF et la similitude des résultats entre les kits de réactifs, les caractéristiques de performance initialement vérifiées et validées avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) avec le module Germline Variant sont déterminées comme étant applicables au NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

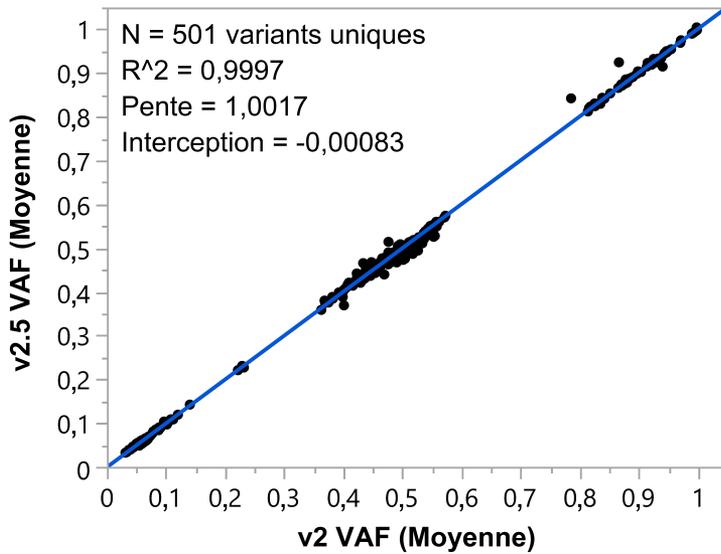
Figure 3 Corrélation de la fréquence allélique des variants (VAF) du module Germline Variant entre le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) et le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



## Variant somatique

Les VAF moyennes pour 501 variants uniques ont été évaluées dans le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) et le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) à l'aide des résultats générés par le module Somatic Variant. Les VAF moyennes ont été calculées à partir de 11 analyses de séquençage (v2.5) et d'une analyse de séquençage (v2). Au moins trois réplicats ont été utilisés pour calculer la moyenne pour chaque variant unique. [Figure 4](#) représente la corrélation de VAF entre les deux kits de réactifs. Sur la base de la corrélation de la VAF et de la similitude des résultats entre les kits de réactifs, les caractéristiques de performance vérifiées et validées avec le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) avec le module Somatic Variant sont déterminées comme étant applicables au NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Figure 4 Corrélation de la fréquence allélique des variants (VAF) du module Somatic Variant entre le NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) et le NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



# Historique des modifications

Document	Date	Description de la modification
Document n° 200031448 v00	Juin 2023	<p>Publication initiale.</p> <p>Le document précédent 1000000030326 a été remplacé par celui-ci.</p> <p>Changements apportés au document 1000000030326 v6 par rapport à ce nouveau document :</p> <ul style="list-style-type: none"><li>• Ajout de contenu pour prendre en charge le serveur Illumina DRAGEN en option pour NextSeq 550Dx.</li><li>• Mise à jour du numéro de référence du filtre à air.</li></ul> <p>Modifications apportées précédemment au document 1000000030326 :</p> <ul style="list-style-type: none"><li>• Mises à jour effectuées pour corriger le contenu ajouté par inadvertance à partir du logiciel source.</li><li>• Ajout d'une déclaration d'avertissements et de précautions concernant le signalement des incidents graves.</li><li>• Ajout d'une déclaration aux principes de procédure spécifiant l'utilisateur concerné.</li><li>• Suppression de la référence au High Output Reagent Kit v2 (300 cycles).</li><li>• Ajout d'une référence au High Output Reagent Kit v2.5 (75 cycles).</li><li>• Ajout d'un tableau sur l'historique des révisions. Adresse du représentant UE autorisé mise à jour.</li></ul>

## Brevets et Marques

Ce document et son contenu sont la propriété exclusive d'Illumina, Inc. et ses filiales (« Illumina »), et sont destinés à un usage contractuel de ses clients en lien avec l'utilisation du ou des produits décrits dans la présente et à aucune autre utilisation. Ce document et son contenu ne seront utilisés ou distribués à aucune autre fin et ne seront communiqués, divulgués ou reproduits d'aucune façon sans le consentement écrit préalable d'Illumina. Par le biais de ce document, Illumina ne cède aucune licence en vertu de son brevet, de sa marque de commerce, de son copyright ou de ses droits traditionnels ni des droits similaires d'un tiers quelconque.

Les instructions présentes dans ce document doivent être strictement et explicitement respectées par le personnel qualifié et correctement formé afin d'assurer une utilisation correcte et sécuritaire du ou des produits décrits dans la présente. Tout le contenu de ce document doit être entièrement lu et compris avant d'utiliser le ou les produits.

LE FAIT DE NE PAS LIRE ENTièrement ET DE NE PAS SUIVRE EXPLICITEMENT TOUTES LES INSTRUCTIONS CONTENUES DANS LA PRÉSENTE PEUT CAUSER DES DOMMAGES AU OU AUX PRODUITS, DES BLESSURES AUX PERSONNES, Y COMPRIS AUX UTILISATEURS OU À D'AUTRES PERSONNES, ET DES DOMMAGES À D'AUTRES BIENS, ET ANNULERA TOUTE GARANTIE APPLICABLE AU OU AUX PRODUITS.

ILLUMINA N'ASSUME AUCUNE RESPONSABILITÉ QUANT AUX DOMMAGES DÉCOULANT D'UNE MAUVAISE UTILISATION DU OU DES PRODUITS DÉCRITS DANS LA PRÉSENTE (Y COMPRIS LES PARTIES DE CELLE-CI OU LE LOGICIEL).

© 2023 Illumina, Inc. Tous droits réservés.

Toutes les marques sont la propriété d'Illumina, Inc. ou de leurs propriétaires respectifs. Pour en savoir plus sur les marques, consultez la page [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

## Coordonnées



Illumina, Inc.  
5200 Illumina Way  
San Diego, Californie 92122 États-Unis  
+(1) 800 809 ILMN (4566)  
+(1) 858 202 4566 (en dehors de l'Amérique du Nord)  
techsupport@illumina.com  
www.illumina.com



Illumina Netherlands B.V.  
Steenoven 19  
5626 DK Eindhoven  
The Netherlands

**Commanditaire australien**  
Illumina Australia Pty Ltd  
Nursing Association Building  
Level 3, 535 Elizabeth Street  
Melbourne, VIC 3000  
Australie

## Étiquette du produit

Pour obtenir des informations détaillées sur les symboles susceptibles d'apparaître sur l'emballage et l'étiquette du produit, consultez la légende des symboles pour votre trousse sur l'onglet *Documentation* à l'adresse [support.illumina.com](http://support.illumina.com).