

Instrument NextSeq^{MC} 550Dx

DESTINÉ AU DIAGNOSTIC IN VITRO UNIQUEMENT

Référence n° 20005715

Utilisation prévue

L'instrument NextSeq 550Dx est destiné au séquençage ciblé de bibliothèques d'ADN provenant d'ADN génomique humain extrait de sang entier périphérique ou de tissus fixés au formol et imprégnés à la paraffine (FPPE), à utiliser avec les tests de diagnostic *in vitro* (DIV). L'instrument NextSeq 550Dx n'est pas destiné au séquençage de génome entier ou de novo. L'instrument NextSeq 550Dx doit être utilisé avec des réactifs de diagnostic *in vitro* enregistrés, certifiés *ou approuvés*, ainsi qu'avec un logiciel d'analyse.

Principes procéduraux

L'instrument NextSeq 550Dx d'Illumina est destiné au séquençage de bibliothèques d'ADN avec les tests de diagnostic *in vitro*. Au départ, l'instrument NextSeq 550Dx utilise des bibliothèques préparées à partir d'ADN pour lesquelles des index d'échantillons et des séquences de saisie sont ajoutés à des cibles amplifiées. Les bibliothèques d'échantillons sont capturées sur la Flow Cell puis séquencées sur l'instrument à l'aide de 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) exécute les analyses d'images et la définition des bases tout en affectant un score de qualité à chacune des bases de chaque cycle de séquençage. Une fois l'analyse primaire terminée, une analyse secondaire peut être faite sur l'instrument pour le traitement des appels de bases. NextSeq 550Dx utilise différents modules pour l'analyse secondaire en fonction du flux de travail. Dans le cas des modules d'analyse des variants germinaux et somatiques, le traitement comprend le démultiplexage, la génération de fichiers FASTQ, l'alignement, l'appel des variants et la génération de fichiers VCF (VCF et gVCF). Les fichiers VCF (VCF et gVCF) contiennent des renseignements sur les variants trouvés à des positions spécifiques dans un génome de référence.

Configuration à amorçage double

NextSeq 550Dx comprend la configuration à amorçage double, qui permet l'utilisation de l'instrument en mode diagnostic ou en mode recherche uniquement. Les tests de séquençage diagnostique *in vitro*, y compris les modules des variants germinaux et somatiques, sont exécutés en mode diagnostic. Seuls les réactifs de séquençage diagnostique *in vitro* peuvent être utilisés en mode diagnostic. Les caractéristiques de performance et les limites de la procédure de l'instrument NextSeq 550Dx ont été établies au moyen des modules des variants germinaux et somatiques en mode diagnostic.

Limites de la procédure

- 1 Destiné au diagnostic *in vitro*.
- 2 Les modules d'analyse des variants germinaux et somatiques, lorsqu'ils sont utilisés avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) ou avec la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) peuvent fournir ce qui suit :
 - ▶ Sortie de séquençage ≥ 90 gigabases (Gb)
 - ▶ Longueur de lecture (en analyse à lecture appariée) de 2 × 150 paires de base (pb)
 - ▶ Bases égales ou supérieures à Q30 ≥ 75 % à une longueur de lecture de 2 x 150 bp

- Au moins 75 % des bases ont un score de qualité Phred supérieur ou égal à 30, ce qui indique une exactitude de la définition des bases supérieure à 99,9 %
- 3 Les lectures comportant des indels (insertions, délétions ou combinaisons) dont la longueur du contenu est supérieure à 25 pb ne sont pas alignées par le logiciel de test. Par conséquent, les indels de plus de 25 pb ne sont pas détectables par le logiciel de test.
 - 4 Le logiciel de test pourrait ne pas aligner les lectures d'amplicons affichant un contenu de variants extrême et que la région soit ainsi signalée comme étant de type sauvage. Les contenus extrêmes comprennent :
 - ▶ les lectures comportant plus de trois indels;
 - ▶ les lectures d'une longueur d'au moins 30 pb dont le contenu de variants mononucléotides (SNV) est supérieur à 4 % de la longueur totale d'amplicon cible (excluant les régions sondées);
 - ▶ les lectures d'une longueur inférieure à 30 pb dont le contenu de SNV est supérieur à 10 % de la longueur totale d'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 omis lorsqu'ils s'étendent sur deux amplicons qui se chevauchent, si la longueur de la délétion est supérieure ou égale au chevauchement entre les amplicons.
 - 7 Le système ne peut détecter les indels s'ils sont directement adjacents à un primer et qu'il n'y a pas d'amplicons se chevauchant. Pour les régions contenant des amplicons se chevauchant, le test ne peut pas détecter les délétions lorsque la zone de chevauchement est plus petite que la délétion à détecter. Par exemple, si la région de chevauchement entre deux amplicons adjacents est de deux bases, le test ne détectera aucune délétion, pas même ces deux bases. Une délétion de base unique au niveau de l'une ou l'autre de ces bases peut être détectée.
 - 8 Comme avec n'importe quel flux de travail de préparation de librairie de type hybridation, les polymorphismes, mutations, insertions ou délétions sous-jacents dans les régions de liaison d'un oligonucléotide peuvent affecter les allèles sondés et les définitions obtenues durant le séquençage. Par exemple :
 - ▶ Un variant en phase avec un variant dans la région du primer peut ne pas être amplifié, ce qui entraîne un faux négatif.
 - ▶ Les variants dans la région du primer pourraient empêcher l'amplification de l'allèle de référence et entraîner une définition de variant homozygote erronée.
 - ▶ Les variants d'indels dans la région du primer peuvent entraîner un faux positif à la fin de la lecture adjacente au primer.
 - 9 Les indels peuvent être filtrés en raison du biais de brin, s'ils se trouvent près de la fin d'une lecture et sont écrêtés doucement durant l'alignement.
 - 10 Les petits MNV n'ont pas été validés et ne sont signalés que dans le module d'analyse des variants somatiques.
 - 11 Les délétions sont signalées dans les fichiers VCF aux coordonnées de la base précédente, selon le format VCF. Par conséquent, il faut tenir compte des variants adjacents avant de déclarer qu'un appel de base est une référence homozygote.
 - 12 Limites spécifiques au flux de travail germinale :
 - ▶ L'instrument NextSeq 550Dx, lorsqu'il est utilisé avec le module d'analyse des variants germinaux de Local Run Manager pour NextSeq 550Dx, est conçu pour produire des résultats qualitatifs aux fins de la définition des variants germinaux (p. ex., homozygote, hétérozygote, type sauvage).
 - ▶ Lorsque le module d'analyse des variants germinaux est utilisé, la couverture minimale par amplicon requise pour obtenir une définition de variants précise est de 150x. Par conséquent, il faut 150 fragments d'ADN pris en charge, ce qui équivaut à 300 lectures appariées se chevauchant. Le nombre d'échantillons et le nombre total de bases ciblées ont une incidence sur la couverture. Le contenu génomique, notamment la teneur en GC, peut avoir une incidence sur la couverture.
 - ▶ La variation du nombre de copies peut déterminer si un variant est identifié comme étant homozygote ou hétérozygote.

- ▶ Dans certains contextes répétitifs, les variants sont filtrés dans les fichiers VCF. Le filtre de répétitions RMxN sert à filtrer les variants lorsque la séquence du variant se retrouve, en totalité ou en partie, à répétition dans le génome de référence adjacent à la position du variant. En ce qui concerne la définition des variants germinaux, il faut qu'il y ait au moins neuf répétitions dans le génome de référence pour qu'un variant soit filtré. Seules les répétitions d'une longueur d'au plus 5 pb seront prises en compte (R5x9).
 - ▶ Si un indel et un SNV sont présents pour un même locus, il se peut qu'un seul variant soit détecté.
- 13 Limites spécifiques au flux de travail somatique.
- ▶ L'instrument NextSeq 550Dx, lorsqu'il est utilisé avec le module d'analyse des variants somatiques de Local Run Manager pour NextSeq 550Dx, est conçu pour produire des résultats qualitatifs aux fins de la définition des variants somatiques (c.-à-d. présence d'un variant somatique dont la fréquence de variant est supérieure ou égale à 0,026, avec une limite de détection de 0,05).
 - ▶ Lorsque le module d'analyse des variants somatiques est utilisé, la couverture minimale par amplicon requise pour obtenir une définition de variants précise est de 450x par pool d'oligonucléotides. Par conséquent, il faut 450 fragments d'ADN pris en charge par pool d'oligonucléotides, ce qui équivaut à 900 lectures appariées se chevauchant. Le nombre d'échantillons et le nombre total de bases ciblées ont une incidence sur la couverture. Le contenu génomique, notamment la teneur en GC, peut avoir une incidence sur la couverture.
 - ▶ En ce qui concerne la définition des variants somatiques, il faut qu'il y ait au moins six répétitions dans le génome de référence pour qu'un variant soit filtré, et seules les répétitions d'une longueur d'au plus 3 pb seront prises en compte (R3x6).
 - ▶ Le module d'analyse des variants somatiques ne peut distinguer les variants germinaux des variants somatiques. Le module est conçu pour détecter les variants dans une fourchette de fréquences de variant, mais elle ne peut servir à différencier les variants somatiques et germinaux.
 - ▶ Le tissu normal dans l'échantillon a une incidence sur la détection des variants. La limite de détection indiquée est fondée sur une fréquence de variant relative à l'ADN total extrait des tissus tumoraux et normaux.

Composants du produit

- 1 Instrument NextSeq 550Dx (référence n° 20005715)
- 2 Les composants logiciels pour l'instrument NextSeq 550Dx comprennent ce qui suit :

Application logicielle	Fonction	Description
NextSeq 550Dx Operating Software (NOS - logiciel d'exploitation NextSeq 550Dx)	Commande le fonctionnement de l'instrument.	L'application logicielle NOS gère le fonctionnement de l'instrument au cours du séquençage et génère des images utilisées par le logiciel Real-Time Analysis (RTA).
Logiciel Real-time Analysis (RTA)	Effectue l'analyse primaire.	L'application logicielle RTA convertit les images générées par NOS pour chaque plaque par cycle de l'analyse de séquençage dans les fichiers de définition des bases qui constituent des entrées pour les modules d'analyse dans Local Run Manager. L'application logicielle RTA ne contient pas d'interface utilisateur.
Local Run Manager	Interface pour la sélection du module	Le logiciel Local Run Manager est une solution intégrée sur instrument, destinée à la gestion des utilisateurs, à la sélection du module d'analyse approprié et à la surveillance de l'état de progression.

Application logicielle	Fonction	Description
Module des variants somatiques	Effectue l'analyse secondaire	Ce module d'analyse de Local Run Manager traite les appels de bases par le biais d'analyses secondaires. Le processus comprend le démultiplexage, la génération de fichiers FASTQ, l'alignement, l'appel des variants et la création de rapports. L'appel des variants (Pisces) génère des fichiers VCF qui contiennent des renseignements sur les variants que l'on trouve à des positions spécifiques dans un génome de référence et comprend la fréquence de variant mesurée.
Module d'analyse des variants germinaux	Effectue l'analyse secondaire	Ce module d'analyse de Local Run Manager traite les appels de bases par le biais d'analyses secondaires. Le processus comprend le démultiplexage, la génération de fichiers FASTQ, l'alignement, l'appel des variants et la création de rapports. L'appel des variants (Pisces) génère des fichiers VCF qui contiennent des renseignements sur les variants que l'on trouve à des positions spécifiques dans un génome de référence et détermine si un variant est identifié comme étant hétérozygote ou homozygote.

Conditions d'utilisation

Élément	Caractéristique
Température	Maintenez la température du laboratoire entre 19 °C et 25 °C (22 °C ± 3 °C), soit 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 %.

Équipement et matériel

Équipement et matériel requis, vendus séparément

Trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles), référence n° 20019554

Trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles), référence n° 20028871

Équipement et matériel nécessaires, non fournis

Consommables fournis par l'utilisateur pour les analyses de séquençage

Consommable	Fournisseur	Utilisation
Tampons imbibés d'alcool isopropylique à 70 % ou Éthanol à 70 %	VWR, n° de référence 95041-714 (ou équivalent) Fournisseur de laboratoire général	Nettoyage de la Flow Cell et usage général
Tissu de laboratoire non pelucheux	VWR, référence n° 21905-026 (ou équivalent)	Nettoyage de la Flow Cell et usage général

Consommables fournis par l'utilisateur pour la maintenance de l'instrument

Consommable	Fournisseur	Utilisation
NaOCl, 5 % (hypochlorite de sodium)	Sigma-Aldrich, n° de référence 239305 (ou produit de catégorie laboratoire équivalent)	Lavage de l'instrument à l'aide de la fonction de lavage manuel après analyse; dilution à 0,12 %
Tween 20	Sigma-Aldrich, n° de référence P7949	Lavage de l'instrument à l'aide des options de lavage manuel, dilution à 0,05 %
Eau de laboratoire	Fournisseur de laboratoire général	Lavage de l'instrument (lavage manuel)
Filtre à air	illumina, référence n° 20022240	Nettoyage de l'air utilisé par l'instrument pour le refroidissement

Recommandations à propos de l'eau de laboratoire

Utilisez toujours de l'eau de laboratoire ou de l'eau désionisée pour réaliser des procédures sur l'instrument. N'utilisez jamais d'eau courante. Utilisez exclusivement les eaux qui suivent ou des eaux de qualité équivalente :

- ▶ Eau désionisée
- ▶ PW1 d'Illumina
- ▶ Eau 18 mégohms (MΩ)
- ▶ Eau Milli-Q
- ▶ Eau Super-Q
- ▶ Eau de qualité biologie moléculaire

Avertissements et précautions



ATTENTION

La loi fédérale américaine n'autorise la vente de ce dispositif que sur ordonnance ou par un médecin ou tout autre professionnel de la santé autorisé par la législation de l'État dans lequel il ou elle exerce à utiliser ou ordonner l'utilisation de cet appareil.

- 1 **Certains composants des réactifs fournis par Illumina pour utilisation avec l'instrument NextSeq 550Dx contiennent des produits chimiques potentiellement dangereux. Des risques de lésions corporelles peuvent survenir par inhalation, ingestion, contact avec la peau et contact avec les yeux. Portez un équipement de protection, y compris des lunettes, des gants et une blouse de laboratoire, adapté à l'exposition à ces risques. Traitez les réactifs usagés comme des déchets chimiques et éliminez-les conformément aux lois et règles régionales, nationales et locales en vigueur.** Pour plus de renseignements relatifs à la protection de l'environnement, à la santé et à la sécurité, consultez les fiches signalétiques (SDS) à l'adresse support.illumina.com/sds.html.
- 2 Manipulez tous les échantillons de sang comme si vous les saviez infectés du virus de l'immunodéficience humaine (VIH), du virus de l'hépatite B humain (VHB) ou d'autres agents pathogènes transmissibles par le sang (précautions universelles).
- 3 Le non-respect des procédures comme décrites peut entraîner des résultats erronés ou une baisse considérable de la qualité des échantillons.
- 4 Utilisez les précautions habituelles en laboratoire. Ne pipettez pas avec la bouche. Ne mangez pas, ne buvez pas et ne fumez pas dans les zones de travail indiquées. Portez des gants jetables et une blouse de laboratoire lors de la manipulation des échantillons et des réactifs de la trousse. Lavez-vous les mains soigneusement après avoir manipulé les échantillons et les réactifs de la trousse.
- 5 Les pratiques de laboratoire appropriées et une bonne hygiène dans le laboratoire sont nécessaires pour empêcher les produits PCR de contaminer les réactifs, les instruments et les échantillons d'ADN génomique. La contamination par des produits PCR peut causer des résultats erronés et non fiables.

- 6 Pour éviter la contamination, veillez à ce que les zones de préamplification et de postamplification aient chacune un équipement et des consommables qui leur soient réservés (p. ex., pipettes, pointes de pipette, blocs chauffants, agitateurs vortex et centrifugeuses).
- 7 La paire index-échantillon doit correspondre exactement au schéma imprimé de la plaque. Le logiciel Local Run Manager inscrit automatiquement les primers d'index associés au nom de chaque échantillon lorsque ces derniers sont entrés dans le module. On recommande à l'utilisateur de vérifier les primers d'index associés aux échantillons avant de lancer l'analyse de séquençage. Toute divergence entre le schéma de l'échantillon et de la plaque entraînera une perte de l'identification positive des échantillons et un rapport de résultats erroné.
- 8 Il est vivement recommandé d'installer un logiciel antivirus fourni par l'utilisateur afin de protéger l'ordinateur des virus. Consultez le manuel de l'utilisateur pour obtenir des instructions relatives à l'installation.
- 9 N'utilisez pas l'instrument NextSeq 550Dx si l'un des panneaux a été retiré. L'utilisation de l'instrument, lorsque l'un des panneaux a été retiré, crée un risque d'exposition à la tension d'alimentation et à plusieurs tensions continues.
- 10 Ne touchez pas la platine de la Flow Cell dans le compartiment de Flow Cell. Le réchauffeur dans ce compartiment fonctionne entre 22 °C et 95 °C et peut causer des brûlures.
- 11 L'instrument pèse environ 185 livres et peut causer des blessures graves s'il tombe ou s'il est manipulé sans précaution.

Mode d'emploi

Le mode d'emploi qui suit sert à l'exécution des modules d'analyse des variants germinaux et somatiques en mode diagnostic sur l'instrument NextSeq 550Dx avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) ou la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).

Saisie des renseignements sur l'analyse

Pour obtenir des instructions détaillées, reportez-vous au Guide de référence de l'instrument NextSeq 550Dx (document n° 100000009513) et au Guide du module de Local Run Manager applicable.

Définir les paramètres

- 1 Connectez-vous à Local Run Manager.
- 2 Sélectionnez **Create Run** (Créer l'analyse), puis sélectionnez **Somatic Variant** (Variant somatique) ou **Germline Variant** (Variant germlinal).
- 3 Entrez un nom qui désignera l'analyse, du séquençage jusqu'à la fin du processus.
Utilisez des caractères alphanumériques, des espaces, des tirets de soulignement ou des tirets.
- 4 **[Facultatif]** Saisissez une description pour identifier l'analyse.
Utilisez des caractères alphanumériques, des espaces, des tirets de soulignement ou des tirets.
- 5 Choisissez le nombre d'échantillons et l'ensemble d'index dans la liste déroulante.
Tenez compte de ce qui suit lorsque vous faites votre sélection.
 - ▶ La liste déroulante contient le nombre d'échantillons et l'ensemble d'index. Par exemple, « 24-Set 1 » indique qu'il y a 24 échantillons à tester avec les échantillons de l'ensemble d'index 1.
 - ▶ Le numéro de l'ensemble d'index fait référence aux différents ensembles des paires d'index i5 et i7. L'ensemble 1 et l'ensemble 2 assurent la diversité des index. Deux ensembles d'index sont offerts pour aider à prévenir la déplétion d'un ensemble.
 - ▶ Choisissez le nombre d'échantillons le plus proche du nombre d'échantillons testés. Si le nombre exact d'échantillons ne figure pas dans la liste, sélectionnez le nombre le plus proche inférieur au nombre d'échantillons testés. Par exemple, si vous souhaitez tester 18 échantillons, sélectionnez 16 échantillons.
 - ▶ Les suggestions de puits d'échantillon et de combinaisons d'index qui répondent aux exigences de diversité sont surlignées en vert.

Importation des fichiers de manifeste à analyser

- 1 Assurez-vous que les fichiers de manifeste à importer se trouvent dans un emplacement réseau accessible ou sur une clé USB.
- 2 Sélectionnez **Import Manifests** (Importer les fichiers de manifeste).
- 3 Accédez aux fichiers de manifeste et sélectionnez les fichiers à ajouter.



REMARQUE


Pour rendre les fichiers de manifeste accessibles pour toutes les analyses réalisées au moyen du module d'analyse des variants germinaux ou du module d'analyse des variants somatiques, ajoutez les fichiers à l'aide de la fonction Module Settings (Paramètres du module). Cette fonction exige un niveau d'accès administrateur. Pour plus de renseignements, consultez le *Guide de référence de l'instrument NextSeq 550Dx (document n° 100000009513)*.

Sélection des échantillons à analyser

Précisez les échantillons à analyser en utilisant l'une des options suivantes, conformément aux directives qui suivent.


- ▶ **Saisie manuelle des échantillons** : Utilisez le tableau vide à l'écran Create Run (Créer l'analyse).
- ▶ **Importer des échantillons** : naviguer vers un fichier dont les valeurs sont séparées par des virgules (*.csv) externe. Il est possible de télécharger un modèle sur l'écran Create Run (Créer l'analyse).

Saisie manuelle des échantillons

- 1 Entrez un nom d'échantillon unique (**module d'analyse des variants somatiques**) ou l'identifiant de l'échantillon (**module d'analyse des variants germinaux**).
Utilisez des caractères alphanumériques, des tirets ou des traits de soulignement.
- 2 **[Facultatif]** Pour les échantillons de contrôle positif ou négatif, faites un clic droit et sélectionnez le type de contrôle.
La saisie du contrôle dans l'un des puits d'échantillon inscrit automatiquement le même contrôle dans le puits correspondant de l'autre pool.
- 3 **[Facultatif]** Entrez la description de l'échantillon dans le champ Description.
Utilisez des caractères alphanumériques, des tirets ou des traits de soulignement.
- 4 Sélectionnez l'adaptateur d'index 1 dans la liste déroulante Index 1 (i7).
Lorsque vous utilisez les puits d'échantillons suggérés, le logiciel inscrit automatiquement les adaptateurs d'index i7 et i5 qui satisfont aux exigences de diversité d'index. Si le nombre exact d'échantillons que vous testez ne figure pas dans la liste, assurez-vous de sélectionner les adaptateurs d'index pour les puits supplémentaires.
- 5 Sélectionnez l'adaptateur d'index 2 dans la liste déroulante Index 2 (i5).
- 6 Sélectionnez un fichier de manifeste dans la liste déroulante Manifest (Manifeste).
Les échantillons du pool A requièrent un fichier de manifeste différent de celui des échantillons du pool B.
- 7 Choisissez l'une des options pour afficher, imprimer ou enregistrer le schéma de la plaque à titre de référence pour la préparation des librairies :
 - Cliquez sur l'icône  **Print** (Imprimer) pour afficher la présentation de la plaque. Sélectionnez **Print** (Imprimer) pour imprimer la présentation de la plaque.
 - Sélectionnez **Export** (Exporter) pour exporter l'information sur les échantillons dans un fichier externe.
- 8 Sélectionnez **Save Run** (Enregistrer l'analyse).

Importation des échantillons

- 1 Sélectionnez **Import Samples** (Importer des échantillons) et rendez-vous à l'emplacement du fichier contenant les informations des échantillons. Il est possible d'importer deux types de fichiers.

- Cliquez sur **Template** (Modèle) à l'écran Create Run (Créer l'analyse) pour établir une nouvelle présentation de plaque. Le fichier de modèle contient les en-têtes de colonnes corrects pour l'importation. Saisissez les informations des échantillons de l'analyse dans chaque colonne. Supprimez l'information donnée en exemple dans les cellules inutilisées, puis enregistrez le fichier.
 - Utilisez un fichier d'information sur les échantillons qui a été exporté du module d'analyse des variants somatiques ou germinaux en utilisant la fonction Export (Exporter).
- 2 Cliquez sur l'icône  **Print** (Imprimer) pour afficher la présentation de la plaque.
 - 3 Sélectionnez **Print** (Imprimer) pour imprimer le schéma de la plaque comme référence pour la préparation des librairies.
 - 4 Sélectionnez **Save Run** (Enregistrer l'analyse).

Préparer la cartouche de réactifs

Assurez-vous de suivre attentivement les directives applicables à la cartouche de réactifs pour un séquençage réussi.

- 1 Retirez la cartouche de réactifs de son lieu de stockage entre -25 et -15 °C.
- 2 Choisissez l'une des méthodes suivantes pour décongeler les réactifs. N'immergez pas la cartouche. Après la décongélation de la cartouche, asséchez-la avant de passer à l'étape suivante.

Température	Durée de décongélation	Période de stabilité
Bain d'eau entre 15 et 30 °C	60 minutes	Ne pas dépasser 6 heures
De 2 à 8 °C	7 heures	Ne pas dépasser 5 jours



REMARQUE

Si plus d'une cartouche est décongelée dans le même bain d'eau, prolongez le temps de décongélation.

- 3 Retournez la cartouche cinq fois pour mélanger les réactifs.
- 4 Inspectez le dessous de la cartouche afin de vous assurer que les réactifs sont décongelés et ne contiennent pas de précipités. Confirmez que les positions 29, 30, 31 et 32 sont décongelées; ce sont les plus grosses et elles prendront plus de temps à décongeler.
- 5 Tapotez doucement sur la paille pour réduire les bulles d'air.
Pour obtenir de meilleurs résultats, chargez directement l'échantillon et configurez l'analyse.

Préparer la Flow Cell

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

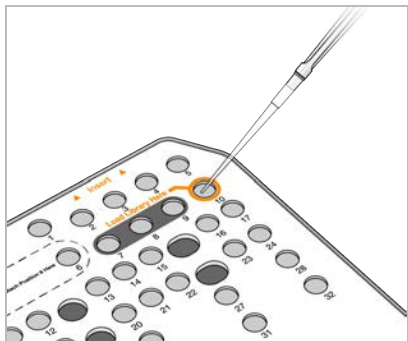
Préparer des librairies pour le séquençage

Dénaturez et diluez les librairies à un volume de chargement de 1,3 ml. En pratique, la concentration de chargement peut varier selon les méthodes de préparation et de quantification des librairies. La dilution des librairies d'échantillons dépend de la complexité des pools d'oligonucléotides. Pour savoir comment préparer les librairies d'échantillons pour le séquençage, y compris la dilution et le regroupement de librairies, consultez la section Mode d'emploi sur la trousse de préparation de librairies applicable. L'optimisation de la densité des amplifiats sur NextSeq 550Dx est requise.

Chargement des librairies sur la cartouche de réactifs

- 1 Nettoyez l'opercule en aluminium recouvrant le réservoir n° 10 étiqueté **Load Library Here** (Charger la librairie ici) à l'aide d'un tissu non pelucheux.
- 2 Percez l'opercule avec une pointe de pipette propre de 1 ml.
- 3 Chargez 1,3 ml de librairie préparée dans le réservoir n° 10 étiqueté **Load Library Here** (Charger la librairie ici). Évitez de toucher l'opercule en aluminium pendant le transfert du produit.

Figure 1 Charger les bibliothèques



Configurer une analyse de séquençage

- 1 Ouvrez une session dans NextSeq 550Dx en entrant votre mot de passe du logiciel Local Run Manager.
- 2 À l'écran d'accueil du logiciel NOS, sélectionnez **Sequence** (Séquencer).
- 3 Sélectionnez une analyse dans la liste, puis cliquez sur **Next** (Suivant).
Une série d'écrans de configuration de l'analyse s'ouvre dans l'ordre suivant : Load Flow Cell (Charger la Flow Cell), Load Buffer Cartridge (Charger la cartouche de tampon), Load Reagents Cartridge (Charger la cartouche de réactifs) et Pre-Run Check (Vérification avant analyse).
- 4 Lorsque l'écran Load Flow Cell (Charger la Flow Cell) s'affiche, nettoyez et chargez la Flow Cell.
 - ▶ Sortez la Flow Cell de son emballage en aluminium.
 - ▶ Ouvrez l'emballage double coque en plastique transparent et sortez la Flow Cell.
 - ▶ Nettoyez la surface en verre de la Flow Cell à l'aide d'une lingette alcoolisée non pelucheuse. Séchez le verre à l'aide d'un chiffon de laboratoire peu pelucheux.
 - ▶ Assurez-vous que la surface de verre de la Flow Cell est propre. Au besoin, répétez les étapes de nettoyage.
 - ▶ Retirez la Flow Cell de l'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 activés.
- 6 Suivez les indications du logiciel pour vider le réservoir à réactifs usagés, charger la cartouche de tampon NextSeq 550Dx et charger la cartouche de réactifs NextSeq 550Dx.
Lorsque les cartouches de tampon et de réactifs NextSeq 550Dx sont chargées, le logiciel lit et enregistre la lecture RFID. Les identifiants des cartouches de tampon et de réactifs s'affichent à l'écran et les capteurs sont vérifiés.
- 7 Une fois que la vérification automatique avant analyse est terminée, sélectionnez **Start** (Démarrer). (étape non requise si le système est configuré pour démarrer automatiquement).
- 8 L'écran Sequencing (Séquençage) s'ouvre lorsque l'analyse démarre. Cet écran fournit une représentation visuelle de l'analyse en cours, y compris les intensités et les scores de qualité.

Résultats

Le logiciel intégré Real-Time Analysis (RTA) exécute les analyses d'images et la définition des bases et affecte un score de qualité à chacune des bases de chaque cycle de séquençage. Une fois l'analyse primaire terminée, le module de Local Run Manager qui a été choisi dans l'instrument NextSeq 550Dx commence automatiquement l'analyse secondaire. Les processus d'analyse secondaire décrits ici concernent les modules d'analyse des variants germinaux et somatiques.

Démultiplexage

Le démultiplexage compare chaque séquence de lecture d'index aux séquences d'indexage définies pour l'analyse. 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 l'analyse.
- ▶ 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 un seul mésappariement par lecture d'index.

Génération des fichiers FASTQ

Après le démultiplexage, le logiciel génère une analyse intermédiaire au format FASTQ, qui est un format texte utilisé pour représenter des séquences. Les fichiers FASTQ contiennent les lectures pour chaque échantillon ainsi que les scores de qualité connexes. 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 compris dans le nom du fichier FASTQ. Dans les modules d'analyse des variants germinaux et somatiques, huit fichiers FASTQ sont générés par échantillon, pour chaque pool d'oligos; quatre proviennent de la lecture 1 et quatre de la lecture 2. Ces résultats donnent respectivement un total de 8 et de 16 fichiers FASTQ par échantillon pour les variants germinaux et somatiques. Les fichiers FASTQ constituent les principales données d'entrée pour l'alignement.

Alignement

Au cours de l'étape de l'alignement, l'algorithme de Smith-Waterman par bande aligne les amplifiats provenant de chacun des échantillons sur les séquences d'amplicon définies dans le fichier de manifeste.

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

Chaque lecture appariée est évaluée en fonction de son alignement sur les séquences de sonde correspondantes.

- ▶ La lecture 1 est évaluée par rapport au complément inverse des oligos spécifiques aux loci en aval (DLSO).
- ▶ La lecture 2 est évaluée par rapport aux oligos spécifiques aux loci en amont (ULSO).
- ▶ Si le début d'une lecture correspond à une séquence de sonde ne présentant pas plus d'un mésappariement, la longueur totale de la lecture est alignée sur l'amplicon cible pour cette séquence.
- ▶ Si le début d'une séquence de lecture correspond à une séquence de sonde ne présentant pas plus de trois différences (mésappariements ou décalages attribuables aux principaux indels), la longueur totale de la lecture est alors alignée sur la séquence d'amplicon cible pour cette séquence.
- ▶ Les indels dans les DLSO et ULSO ne sont pas observés étant donné la chimie du test.

Les alignements sont filtrés en fonction des résultats de l'alignement basés sur les taux de mésappariements par rapport à la région d'intérêt ou à l'amplicon complet, selon la longueur de l'amplicon. Les alignements filtrés sont notés dans les fichiers d'alignement comme étant non alignés et ne sont pas utilisés pour l'appel des variants.

Appel des variants

Le paramètre d'appel des variants *Pisces* est conçu pour générer des appels de SNV et de variants d'indels provenant des bibliothèques préparées pour l'instrument.

Rapports et fichiers de sortie supplémentaires

Les modules d'analyse des variants produisent des rapports en format PDF et sous la forme de fichiers de texte séparé par des tabulations (*.txt) qui contiennent les indicateurs, comme la profondeur de séquençage et le nombre de variants. Les modules produisent aussi des fichiers de sortie, comme des fichiers VCF et des fichiers gVCF (genome Variant Call Format).

Procédures de contrôle qualité

Le logiciel NextSeq 550Dx évalue chaque analyse, échantillon et définition des bases par rapport aux indicateurs 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, alors il y a peut-être eu une erreur dans le suivi de l'échantillon, un enregistrement incorrect des primers d'indexage 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 y a peut-être eu une erreur dans le suivi de l'échantillon ou un enregistrement incorrect des primers d'indexage.

Caractéristiques de performance

Les caractéristiques de performance de l'instrument NextSeq 550Dx ont été établies en utilisant les modules d'analyse des variants germinaux et somatiques, de la trousse d'amplicons personnalisés TruSeq Dx et de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) et ont été confirmées en utilisant la trousse de réactifs NextSeq 550Dx à débit élevé 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 la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) ont été conçues pour évaluer les allégations de performance précédemment établies avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles). Les résultats montrent que les trousse de réactifs (v2 et v2.5) ont des performances comparables en utilisant la trousse d'amplicons personnalisés TruSeq Dx. Consultez la *notice d'accompagnement de la trousse d'amplicons personnalisés TruSeq* pour connaître 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 performance

- 1 La concordance positive en pourcentage (CPP) correspond à la proportion de loci classifiés comme étant des variants par une méthode de référence que le test a correctement analysés.
 - ▶ $(\text{nombre de loci variants correctement analysés par le test}) / (\text{nombre total de loci variants})$
Les loci variants analysés par le test et qui concordent avec la méthode de référence sont des vrais positifs (TP). Les loci variants analysés comme des appels de référence ou des appels de variant différents par le test sont des faux négatifs (FN).
- 2 La concordance négative en pourcentage (CNP) est calculée comme correspondant à la proportion de loci classifiés comme étant de type sauvage par une méthode de référence que le test a correctement analysés.
 - ▶ $(\text{nombre de loci de type sauvage correctement analysés par le test}) / (\text{nombre total de loci de type sauvage})$
Les loci de type sauvage analysés par le test et qui concordent avec la méthode de référence sont des vrais négatifs (TN). Les loci de type sauvage analysés comme étant des variants par le test sont des faux positifs (FP).
- 3 Le pourcentage global de concordance (PGC) correspond à la proportion de loci correctement analysés par le test relativement à une méthode de référence.

- ▶ $((\text{nombre de loci variants correctement analysés par le test}) + (\text{nombre de loci de type sauvage correctement analysés par le test})) / ((\text{nombre total de loci variants}) + (\text{nombre total de loci de type sauvage}))$
- 4 Les calculs de la CPP, de la CNP et du PGC ne comprennent pas les absences d'appels (loci variants ou loci de référence qui ne passent pas un ou plusieurs des filtres de qualité).
- 5 Le débit d'appel autosome est calculé comme correspondant au nombre total de loci qui passent les filtres, divisé par le nombre total de positions séquencées pour les chromosomes 1 à 22; les chromosomes X et Y étant exclus. Ce paramètre ne tient pas compte de la concordance des appels avec la méthode de référence.

Performance de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles)

Indexage de l'échantillon

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'indexage de l'échantillon est utilisé pour les flux de travail germinale et somatique. L'objectif de cette étude était d'établir le nombre d'échantillons minimal (8) et maximal (96) pouvant être traités lors d'une seule analyse de séquençage par l'instrument NextSeq 550Dx. Huit échantillons uniques de Platinum Genome ont été analysés avec 12 combinaisons différentes de primers d'indexage par échantillon. Les résultats d'analyse des échantillons provenant de quatre analyses de séquençage utilisant le module d'analyse des variants germinaux ont été comparés aux données de Platinum Genomes version 2016-1.0.

Pour la première série d'analyses, 96 bibliothèques d'échantillons à index unique ont été analysées à l'aide d'un test représentatif conçu pour étudier une variété de gènes couvrant 12 588 bases par brin sur les 23 chromosomes humains pour vérifier la capacité du test à faire systématiquement un appel de génotypage pour un échantillon donné sur différentes combinaisons de primers d'indexage. Pour la deuxième série d'analyses, huit bibliothèques d'échantillons à index unique ont été séquencées en deux analyses de séquençage afin de vérifier le nombre minimal d'index pris en charge.

Dans le cas des analyses de 96 index, la CPP pour les SNV allait de 98,7 % à 100 %, la CPP pour les insertions et les délétions était de 100 % et la CNP était de 100 % pour chacune des 96 combinaisons d'index. Les analyses de huit index avaient une CPP de 100 % (SNV, insertions et délétions) et une CNP de 100 % pour chacune des huit combinaisons d'index.

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 mâle et un femelle, 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 d'analyse des variants germinaux. La contamination par transfert des échantillons mâles aux échantillons femelles a été observée par la présence de lectures d'amplicons du chromosome Y dans les échantillons femelles.

La contamination par transfert entre échantillons réalisée 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 la contamination par transfert entre échantillons au cours d'une analyse de séquençage, un regroupement de bibliothèques composé de 46 réplicats d'échantillons mâles et femelles et de quatre contrôles négatifs a été séquencé une fois sur l'instrument NextSeq 550Dx. La contamination par transfert entre échantillons au cours d'une analyse de séquençage a été évaluée par la comparaison de la couverture des amplicons du chromosome Y de chaque réplicat femelle avec la couverture moyenne des amplicons du chromosome Y de tous les réplicats mâles du regroupement. La moyenne observée de contamination par transfert entre échantillons au cours d'une analyse de séquençage était de 0,084 %.

Pour tester la contamination par transfert entre échantillons entre les analyses de séquençage, deux regroupements de bibliothèques ont été préparés et séquencés de façon consécutive sur l'instrument NextSeq 550Dx. Le premier regroupement contenait 46 réplicats d'échantillons femelles, plus deux contrôles négatifs. Le deuxième regroupement contenait 46 réplicats d'échantillons mâles, plus deux contrôles négatifs. 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 entre les analyses de séquençage a été évaluée par la comparaison de la couverture des amplicons du chromosome Y entre les réplicats correspondant de la deuxième analyse de séquençage du regroupement des échantillons femelles et de l'analyse de séquençage du regroupement des échantillons mâles. La moyenne observée de contamination par transfert entre échantillons entre les analyses de séquençage était de 0,0076 %.

Entrée d'ADN

Sang (variant germinal)

La plage d'entrées d'ADN provenant d'échantillons de sang pour la préparation des bibliothèques de la trousse d'amplicons personnalisés TruSeq Dx au moyen du flux de travail du module d'analyse des variants germinaux a été établie pour l'instrument NextSeq 550Dx. Cette plage a été évaluée en effectuant une étude de dilution en série au moyen de 13 échantillons de Platinum Genome et d'un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases sur 23 chromosomes différents. La bibliothèque a été séquencée sur deux instruments NextSeq 550Dx avec un lot de la trousse de réactifs NextSeq 550Dx à débit élevé 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 dans un seul réplicat 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 concordance positive en pour cent (CPP) pour chaque type de variant (SNV, insertions et délétions) est présentée dans le [Tableau 1](#); la concordance négative en pour cent (CNP) est présentée dans le [Tableau 2](#). Tous les niveaux d'entrée ont affiché une précision similaire. Le niveau d'entrée d'ADN recommandé pour la trousse d'amplicons personnalisés TruSeq Dx est 50 ng, les niveaux 25 ng et 100 ng représentant les limites inférieures et supérieures permettant de respecter les caractéristiques de performance.

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

Entrée d'ADN (ng)	Type de variant	Variants prévus	TP	FN	Absence d'appel de variant	CPP (%)
12	SNV	2412	2381	31	0	98,7
25			2404	8	0	99,7
50			2403	9	0	99,6
100			2412	0	0	100
250			2412	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 CNP pour chaque niveau d'entrée d'ADN

Entrée d'ADN (ng)	TN	FP	Absence d'appel dans la référence	CNP (%)
12	430940	4	26	> 99,9
25	430936	0	34	100
50	430936	2	32	> 99,9
100	430942	0	28	100
250	430942	0	28	100

FFPE (variant somatique)

La plage d'entrées d'ADN provenant de tissus fixés au formol et imprégnés à la paraffine (FFPE) pour la préparation des bibliothèques de la trousse d'amplicons personnalisés TruSeq Dx au moyen du flux de travail du module d'analyse des variants somatiques a été établie pour l'instrument NextSeq 550Dx. La plage d'entrées d'ADN a été évaluée en effectuant une étude de dilution en série au moyen de trois échantillons de Platinum Genome et d'un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases sur 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. La lignée GM12878 a été diluée avec la lignée GM12877, de sorte que les fréquences alléliques des variants (FAV) de 81 variants (55 SNV, 10 insertions et 16 délétions) se situent près de 0,025, 0,05 ou 0,10. De plus, chaque échantillon comptait 91 variants dont la FAV était supérieure et pouvait aller jusqu'à 1,0. Les échantillons ont été traités en double à cinq niveaux d'entrée d'ADN dont la moyenne du cycle de quantification delta (dCq) était de 2,1, 3,6, 4,6, 6,0 et 7,8, selon la mesure de la trousse d'amplicons personnalisés TruSeq Dx – FFPE CQ. Chaque bibliothèque a été séquencée sur deux instruments NextSeq 550Dx avec deux lots de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles). Pour la détermination de la précision, les appels de variant des échantillons ont été comparés aux données de Platinum Genomes version 2016-1.0. La CPP pour chaque type de variant (SNV, insertions et délétions) est présentée dans le [Tableau 3](#); la CNP est présentée dans le [Tableau 4](#). La recommandation en matière d'entrée d'ADN pour les variants dont la FAV est de 0,05 ou plus correspond à un dCq inférieur ou égal à 4; 4,6 procurant une limite inférieure permettant de respecter les caractéristiques de performance.

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

dCq moyen	Type de variant	Variants prévus	Absences d'appel prévues	Dilution cible – FAV					
				0,025		0,05		0,10	
				Absence d'appel de variant	CPP (%)	Absence d'appel de variant	CPP (%)	Absence d'appel de variant	CPP (%)
2,1	SNV	808	Sans objet.	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 CNP pour chaque niveau d'entrée d'ADN

dCq moyen	Type sauvage attendu	Dilution cible – FAV					
		0,025		0,05		0,10	
		Absence d'appel dans la référence	CNP (%)	Absence d'appel dans la référence	CNP (%)	Absence d'appel dans la référence	CNP (%)
2,1	93688	344	100	260	100	324	100
3,6		400	100	332	100	380	100
4,6		1308	100	1336	100	784	100
6,0		3900	> 99,9	3296	> 99,9	2996	100
7,8		3020	> 99,9	2880	> 99,9	2448	> 99,9

Sensibilité analytique (limite de blanc [LB] et limite de détection [LD])

Cette étude a été menée afin d'évaluer la limite de blanc (LB) et la limite de détection (LD) pour le module d'analyse des variants somatiques sur l'instrument NextSeq 550Dx. L'étude a été réalisée avec un test représentatif conçu pour étudier divers gènes couvrant 12 588 bases sur 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. GM12878 a été dilué avec GM12877 de façon à ce que la fréquence de 74 variants (53 SNV, sept insertions et 14 délétions) soit de $0,05 \pm 0,02$. GM12877 et GM12878 dilué (GM12878-D) ont été testés au cours de six journées consécutives avec un seul instrument, en alternance entre deux lots de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles), pour un total de six analyses de séquençage. Il en a résulté 60 réplicats pour chaque variant de GM12878-D et 72 réplicats pour chaque coordonnée correspondante de type sauvage de GM12877, pour chaque lot de réactifs. La LB et la LD ont été calculées au moyen des méthodes habituelles énoncées dans la norme CLSI EP17-A2 en utilisant l'option non paramétrique. La LB et la LD ont été calculées séparément pour les SNV, les insertions et les délétions en regroupant les fréquences de variants pour chaque type de variant. L'erreur de type I a été définie à 0,01 et l'erreur de type II, à 0,05.

Pour la LB, les fréquences de variants regroupées ont été classées de la moins élevée à la plus élevée, et la position arrivant au 99e rang a été calculée pour chaque lot de réactifs, pour chaque type de variant (Tableau 5). Le module d'analyse des variants somatiques utilise une limite (la LB appliquée) de 0,026 pour la fréquence allélique des variants (FAV) afin de déterminer la détection qualitative des variants. La LB calculée servait à vérifier que cette limite n'entraîne que des erreurs de type I ne dépassant pas 0,01.

Tableau 5 Limite de blanc

Type de variant	Observations totales	LB, Lot de réactifs 1 (en %)	LB, Lot de réactifs 2 (en %)
SNV	3816	0,77	0,77
Insertion	504	0,56	0,56
Délétion	1008	1,20	1,20

Pour la LD, le pourcentage de la fréquence de chaque mutation pour chaque lot de réactifs et pour chaque type de variant tombant sous la limite de 0,026 a été calculé (Tableau 6). Puisque les pourcentages étaient inférieurs à ceux des erreurs de type II, soit 5 % (0,05), la médiane des fréquences de variants combinées a été établie

comme LD (**Tableau 6**). La LD de chaque type de variant a été utilisée comme la plus élevée des deux valeurs calculées 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 de la FAV < 2,6 %	% de mesures de la FAV < 2,6 %	Limite de détection (%)
1	SNV	3180	53	1,7	4,94
	Insertion	420	6	1,4	5,08
	Délétion	840	7	0,8	5,22
2	SNV	3180	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 d'appel de variant du module des variants germinaux sur l'instrument NextSeq 550Dx à l'aide de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles). 13 échantillons uniques de Platinum Genome 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. La précision a été déterminée pour les SNV, les insertions et les délétions par la comparaison des résultats avec ceux de la méthode de référence composite bien caractérisée, Platinum Genomes version 2016-1.0. Les régions génomiques de fiabilité ont été définies sur la base de cette méthode de référence, à moins d'indication contraire.

Tableau 7 Sommaire de la concordance germinale

Critères	Observations totales ¹	Résultat en fonction de l'observation ²	Résultat en fonction de l'analyse ³
CPP pour les SNV	819	98,7	> 99,9
CPP pour les insertions	819	95,0	98,9
CPP pour les délétions	819	100	100
CNP	819	100	100
PGC	819	> 99,9	> 99,9

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

² Valeur la plus faible observée par réplicat d'échantillon lors des neuf analyses.

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

Le **Tableau 8** montre les données de l'étude présentées avec la concordance positive et négative en pour cent, par échantillon, où les résultats des variants sont comparés aux données de Platinum Genomes version 2016-1.0 pour le calcul de la concordance positive en pour cent (CPP). Les trois types de variants (SNV, insertions et délétions) sont combinés. Étant donné que la méthode de référence ne donne des résultats que pour les variants mononucléotides et les insertions/délétions, les résultats des bases sans variants sont comparés à la séquence de référence du génome humain version hg19 pour le calcul de la concordance négative en pour cent (CNP).

Tableau 8 Concordance des variants germinaux, par échantillon

Échantillon	Taux d'appel moyen	Variants prévus ¹	TP	FN	Absence d'appel de variant	TN	FP	CPP	CNP	PGC
NA12877	> 99,9	4788	4788	0	0	756762	0	100	100	100
NA12878	> 99,9	8505	8379	1	125	751464	0	> 99,9	100	> 99,9
NA12879	> 99,9	6048	5985	5	58	757701	0	99,9	100	> 99,9
NA12880	> 99,9	6993	6930	0	63	757638	0	100	100	100
NA12881	> 99,9	7875	7811	3	61	751653	0	> 99,9	100	> 99,9
NA12882	> 99,9	6300	6174	3	123	754803	0	> 99,9	100	> 99,9
NA12883	> 99,9	7119	7056	0	63	751905	0	100	100	100
NA12884	> 99,9	7182	7119	6	57	754146	0	99,9	100	> 99,9
NA12885	> 99,9	7686	7560	2	124	754173	0	> 99,9	100	> 99,9
NA12886	> 99,9	7245	7182	7	56	752469	0	99,9	100	> 99,9
NA12887	> 99,9	7119	7119	0	0	750645	0	100	100	100
NA12888	> 99,9	6804	6804	0	0	756065	0	100	100	100
NA12893	> 99,9	7434	7371	1	62	750015	0	> 99,9	100	> 99,9

¹ Nombre total de variants dans tous les réplicats d'échantillons des neuf analyses.

Le **Tableau 9** montre les données de l'étude présentées par échantillon, où les résultats des variants sont comparés avec la méthode de référence composite bien caractérisée. La détection est évaluée séparément pour chaque type de variant (SNV, insertions et délétions). 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	2331	2331	0	1323	1323	0	1134	1134	0
NA12878	5733	5733	0	1260	1197	1	1512	1449	0
NA12879	3591	3591	0	1323	1260	5	1134	1134	0
NA12880	4221	4221	0	1512	1512	0	1260	1197	0
NA12881	4914	4913	1	1512	1449	2	1449	1449	0
NA12882	3717	3717	0	1386	1323	3	1197	1134	0
NA12883	4284	4284	0	1449	1449	0	1386	1323	0
NA12884	4284	4284	0	1575	1512	6	1323	1323	0
NA12885	4725	4725	0	1575	1512	2	1386	1323	0
NA12886	4347	4347	0	1449	1386	7	1449	1449	0
NA12887	4284	4284	0	1323	1323	0	1512	1512	0
NA12888	4158	4158	0	1449	1449	0	1197	1197	0
NA12893	4599	4599	0	1386	1323	1	1449	1449	0

Les échantillons ont également été analysés à la recherche de petites insertions et délétions (indels).

Un sommaire global est présenté au **Tableau 10**. Il y avait au total 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 Sommaire de la détection des indels germinaux

Type de variant	Variants prévus	TP	FN	Absence d'appel de variant	CPP
Insertion	18522	18018	27	477	99,9
Délétion	17388	17073	0	315	100

Le test représentatif comprenait 150 amplicons conçus pour couvrir un contenu génomique varié. La teneur en GC des amplicons allait de 0,19 à 0,87. Les amplicons comprenaient également diverses répétitions de mononucléotides (p. ex., PolyA, PolyT), de dinucléotides et de trinucléotides. Les données ont été compilées par amplicon (Tableau 11) de façon à déterminer l'effet du contenu génomique sur le pourcentage d'appels exacts. Le pourcentage d'appels exacts se rapporte aux appels de variants et aux appels de référence et est inférieur à 100 % en cas d'appels inexacts ou d'absences d'appel.

Tableau 11 Variants germinaux – Précision au niveau de l'amplicon

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
1	1	36450499	36450591	93	93	Indel	0,22	76167	0	0	100
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	64701	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	74529	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	75348	0	0	100
5	1	228526602	228526682	81	81	Poly G (5)	0,69	66339	0	0	100
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	57330	0	0	100
7	1	247812041	247812128	88	88	Poly A (5), CT(3), TAA(3), indel	0,27	72072	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	73710	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	65520	0	0	100
10	2	177016721	177016805	85	81	S. O.	0,65	66339	0	0	100
11	2	186625727	186625801	75	75	Poly A (8)	0,35	61425	0	0	100
12	2	190323504	190323591	88	88	Poly T (5)	0,42	72072	0	0	100
13	2	200796740	200796826	87	87	Poly T (5), indel	0,31	71253	0	0	100
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	74529	0	0	100
15	2	228147052	228147144	93	93	Indel	0,43	76167	0	0	100
16	2	235016350	235016422	73	73	Poly T (5), indel	0,42	59787	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	74823	0	1344	98,2
18	3	46620561	46620643	83	83	S. O.	0,43	67977	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	57330	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG (3)	0,41	72072	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	60543	0	63	99,9
22	4	2233667	2233744	78	78	Poly A (6)	0,26	63882	0	0	100
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	79443	0	0	100
24	4	15688604	15688681	78	78	S. O.	0,29	63882	0	0	100
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	50778	0	0	100
26	4	102839244	102839314	71	69	Poly A (5)	0,46	56511	0	0	100
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	50778	0	0	100
28	5	1882081	1882158	78	75	S. O.	0,78	61425	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	68796	0	0	100
30	5	41069808	41069871	64	64	S. O.	0,39	52416	0	0	100
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	67977	0	0	100
32	5	147475343	147475409	67	67	Poly T (5)	0,37	54873	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	74529	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	61425	0	0	100
35	6	6318713	6318814	102	102	Poly G (6)	0,68	83538	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	75348	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	76608	0	378	99,5
38	6	32147987	32148084	98	98	Poly T (5), TCT(3), CTT(3)	0,55	80262	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	77805	0	0	100
40	6	33408498	33408583	86	86	Poly C (6)	0,7	70434	0	0	100
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	76986	0	0	100
42	6	112435865	112435955	91	91	Poly A (5)	0,44	74529	0	0	100
43	7	22202076	22202148	73	73	S. O.	0,44	59787	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
44	7	66276100	66276187	88	88	Indel	0,35	72072	0	0	100
45	7	77365735	77365821	87	87	Poly A (7), AG(4)	0,26	71253	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	69615	0	0	100
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	73710	0	0	100
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	74529	0	0	100
49	7	154404519	154404599	81	66	S. O.	0,31	54054	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	76167	0	0	100
51	8	1817312	1817394	83	83	S. O.	0,42	67977	0	0	100
52	8	24811020	24811109	90	89	Poly G (7), CTC(4), indel	0,61	72171	0	720	99,0
53	8	76518625	76518691	67	67	Indel	0,3	54873	0	0	100
54	9	103054909	103055006	98	98	Poly G (6)	0,67	80262	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	53235	0	0	100
56	9	107620823	107620918	96	96	S. O.	0,49	78624	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	67977	0	0	100
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	79443	0	0	100
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	63882	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	74529	0	0	100
61	10	27317777	27317855	79	79	Poly T (5)	0,3	64701	0	0	100
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	73710	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	77805	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	71747	0	325	99,5
65	10	101611250	101611329	80	80	S. O.	0,49	65520	0	0	100
66	10	118351373	118351453	81	81	S. O.	0,51	66339	0	0	100
67	11	8159816	8159912	97	96	S. O.	0,45	78624	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	57330	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
69	11	47470345	47470444	100	100	S. O.	0,65	81900	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	50778	0	0	100
71	11	64418856	64418957	102	102	S. O.	0,59	83538	0	0	100
72	11	93529612	93529684	73	73	Poly A (5)	0,4	59787	0	0	100
73	11	101347052	101347136	85	85	S. O.	0,42	69615	0	0	100
74	11	102477336	102477426	91	91	Poly G (6)	0,55	74529	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	69615	0	0	100
76	11	120357801	120357885	85	85	Poly A (5), CA(3), indel	0,34	69615	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	69615	0	0	100
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	68796	0	0	100
79	12	26811004	26811096	93	93	Poly A (7), AC(4)	0,33	76167	0	0	100
80	12	30881766	30881846	81	81	S. O.	0,49	66339	0	0	100
81	12	88474105	88474175	71	71	Poly A (6)	0,35	58149	0	0	100
82	12	120966872	120966966	95	95	Poly G (5)	0,68	77805	0	0	100
83	13	24167504	24167576	73	73	S. O.	0,52	59787	0	0	100
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	72072	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	72891	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	63063	0	0	100
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	54873	0	0	100
88	14	39517884	39517966	83	83	S. O.	0,25	67977	0	0	100
89	14	46958962	46959034	73	72	Poly T (5), indel	0,19	58642	0	326	99,4
90	14	58050030	58050110	81	81	Indel	0,38	66339	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	74529	0	0	100
92	14	92549544	92549609	66	66	Poly A (5)	0,41	54054	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	76986	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
94	15	43170751	43170848	98	96	Poly C (5)	0,45	78624	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	55692	0	0	100
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	76167	0	0	100
97	15	81625334	81625428	95	95	Poly T (6)	0,43	77805	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	58149	0	0	100
99	15	89817413	89817503	91	91	S. O.	0,36	74529	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	57330	0	0	100
101	16	1894910	1894972	63	63	S. O.	0,27	51597	0	0	100
102	16	28997904	28997998	95	95	Poly C (5)	0,67	77805	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	71253	0	0	100
104	16	57954406	57954509	104	104	Poly C (5)	0,67	85176	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	74529	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	72891	0	0	100
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	71247	0	6	100
108	17	3970090	3970180	91	91	Indel	0,46	74529	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	76167	0	0	100
110	17	33998759	33998849	91	89	Poly T (5)	0,54	72891	0	0	100
111	17	39589691	39589774	84	82	Poly A (13), indel (x2)	0,29	66343	27	788	98,8
112	17	41244394	41244484	91	91	Poly A (5)	0,34	74529	0	0	100
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT(4), AT(4), indel	0,26	75348	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	64413	0	288	99,6
115	17	64023582	64023667	86	86	Poly T (7)	0,22	70434	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	68796	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	54873	0	0	100
118	18	6980478	6980568	91	91	S. O.	0,37	74529	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	56511	0	0	100
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	61425	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	66339	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	69615	0	0	100
123	18	59773996	59774060	65	65	S. O.	0,48	53235	0	0	100
124	19	625143	625241	99	99	S. O.	0,59	81081	0	0	100
125	19	18121418	18121491	74	74	S. O.	0,68	60605	1	0	100
126	19	18186574	18186643	70	70	S. O.	0,64	57330	0	0	100
127	20	746056	746149	94	94	S. O.	0,61	76986	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	67158	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	62244	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	57330	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	82719	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	54873	0	0	100
133	20	62331904	62331994	91	88	Poly G (6)	0,73	72072	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	71253	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	54054	0	0	100
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	80262	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	71253	0	0	100
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	56439	0	72	99,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	73710	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	81900	0	0	100
141	22	32439233	32439329	97	97	S. O.	0,68	79443	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	79443	0	0	100
143	22	37637596	37637694	99	99	S. O.	0,6	81081	0	0	100

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
144	22	47081347	47081438	92	92	Indel	0,66	75348	0	0	100
145	X	15870424	15870492	69	69	Poly T (5)	0,26	56511	0	0	100
146	X	135288543	135288611	69	69	Poly C (5)	0,62	56511	0	0	100
147	X	135290777	135290847	71	71	S. O.	0,52	58149	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.
150	Y	2655609	2655679	71	0	Poly A (5)	0,37	0	0	0	S. O.

Les résultats de séquençage pour l'échantillon NA12878 ont été comparés à un génotype d'une grande fiabilité établi pour NA12878 par le National Institutes of Standards and Technology (NIST, Institut national des normes et de la technologie) (v.2.19). Parmi les 150 amplicons, 92 étaient pleinement compris dans les régions génomiques de grande fiabilité, 41 montraient des chevauchements partiels et 17 ne comportaient aucun chevauchement avec la séquence du NIST. Il en a résulté 10 000 coordonnées par réplicat, pour comparaison. Les définitions des bases sans variant ont été comparées à la séquence du génome humain de référence version hg19. Les résultats sur la précision sont affichés dans le [Tableau 12](#).

Tableau 12 Concordance germinale de l'échantillon NA12878 avec la base de données du NIST

Échantillon	Nombre d'amplicons	Taux d'appel moyen	TP	FN	TN	FP	CPP	CNP	PGC
NA12878	133	> 99,9	6552	1	610470	0	> 99,9	100	> 99,9

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

- ▶ Teneur en GC ≥ 19 % (toutes les bases définies dans 819 amplicons séquencés avec une teneur en GC de 19 % ont été appelées correctement et le taux d'absences d'appel est de 0,6 %)
- ▶ Teneur en GC ≤ 87 % (toutes les bases définies dans 819 amplicons séquencés avec une teneur en GC de 87 % ont été appelées correctement et il n'y a eu aucune absence d'appel)
- ▶ Longueurs de PolyA ≤ 9 (toutes les bases définies dans 819 amplicons séquencés avec une répétition PolyA de neuf nucléotides ont été appelées correctement et il n'y a eu aucune absence d'appel)
- ▶ Longueurs de PolyT ≤ 10 (toutes les bases définies dans 819 amplicons séquencés avec une répétition PolyT de 10 nucléotides ont été appelées correctement et il n'y a eu aucune absence d'appel)
- ▶ Longueurs de PolyG ≤ 7 (toutes les bases définies dans 819 amplicons séquencés avec une répétition PolyG de sept nucléotides ont été appelées correctement et le taux d'absences d'appel est de 1,0 %)
- ▶ Longueurs de PolyC ≤ 6 (toutes les bases définies dans 2 457 amplicons séquencés avec une répétition PolyC de 6 nucléotides ont été appelées correctement et il n'y a eu aucune absence d'appel)
- ▶ Longueurs de répétition de dinucléotides $\leq 11x$ (toutes les bases définies dans 819 amplicons séquencés avec une répétition de dinucléotides de 11x ont été appelées correctement et le taux d'absences d'appel est de 0,5 %)
- ▶ Longueurs de répétition de trinucléotides $\leq 5x$ (toutes les bases définies dans 819 amplicons séquencés avec une répétition de trinucléotides de 5x ont été appelées correctement et le taux d'absences d'appel est de 0,5 %)
- ▶ Longueurs d'insertion ≤ 24 (66 343 bases définies sur 66 370 dans 819 amplicons séquencés avec une insertion de 24 nucléotides ont été appelées correctement et le taux d'absences d'appel est de 1,2 %; aucun appel inexact n'est survenu dans la région contenant l'insertion de 24 nucléotides)
- ▶ Longueurs de la délétion ≤ 25 (toutes les bases définies dans 2 457 amplicons séquencés avec une délétion de 25 nucléotides ont été appelées correctement et il n'y a eu aucune absence d'appel)

Variant somatique

L'étude décrite ici a servi à évaluer la précision de la définition de variants du module d'analyse des variants somatiques sur l'instrument NextSeq 550Dx avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles).

Cette étude a utilisé un test représentatif conçu pour étudier une variété de gènes couvrant 12 588 bases (150 amplicons) sur 23 chromosomes différents. 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 la fréquence des variants est d'environ 5 % et 7 %. L'échantillon d'ADN GM12878 a été dilué de façon similaire avec l'échantillon d'ADN GM12877 pour créer les échantillons GM12878-D5 et GM12878-D7. Chacun des échantillons a été testé trois fois, 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. La précision a été déterminée pour les SNV, les insertions et les délétions par la comparaison des résultats avec ceux de la méthode de référence composite bien caractérisée, Platinum Genomes version 2016-1.0. Les régions génomiques de fiabilité ont été définies sur la base de cette méthode de référence, à moins d'indication contraire.

Tableau 13 Sommaire de la concordance somatique

Critères	Observations totales ¹	Résultat en fonction de l'observation ²	Résultat en fonction de l'analyse ³
CPP pour les SNV	378	98,9	99,9
CPP pour les insertions	378	96,9	99,9
CPP pour les délétions	378	97,1	99,9
CNP	378	> 99,9	> 99,9
PGC	378	> 99,9	> 99,9

¹Calculées en fonction du nombre d'échantillons par analyse (91) x nombre d'analyses (9) = 378.

²Valeur la plus faible observée par réplicat d'échantillon lors des neuf analyses.

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

Le **Tableau 14** montre les données de l'étude présentées avec la concordance positive et négative exprimée en pourcentage, par échantillon, où les résultats des variants sont comparés avec la méthode de référence composite bien caractérisée pour le calcul de la concordance positive en pourcentage (CPP). Les trois types de variants (SNV, insertions et délétions) sont combinés. Étant donné que la méthode de référence ne donne des résultats que pour les variants mononucléotides et les insertions/délétions, les résultats des bases sans variants sont comparés à la séquence de référence du génome humain version hg19 pour le calcul de la concordance négative en pour cent (CNP).

Tableau 14 Concordance des variants somatiques, par échantillon

Échantillon	Taux d'appel moyen	Prévu	TP	FN	Absence d'appel de variant	TN	FP	CPP	CNP	PGC
GM12877	98,7	2052	2025	0	27	318682	15	100	> 99,9	> 99,9
GM12878	98,8	3645	3564	0	81	317645	0	100	100	100
GM12879	99,8	2592	2538	0	54	323614	2	100	> 99,9	> 99,9
GM12884	99,8	3078	3024	0	54	322038	5	100	> 99,9	> 99,9
GM12885	99,8	3294	3213	0	81	322121	0	100	100	100
GM12888	99,8	2916	2889	0	27	323048	2	100	> 99,9	> 99,9
GM12877-D5	99,8	9288	8930	0	358	630621	0	100	100	100
GM12877-D7	99,7	9288	9032	0	256	629719	0	100	100	100
GM12878-D5	99,5	9288	8699	42	547	628582	0	99,5	100	> 99,9
GM12878-D7	99,7	9288	9108	0	180	629803	0	100	100	100

Le **Tableau 15** montre les données de l'étude présentées par échantillon, où les résultats des variants sont comparés avec la méthode de référence composite bien caractérisée. La détection est évaluée séparément pour chaque type de variant (SNV, insertions et délétions). 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	2457	2457	0	540	513	0	648	594	0
GM12879	1539	1539	0	567	540	0	486	459	0
GM12884	1836	1836	0	675	648	0	567	540	0
GM12885	2025	2025	0	675	648	0	594	540	0
GM12888	1782	1782	0	621	621	0	513	486	0
GM12877-D5	5454	5392	0	1782	1647	0	2052	1891	0
GM12877-D7	5454	5406	0	1782	1728	0	2052	1898	0
GM12878-D5	5454	5192	28	1782	1651	9	2052	1856	5
GM12878-D7	5454	5445	0	1782	1719	0	2052	1944	0

Les dix échantillons ont également été analysés à la recherche de petites insertions et délétions (indels) (Tableau 16). Il y avait au total 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 Sommaire de la détection des indels somatiques

Type de variant	Variants prévus	TP	FN	Absence d'appel de variant	CPP
Insertion	10773	10282	9	482	99,2
Délétion	11502	10667	5	830	> 99,9

Les 150 amplicons ont été conçus pour couvrir un contenu génomique varié. La teneur en GC des amplicons allait de 0,19 % à 0,87 %. Les amplicons comprenaient également diverses répétitions de mononucléotides (p. ex., PolyA, PolyT), de dinucléotides et de trinucléotides. Les données ont été compilées par amplicon (Tableau 17) de façon à déterminer l'effet du contenu génomique sur le pourcentage d'appels exacts. Le pourcentage d'appels exacts se rapporte aux appels de variants et aux appels de référence et est inférieur à 100 % en cas d'appels inexacts ou d'absences d'appel.

Tableau 17 Variants somatiques – Précision au niveau de l'amplicon

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
1	1	36450499	36450591	93	93	Indel	0,22	35066	0	88	99,7
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	29827	0	35	99,9
3	1	218353867	218353957	91	91	Indel	0,4	34202	0	283	99,2
4	1	223906657	223906748	92	92	Indel	0,49	34613	0	163	99,5
5	1	228526602	228526682	81	81	Poly G (5)	0,69	30571	0	47	99,8
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	26452	0	8	100,0
7	1	247812041	247812128	88	88	Poly A (5), CT(3), TAA(3), indel	0,27	33148	0	116	99,7
8	2	55862774	55862863	90	90	Indel	0,28	33928	0	92	99,7
9	2	87003930	87004009	80	80	Indel	0,38	30218	0	22	99,9
10	2	177016721	177016805	85	81	S. O.	0,65	30616	0	2	> 99,9
11	2	186625727	186625801	75	75	Poly A (8)	0,35	28017	0	499	98,3
12	2	190323504	190323591	88	88	Poly T (5)	0,42	33207	0	57	99,8
13	2	200796740	200796826	87	87	Poly T (5), indel	0,31	32524	9	718	97,8
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	33972	0	456	98,7
15	2	228147052	228147144	93	93	S. O.	0,43	35051	0	103	99,7
16	2	235016350	235016422	73	73	Poly T (5), indel	0,42	27459	0	136	99,5
17	3	4466229	4466321	93	93	AT(3), indel	0,27	34534	0	620	98,2
18	3	46620561	46620643	83	83	S. O.	0,43	31339	0	44	99,9
19	3	49851331	49851400	70	70	CT(3), indel	0,49	26373	0	87	99,7

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG (3)	0,41	32829	0	857	97,5
21	3	190106030	190106104	75	74	Indel	0,57	27925	0	47	99,8
22	4	2233667	2233744	78	78	Poly A (6)	0,26	29327	4	162	99,4
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	36585	0	117	99,7
24	4	15688604	15688681	78	78	S. O.	0,29	29427	0	57	99,8
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	23356	5	75	99,7
26	4	102839244	102839314	71	69	Poly A (5)	0,46	25942	0	140	99,5
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	22944	0	560	97,6
28	5	1882081	1882158	78	75	S. O.	0,78	28299	0	53	99,8
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	31658	0	94	99,7
30	5	41069808	41069871	64	64	S. O.	0,39	24120	0	72	99,7
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	31297	0	77	99,8
32	5	147475343	147475409	67	67	Poly T (5)	0,37	25277	0	55	99,8
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	34308	0	90	99,7
34	5	155662213	155662287	75	75	Indel	0,43	28266	0	163	99,4
35	6	6318713	6318814	102	102	Poly G (6)	0,68	38489	0	67	99,8
36	6	24949983	24950074	92	92	Indel	0,63	34730	0	46	99,9
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	35057	0	483	98,6
38	6	32147987	32148084	98	98	Poly T (5), TCT(3), CTT(3)	0,55	36647	0	406	98,9
39	6	32986864	32986958	95	95	Indel	0,53	35681	0	238	99,3
40	6	33408498	33408583	86	86	Poly C (6)	0,7	32438	0	70	99,8
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	35441	0	91	99,7
42	6	112435865	112435955	91	91	Poly A (5)	0,44	34354	0	44	99,9
43	7	22202076	22202148	73	73	S. O.	0,44	27575	0	28	99,9

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
44	7	66276100	66276187	88	88	Indel	0,35	33060	0	213	99,4
45	7	77365735	77365821	87	87	Poly A (7), AG(4)	0,26	32423	0	489	98,5
46	7	110939946	110940030	85	85	Indel	0,38	32074	0	56	99,8
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	33791	0	281	99,2
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	34316	0	82	99,8
49	7	154404519	154404599	81	66	S. O.	0,31	24901	0	47	99,8
50	7	156476507	156476599	93	93	Indel	0,35	35067	0	87	99,8
51	8	1817312	1817394	83	83	S. O.	0,42	31365	0	9	> 99,9
52	8	24811020	24811109	90	89	Poly G (7), CTC(4), indel	0,61	32781	0	890	97,4
53	8	76518625	76518691	67	67	Indel	0,3	25228	0	146	99,4
54	9	103054909	103055006	98	98	Poly G (6)	0,67	36968	0	76	99,8
55	9	105586150	105586214	65	65	Indel	0,32	24472	0	100	99,6
56	9	107620823	107620918	96	96	S. O.	0,49	36203	0	85	99,8
57	9	123769149	123769231	83	83	AT(3)	0,37	31329	0	45	99,9
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	36472	0	201	99,5
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	29473	0	11	> 99,9
60	10	11784629	11784726	98	91	GC(3)	0,87	34188	0	213	99,4
61	10	27317777	27317855	79	79	Poly T (5)	0,3	29843	0	19	99,9
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	33968	0	68	99,8
63	10	45084159	45084253	95	95	Indel	0,35	35829	0	81	99,8
64	10	55892599	55892687	89	88	AC(11), indel	0,42	32098	88	2048	93,8
65	10	101611250	101611329	80	80	S. O.	0,49	30217	0	28	99,9
66	10	118351373	118351453	81	81	S. O.	0,51	30531	0	96	99,7
67	11	8159816	8159912	97	96	S. O.	0,45	36105	0	192	99,5
68	11	30177648	30177717	70	70	Indel	0,46	26318	0	153	99,4

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
69	11	47470345	47470444	100	100	S. O.	0,65	37785	0	24	99,9
70	11	59837679	59837740	62	62	Indel	0,37	23368	0	68	99,7
71	11	64418856	64418957	102	102	S. O.	0,59	38546	0	10	> 99,9
72	11	93529612	93529684	73	73	Poly A (5)	0,4	27516	0	78	99,7
73	11	101347052	101347136	85	85	S. O.	0,42	32083	0	48	99,9
74	11	102477336	102477426	91	91	Poly G (6)	0,55	34047	0	369	98,9
75	11	118406285	118406369	85	85	Indel	0,53	32065	0	74	99,8
76	11	120357801	120357885	85	85	Poly A (5), CA(3), indel	0,34	32083	0	47	99,9
77	11	125769313	125769397	85	85	GA(3)	0,52	32103	0	27	99,9
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	31645	16	525	98,3
79	12	26811004	26811096	93	93	Poly A (7), AC(4)	0,33	34824	0	330	99,1
80	12	30881766	30881846	81	81	S. O.	0,49	30497	0	121	99,6
81	12	88474105	88474175	71	71	Poly A (6)	0,35	26773	0	65	99,8
82	12	120966872	120966966	95	95	Poly G (5)	0,68	35830	9	72	99,8
83	13	24167504	24167576	73	73	S. O.	0,52	27498	0	114	99,6
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	32824	0	566	98,3
85	13	44880112	44880200	89	89	Indel	0,49	33574	0	77	99,8
86	13	77665218	77665294	77	77	Indel	0,39	29075	0	31	99,9
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	25313	0	13	99,9
88	14	39517884	39517966	83	83	S. O.	0,25	31360	0	22	99,9
89	14	46958962	46959034	73	72	Poly T (5), indel	0,19	26499	0	717	97,4
90	14	58050030	58050110	81	81	Indel	0,38	30494	0	133	99,6
91	14	82390559	82390649	91	91	Indel	0,35	34313	0	86	99,7
92	14	92549544	92549609	66	66	Poly A (5)	0,41	24555	0	1527	94,1
93	14	102808496	102808589	94	94	Indel	0,62	35472	0	69	99,8

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
94	15	43170751	43170848	98	96	Poly C (5)	0,45	36264	0	24	99,9
95	15	63446149	63446216	68	68	Indel	0,25	25667	0	37	99,9
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	34745	0	432	98,8
97	15	81625334	81625428	95	95	Poly T (6)	0,43	35870	0	40	99,9
98	15	85438263	85438334	72	71	Indel	0,65	26762	0	76	99,7
99	15	89817413	89817503	91	91	S. O.	0,36	34286	0	112	99,7
100	15	89864274	89864343	70	70	Indel	0,56	26449	0	11	> 99,9
101	16	1894910	1894972	63	63	S. O.	0,27	23809	0	5	> 99,9
102	16	28997904	28997998	95	95	Poly C (5)	0,67	35860	0	50	99,9
103	16	53682908	53682994	87	87	TA(3)	0,41	32835	0	60	99,8
104	16	57954406	57954509	104	104	Poly C (5)	0,67	39177	0	144	99,6
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	34075	0	323	99,1
106	17	3563920	3564008	89	89	GC(3)	0,64	33632	0	11	> 99,9
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	32752	0	134	99,6
108	17	3970090	3970180	91	91	Indel	0,46	34343	0	82	99,8
109	17	16084945	16085037	93	93	Indel	0,26	35077	0	78	99,8
110	17	33998759	33998849	91	89	Poly T (5)	0,54	33553	0	89	99,7
111	17	39589691	39589774	84	82	Poly A (13), indel (x2)	0,29	30554	53	2296	92,9
112	17	41244394	41244484	91	91	Poly A (5)	0,34	34360	0	38	99,9
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT(4), AT(4), indel	0,26	34367	0	418	98,8
114	17	61502432	61502510	79	79	Indel	0,41	29751	0	119	99,6
115	17	64023582	64023667	86	86	Poly T (7)	0,22	32176	0	340	99,0
116	17	72308237	72308320	84	84	GAG(3)	0,62	31604	7	141	99,5
117	18	2616456	2616522	67	67	GA(3)	0,31	25273	8	45	99,8
118	18	6980478	6980568	91	91	S. O.	0,37	34386	0	12	> 99,9

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	25692	0	399	98,5
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	27923	0	893	96,9
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	30598	0	20	99,9
122	18	54815665	54815749	85	85	CT(3), indel	0,45	31969	0	161	99,5
123	18	59773996	59774060	65	65	S. O.	0,48	24531	0	48	99,8
124	19	625143	625241	99	99	S. O.	0,59	37298	0	124	99,7
125	19	18121418	18121491	74	74	S. O.	0,68	27881	0	109	99,6
126	19	18186574	18186643	70	70	S. O.	0,64	26442	0	26	99,9
127	20	746056	746149	94	94	S. O.	0,61	35501	0	31	99,9
128	20	10633195	10633276	82	82	AC(3)	0,59	30951	0	72	99,8
129	20	17705633	17705708	76	76	CT(3)	0,58	28686	0	42	99,9
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	26372	0	88	99,7
131	20	25278421	25278521	101	101	Indel	0,63	38159	0	20	99,9
132	20	50897302	50897368	67	67	Indel	0,36	25188	0	544	97,9
133	20	62331904	62331994	91	88	Poly G (6)	0,73	32969	0	309	99,1
134	20	62690860	62690946	87	87	Indel	0,57	32818	0	77	99,8
135	21	30300823	30300888	66	66	Indel	0,35	24758	9	181	99,2
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	36902	0	160	99,6
137	21	36710706	36710792	87	87	GT(3), indel	0,39	32841	0	48	99,9
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	25939	0	280	98,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	33942	0	78	99,8
140	22	25750774	25750873	100	100	Indel	0,63	37733	0	86	99,8
141	22	32439233	32439329	97	97	S. O.	0,68	36617	0	49	99,9
142	22	37409844	37409940	97	97	Indel	0,46	36525	0	162	99,6
143	22	37637596	37637694	99	99	S. O.	0,6	37398	0	24	99,9

Amplicon	Chromosome	Début de l'amplicon	Fin de l'amplicon	Taille du fragment analysé	Bases dans les régions de fiabilité	Contenu génomique de l'amplicon	Teneur en GC	Appels exacts	Appels inexacts	Absence d'appel	% d'appels exacts
144	22	47081347	47081438	92	92	Indel	0,66	34754	0	22	99,9
145	X	15870424	15870492	69	69	Poly T (5)	0,26	26046	0	36	99,9
146	X	135288543	135288611	69	69	Poly C (5)	0,62	26019	0	63	99,8
147	X	135290777	135290847	71	71	S. O.	0,52	26780	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.
150	Y	2655609	2655679	71	0	Poly A (5)	0,37	0	0	0	S. O.

Les résultats de séquençage pour l'échantillon GM12878 ont été comparés à un génotype d'une grande fiabilité établi pour NA12878 par le National Institutes of Standards and Technology (NIST, Institut national des normes et de la technologie) (v.2.19). Parmi les 150 amplicons, 92 étaient pleinement compris dans les régions génomiques de grande fiabilité, 41 montraient des chevauchements partiels et 17 ne comportaient aucun chevauchement avec la séquence du NIST. Il en a résulté 10 000 coordonnées par réplicat, pour comparaison. Les définitions des bases sans variant ont été comparées à la séquence du génome humain de référence version hg19. Les résultats sur la précision sont affichés dans le [Tableau 18](#).

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

Échantillon	Nombre d'amplicons	Taux d'appel moyen	TP	FN	TN	FP	CPP	CNP	PGC
GM12878	133	98,8	2808	0	258488	0	100	100	100

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

- ▶ Teneur en GC ≥ 19 % (toutes les bases définies dans 378 amplicons séquencés avec une teneur en GC de 19 % ont été appelées correctement et le taux d'absences d'appel est de 2,6 %)
- ▶ Teneur en GC ≤ 87 % (toutes les bases définies dans 378 amplicons séquencés avec une teneur en GC de 87 % ont été appelées correctement et le taux d'absences d'appel est de 0,6 %)
- ▶ Longueurs de PolyA ≤ 9 (toutes les bases définies dans 378 amplicons séquencés avec une répétition PolyA de neuf nucléotides ont été appelées correctement et le taux d'absences d'appel est de 2,5 %)
- ▶ Longueurs de PolyT ≤ 10 (toutes les bases définies dans 378 amplicons séquencés avec une répétition PolyT de dix nucléotides ont été appelées correctement et le taux d'absences d'appel est inférieur à 0,1 %)
- ▶ Longueurs de PolyG ≤ 6 (toutes les bases définies dans 2 268 amplicons séquencés avec une répétition PolyG de six nucléotides ont été appelées correctement et le taux d'absences d'appel est de 0,5 %)
- ▶ Longueurs de PolyC ≤ 6 (toutes les bases définies dans 756 amplicons séquencés avec une répétition PolyC de six nucléotides ont été appelées correctement et le taux d'absences d'appel est de 0,4 %)
- ▶ Longueurs de répétition de dinucléotides $\leq 4x$ (toutes les bases définies dans 1 890 amplicons séquencés avec une répétition de dinucléotides de 4x ont été appelées correctement et le taux d'absences d'appel est de 0,9 %)
- ▶ Longueurs de répétition de trinucléotides $\leq 5x$ (toutes les bases définies dans 378 amplicons séquencés avec une répétition de trinucléotides de 5x ont été appelées correctement et le taux d'absences d'appel est de 1,4 %)
- ▶ Longueurs de l'insertion ≤ 23 (toutes les bases définies dans 378 amplicons séquencés avec une insertion de 23 nucléotides ont été appelées correctement et le taux d'absences d'appel est de 0,8 %)
- ▶ Longueurs de la délétion ≤ 25 (toutes les bases définies dans 1 134 amplicons séquencés avec une délétion de 25 nucléotides ont été appelées correctement et le taux d'absences d'appel est de 0,7 %)

Précision

La précision de l'instrument NextSeq 550Dx a été déterminée en testant 13 échantillons uniques de Platinum Genome au moyen de trois instruments, trois lots de réactifs et trois opérateurs afin de générer neuf analyses de séquençage sur une période de cinq jours. Les tests, échantillons et méthodes de référence représentatifs étaient les mêmes que ceux qui ont été utilisés pour l'étude de précision germinale. Les contributions relatives à la précision ont été déterminées par l'analyse des composantes de variance au moyen de la FAV comme variable dépendante et le calcul des écarts-types au niveau de l'instrument, du lot de réactifs, de l'opérateur et de la période ([Tableau 19](#)). Le nombre total d'observations utilisées dans l'analyse pour chacun des composants de variabilité, soit l'instrument, l'opérateur ou le lot de réactifs, était respectivement de 699, 176 et 235 pour les SNV, les insertions et les délétions.

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

Composant	Type de variant	Écart-type du composant		Écart-type 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)

Des échantillons de sang entier et des échantillons FFPE ont été analysés sur les instruments NextSeq 550Dx et MiSeqDx au moyen du flux de travail germinal ou somatique de la trousse d'amplicons personnalisés TruSeq Dx. La concordance de la fréquence de variants pour les échantillons de sang et les échantillons FFPE a été évaluée au moyen de multiples tests représentatifs. La [Figure 2](#) représente la corrélation de la FAV entre les deux instruments pour un test représentatif et le [Tableau 20](#) résume cette corrélation en fonction du panel de tests. D'après la forte corrélation entre les instruments MiSeqDx et NextSeq 550Dx, les caractéristiques de performance liées aux facteurs préanalytiques (p. ex., les méthodes d'extraction ou les substances interférentes) sont déterminées comme étant applicables aux deux instruments. Pour obtenir des précisions supplémentaires, consultez la notice d'accompagnement de la trousse d'amplicons personnalisés TruSeq Dx.

Figure 2 Corrélation de la FAV selon les instruments MiSeqDx et NextSeq 550Dx pour les échantillons FFPE (gauche) et les échantillons de sang (droite) avec le test 1

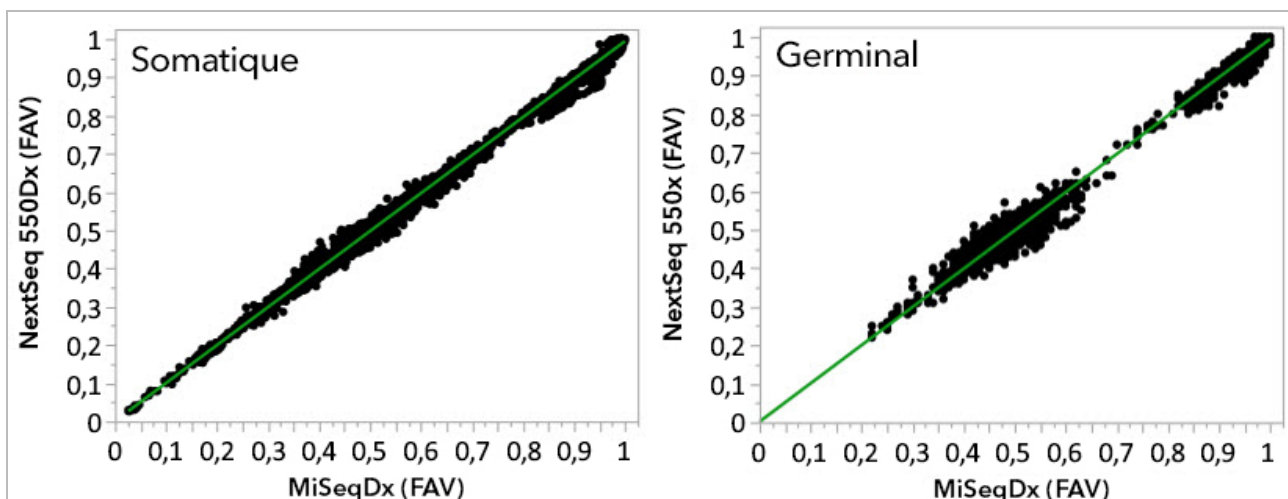


Tableau 20 Résultats de la comparaison des méthodes utilisant des échantillons de sang uniques et des échantillons FFPE

Source d'ADNg	Test (panel d'oligos)	Réplicats biologiques (échantillons)	Réplicats techniques (par échantillon)	Observations (nbre de variants)	Inclinaison	Interception	Corrélation (R ²)
Sang	Test 1	45	2	8369 ¹	0,992	0,002	0,995 ²
Sang	Test 2	45	2	5457	0,995	0,005	0,981
FFPE	Test 1	46	2	8319	0,993	0,000	0,997 ²
FFPE	Test 3	40	1	280	0,969	0,015	0,978

¹ Deux points de données ont été retirés en raison de la limite indiquée pour le module d'analyse des variants germinaux.

² Coefficients de détermination des courbes de la FAV, illustrée dans le tableau 2.

Reproductibilité

La reproductibilité de l'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. Le test des variants germinaux comprenait sept réplicats de 13 échantillons; le test des variants somatiques comprenait six réplicats de sept échantillons à différents niveaux de FAV. Les échantillons ont été préparés à l'aide de la trousse d'amplicons personnalisés TruSeq Dx.

Le test a été effectué dans trois sites externes avec un lot de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles). Un seul instrument NextSeq 550Dx a été utilisé à chacun des sites. Deux opérateurs ont effectué les tests à 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 dans les trois sites. Il en a résulté 18 analyses pour chacun des flux de travail germinaux et somatiques.

Variant germinaux

Les variants germinaux dont la FAV est supérieure ou égale à 0,2 sont déclarés comme étant (des variants) positifs. Dans le cas des variants germinaux positifs prévus, les données ont été évaluées pour le taux d'absence d'appel et le taux d'appels positifs exacts de chacun des types 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 appels de variants germinaux pour les résultats positifs prévus selon le type de variant

Type de variant	Absence d'appel			Appel positif exact				
	Observé	Total	Pourcentage	Observé	Total	Pourcentage	NCI 95 %	NCS 95 %
SNV	16	110 376	0,014	110 349	110 360	99,99	99,98	99,99
Insertions	1026	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 dont la FAV est inférieure à 0,2 sont déclarés comme étant négatifs (type sauvage). Dans le cas des emplacements des variants germinaux négatifs prévus, les données ont été évaluées selon le taux d'absence d'appel et le taux d'appels exacts 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.

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

Type de variant	Absence d'appel			Appel négatif exact				
	Observé	Total	Pourcentage	Observé	Total	Pourcentage	NCI 95 %	NCS 95 %
Type sauvage	4883	19 600 182	0,025	19 595 299	19 595 299	100	100,00	100,00

Les variants germinaux ayant une FAV supérieure ou égale à 0,2 et inférieure à 0,7 sont qualifiés d'hétérozygotes positifs pour le variant et les variants ayant une FAV supérieure ou égale à 0,7 sont qualifiés d'homozygotes positifs pour le variant. Des échantillons germinaux ayant des variants hétérozygotes ont été utilisés pour déterminer si la variabilité inhérente au test influe sur le typage génotypique. Le Cx a été déterminé pour les deux limites (0,2 pour les génotypes hétérozygotes et 0,7 pour les génotypes homozygotes), où x correspond à la proportion des tests répétés ayant excédé la limite. Pour la limite inférieure de FAV de 0,2, le Cx était supérieur ou égal à 99,999 %, ce qui indique qu'une proportion supérieure ou égale à 99,999 % des variants hétérozygotes serait définie comme hétérozygote. En ce qui a trait à la limite supérieure de FAV de 0,7, le Cx était inférieur ou égal à 0,001 %, ce qui indique qu'une proportion inférieure ou égale à 0,001 % des variants hétérozygotes serait définies comme homozygote. Le [Tableau 23](#) résume les résultats par type de variant.

Les variants germinaux ayant une FAV supérieure ou égale à 0,2 et inférieure à 0,7 sont qualifiés d'hétérozygotes positifs pour le variant et les variants ayant une FAV supérieure ou égale à 0,7 sont qualifiés d'homozygotes positifs pour le variant. Des échantillons germinaux ayant des variants hétérozygotes ont été utilisés pour déterminer si la variabilité inhérente au test influe sur le typage génotypique. Le Cx a été déterminé pour les deux limites (0,2 pour les génotypes hétérozygotes et 0,7 pour les génotypes homozygotes), où x correspond à la proportion des tests répétés ayant excédé la limite. En ce qui a trait à la limite inférieure de FAV de 0,2, le Cx était supérieur ou égal à 99,999 %, ce qui indique qu'une proportion supérieure ou égale à 99,999 % des variants hétérozygotes serait définie comme hétérozygote. Pour la limite supérieure de FAV de 0,7, le Cx était inférieur ou égal à 0,001 %, ce qui indique qu'une proportion inférieure ou égale à 0,001 % des variants hétérozygotes serait définies comme homozygote. Le [Tableau 23](#) résume les résultats par type de variant.

Tableau 23 Valeurs Cx germinales pour les variants hétérozygotes

Type de variant	Limite de FAV de 0,2	Limite de FAV de 0,7
	≥ C99,999 %	≤ 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 dont la FAV est supérieure ou égale à 0,026 sont déclarés comme étant (des variants) positifs. Les observations de FAV supérieure ou égale à 0,01 et inférieure à 0,026 ont été considérées comme équivoques aux fins de cette analyse (ni positives ni négatives, marquées comme étant une fréquence de variants faible). Pour évaluer la performance, les résultats ont été calculés de trois façons :

- ▶ Meilleur cas : tout résultat équivoque a été considéré comme étant un appel positif exact (concordance avec les résultats prévus)
- ▶ Pire cas : tout résultat équivoque a été considéré comme étant un appel inexact (non-concordance avec les résultats prévus)
- ▶ Cas d'exclusion : tout résultat équivoque a été exclu de l'analyse

Trois tableaux, soit le [Tableau 24](#), le [Tableau 25](#) et le [Tableau 26](#), résument les résultats des appels, respectivement selon le scénario du meilleur cas, du pire cas et du cas d'exclusion, ainsi que les niveaux de confiance inférieurs (NCI) et supérieurs (NCS) à 95 %, calculés selon la méthode de Wilson.

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

Type de variant	Appel positif exact				
	Observé	Total	Pourcentage	NCI 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 appels de variants somatiques pour les résultats positifs prévus selon le type de variant (pire cas)

Type de variant	Appel positif exact				
	Observé	Total	Pourcentage	NCI 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 des appels de variants somatiques pour les résultats positifs prévus selon le type de variant (appels équivoques retirés)

Type de variant	Appel positif exact				
	Observé	Total	Pourcentage	NCI 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 ayant une FAV inférieure à 0,01 sont signalés comme étant des appels négatifs (type sauvage). Dans le cas des emplacements des variants somatiques négatifs prévus, les données ont été évaluées selon le taux d'absence d'appel et le taux d'appels exacts de type sauvage. Les appels exacts de type sauvage ont été déterminés en excluant le taux d'absence d'appel et en soustrayant du total les appels observés qui tombaient dans la zone équivoque ($FAV \geq 0,01$ et $< 0,026$) ainsi que les appels inexacts supérieurs à la limite ($FAV \geq 0,026$). Le [Tableau 27](#) résume les résultats observés, les résultats totaux et les résultats en pourcentage pour les emplacements de variants somatiques négatifs dans le cas du taux d'absence d'appel et du taux d'appels exacts de type sauvage ainsi que les niveaux de confiance inférieurs (NCI) et supérieurs (NCS) à 95 % calculés selon la méthode de Wilson.

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

Type de variant	Absence d'appel			Appel exact						
	Observé	Total	Pourcentage	Équivoque	Inexact	Exact	Total	Pourcentage	NCI 95 %	NCS 95 %
Type sauvage	36 326	8 909 676	0,408	2 254	121	8 870 975	8 873 350	99,97	99,972	99,974

Les échantillons somatiques situés à des niveaux de FAV différents pour le même variant ont été évalués en vue de déterminer le C95 du test (pour chacun des types de variant). Pour l'évaluation de la variabilité près de la limite du test, les échantillons dont la FAV prévue se situait entre 0,02 et 0,07 ont été utilisés. Le C95 a été déterminé pour chacun des variants et le C95 le plus élevé pour chacun des types de variant est indiqué dans le [Tableau 28](#).

Tableau 28 Résumé du C95 des variants somatiques

Type de variant	N	C95
SNV	74	0,0613
Insertion	24	0,0573
Délétion	33	0,0575

Performance de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles)

Vue d'ensemble

NextSeq 550Dx est appuyé par deux trousse de réactifs : la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) et la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles). Dans le but de démontrer que la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) peut répondre aux exigences de performance analytique vérifiées et validées avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles), des études ont été menées avec la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles). Deux préparations de bibliothèques utilisant la trousse d'amplicons personnalisés TruSeq Dx ont été effectuées, l'une avec le flux de travail germinale et l'autre avec le flux de travail somatique. Les bibliothèques de chaque flux de travail ont été testées avec trois lots de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) en utilisant trois instruments NextSeq 550Dx. En outre, les tests pour chaque flux de travail comprenaient une seule analyse avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles).

Sensibilité analytique (limite de blanc [LB] et limite de détection [LD])

La vérification menée avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) a démontré que l'instrument NextSeq 550Dx pouvait détecter des variants ayant une FAV de 0,05 avec une erreur de type II inférieure ou égale à 0,05 et que le seuil de FAV de 0,026 utilisé par le module des variants somatiques (LB effective) supporte une erreur de type I inférieure ou égale à 0,01. Sur la base de ces affirmations, on estime qu'une variante ayant une FAV de 0,05 est supérieure ou égale à une FAV de 0,026 95 % du temps et qu'une position de type sauvage est inférieure à une FAV de 0,026 99 % du temps. Pour s'assurer que ces affirmations étaient satisfaites avec la trousse de réactifs NextSeq 550Dx à débit élevé 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 0,05 FAV (échantillons LD) en utilisant la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles). La proportion d'appels au-dessus et au-dessous du seuil de 0,026 a ensuite été comparée aux affirmations établies avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles).

Les tests comprenaient deux échantillons LD, chacun avec un ensemble unique de variants ciblés ayant une FAV de 0,05 et des échantillons LB correspondants qui étaient de type sauvage pour les variants ciblés. Pour la préparation des bibliothèques, les échantillons LD et LB ont été traités en réplicats de huit et sept, respectivement, en utilisant la trousse d'amplicons personnalisés TruSeq Dx. Les bibliothèques ont été d'abord séquencées à l'aide de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) afin d'identifier les variants/coordonnées génomiques pour l'évaluation LB/LD avec la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles). Tous les variants dont la FAV moyenne se situe entre 0,045 et 0,055 (variants LD) sur la base des résultats de la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) ont été utilisés pour l'analyse LD (N = 51 variants). Pour l'analyse LB, les 51 coordonnées génomiques correspondantes ont été évaluées.

Pour l'évaluation de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles), les bibliothèques ont été séquencées en trois analyses sur trois jours consécutifs en utilisant le même instrument et le même lot de la trousse de réactifs. Ces tests se sont élevés à 24 réplicats pour chacun des 51 variants LD et à 21 réplicats pour chacune des positions de type sauvage correspondantes. La proportion d'appels de type sauvage avec une FAV inférieure à 0,026 est indiquée dans le [Tableau 29](#). La proportion d'appels de variants LD avec une FAV supérieure ou égale à 0,026 est indiquée dans le [Tableau 30](#).

Tableau 29 Proportion d'appels inférieurs à 0,026 pour les postes de type sauvage (évaluation des affirmations LB)

Type de variant	Positions évaluées	Observations totales	Nombre de mesures de la FAV $\geq 2,6$ %	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 d'appels supérieurs ou égaux à une FAV de 0,026 pour les variants LD (évaluation des affirmations LD)

Type de variant	Positions évaluées	Observations totales	Nombre de mesures de la FAV < 2,6 %	Nombre de mesures de la FAV $\geq 2,6$ %	Proportion $\geq 2,6$ %	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 qui suit a été menée afin d'évaluer la précision de la définition de variants du module d'analyse des variants germinaux au moyen de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles). Douze échantillons uniques du Platinum Genome ont été testés à l'aide d'un test représentatif. Un total de 11 analyses ont été réalisées en utilisant trois instruments NextSeq 550Dx et trois trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).

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

Tableau 31 Sommaire de la concordance germinale

Critères	Observations totales (v2.5) ¹	Résultat en fonction de l'observation (v2.5) ²	Résultat en fonction de l'observation (v2) ³	Résultat en fonction de l'analyse (v2.5) ⁴	Résultat en fonction de l'analyse (v2) ⁴
CPP pour les SNV	1056	98,7	98,7	> 99,9	> 99,9
CPP pour les insertions	1056	100	100	100	98,9
CPP pour les délétions	1056	95,2	95,2	> 99,9	100
CNP	1056	100	100	100	100
PGC	1056	> 99,9	> 99,9	> 99,9	> 99,9

¹Calculé comme le nombre d'échantillons par analyse x le nombre d'analyse (96 échantillons par analyse x 11 analyses = 1056 observations).

²Valeur la plus faible observée par échantillon de réplicat dans toutes les analyses (sur la base de 11 analyses pour la trousse de réactifs NextSeq 550Dx à débit élevé v2.5).

³Valeur la plus faible observée par réplicat d'échantillon lors d'une analyse (96 observations totales).

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

Variant somatique

L'étude qui suit a été menée afin d'évaluer la précision de la définition de variants du module d'analyse des variants somatiques sur l'instrument NextSeq 550Dx au moyen de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles). Dix échantillons FFPE du Platinum Genome (deux avec des variants dilués à une FAV de 0,05) ont été testés à l'aide d'un test représentatif. Un total de 11 analyses ont été réalisées en utilisant trois instruments NextSeq 550Dx et trois lots de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).

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

Tableau 32 Sommaire de la concordance somatique

Critères	Observations totales (v2.5) ¹	Résultat en fonction de l'observation (v2.5) ²	Résultat en fonction de l'observation (v2) ³	Résultat en fonction de l'analyse (v2.5) ⁴	Résultat en fonction de l'analyse (v2) ⁴
CPP pour les SNV	528	100	100	100	100
CPP pour les insertions	528	96,9	96,9	> 99,9	> 99,9
CPP pour les délétions	528	100	100	100	100
CNP	528	> 99,9	> 99,9	> 99,9	> 99,9
PGC	528	> 99,9	> 99,9	> 99,9	> 99,9

¹Calculé comme le nombre d'échantillons par analyse x le nombre d'analyse (48 échantillons par analyse x 11 analyses = 528 observations).

²Valeur la plus faible observée par échantillon de réplicat dans toutes les analyses (sur la base de 11 analyses pour la trousse de réactifs NextSeq 550Dx à débit élevé v2.5).

³Valeur la plus faible observée par réplicat d'échantillon lors d'une analyse (96 observations totales).

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

Précision

Variant germinale

La précision de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) avec le module de variants germinaux a été évaluée à l'aide d'échantillons de Platinum Genome et d'un test représentatif. Les tests ont consisté en une seule préparation de librairie utilisant la trousse d'amplicons personnalisés TruSeq Dx et comprenaient 12 échantillons traités avec huit réplicats chacun. Les librairies ont été séquencées avec trois lots de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) et trois instruments NextSeq 550Dx pour un total de neuf analyses de séquençage.

Des échantillons ayant des variants hétérozygotes ont été utilisés pour déterminer si la variabilité inhérente au test influe sur le typage génotypique (N = 153 variants hétérozygotes uniques). Le Cx a été déterminé pour les deux limites du module des variants germinaux (0,2 pour les génotypes hétérozygotes et 0,7 pour les génotypes homozygotes), où x correspond à la proportion des tests répétés ayant excédé la limite. Pour la limite inférieure de la FAV à 0,2, le variant avec le Cx minimum pour la trousse de réactifs NextSeq 550Dx v2.5 (300 cycles) était supérieur à 99,9 %, ce qui indique que plus de 99,9 % des variants hétérozygotes seraient appelés hétérozygotes. Pour la limite supérieure de la FAV à 0,7, le variant avec le Cx maximum pour la trousse de réactifs NextSeq 550Dx v2.5 (300 cycles) était inférieur à 1,5 %, ce qui indique que moins de 1,5 % des variants

hétérozygotes seraient appelés homozygotes. Le **Tableau 33** résume les résultats par type de variant. Les valeurs de Cx du séquençage unique utilisant la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) sont fournies à titre de référence.

Tableau 33 Valeurs Cx germinales pour les variants hétérozygotes

Type de variant	N	Limite de FAV de 0,2		Limite de FAV de 0,7	
		Cx min (v2.5) ¹	Cx min (v2) ²	Cx max (v2.5) ¹	Cx max (v2) ²
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 %

¹Valeurs de Cx basées sur les estimations de l'écart-type total à partir de l'analyse des composantes de variance.

²Valeurs de Cx basées sur les écarts-types des échantillons.

Variant somatique

La précision de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) avec le module de variants germinaux a été évaluée à l'aide d'échantillons fpe de Platinum Genome et d'un test représentatif. Les tests ont consisté en une seule préparation de librairie utilisant la trousse d'amplicons personnalisés TruSeq Dx et comprenaient deux échantillons avec huit réplicats chacun. Les librairies ont été séquencées en utilisant trois lots de la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) et trois instruments NextSeq 550Dx pour un total de neuf analyses de séquençage.

Des variants somatiques avec des niveaux de FAV attendus inférieurs à 0,10 (N = 131 variants uniques) ont été utilisés pour évaluer la variabilité de l'instrument près de la limite de FAV du module d'analyse des variants somatiques (les variants somatiques avec un niveau de FAV supérieur ou égal à 0,026 sont dites positives pour le variant). Les valeurs de C95 ont été déterminées pour chacun des variants somatiques. Les valeurs de C95 représentent la FAV pour laquelle la probabilité d'être supérieure à la limite de FAV du module des variants somatiques est de 95 %. Les valeurs les plus élevées de C95 par type de variant sont indiquées dans le **Tableau 34**. Les valeurs de C95 du séquençage unique utilisant la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) sont fournies à titre de référence.

Tableau 34 Résumé du C95 des variants somatiques

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

¹Valeurs de C95 basées sur les estimations de l'écart-type total à partir de l'analyse des composantes de variance.

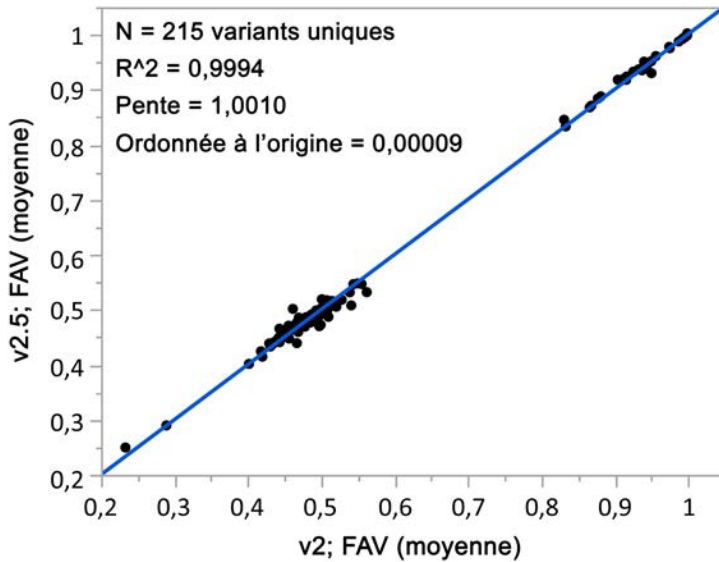
²Valeurs de C95 basées sur les écarts-types des échantillons.

Comparaison des méthodes (trousse de réactifs)

Variant germinale

Les FAV moyennes de 215 variants uniques ont été évaluées avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) et la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) en utilisant les résultats générés par le module des variants germinaux. Les FAV 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 de chaque variant. La **Figure 3** montre la corrélation de la FAV entre les deux trousse de réactifs. Sur la base de la forte corrélation linéaire de la FAV et de la similitude des résultats entre les trousse de réactifs, les caractéristiques de performance initialement vérifiées et validées avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) avec le module des variants germinaux sont déterminées comme étant applicables à la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).

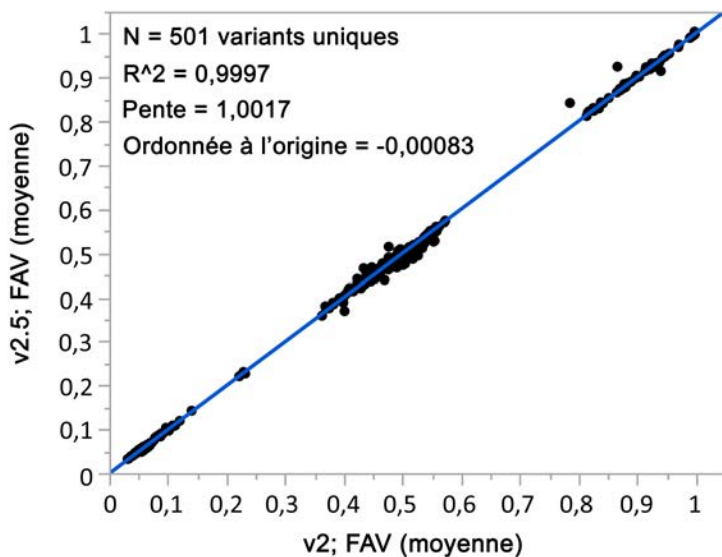
Figure 3 Corrélation de la fréquence allélique de variants (FAV) du module des variants germinaux entre la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) et la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).



Variant somatique

Les FAV moyennes de 501 variants uniques ont été évaluées avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) et la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles) en utilisant les résultats générés par le module des variants somatiques. Les FAV 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 de chaque variant. La [Figure 4](#) montre la corrélation de la FAV entre les deux trousse de réactifs. Sur la base de la corrélation de la FAV et de la similitude des résultats entre les trousse de réactifs, les caractéristiques de performance vérifiées et validées avec la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) avec le module des variants somatiques sont déterminées comme étant applicables à la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).

Figure 4 Corrélation de la fréquence allélique de variants (FAV) du module des variants somatiques entre la trousse de réactifs NextSeq 550Dx à débit élevé v2 (300 cycles) et la trousse de réactifs NextSeq 550Dx à débit élevé v2.5 (300 cycles).



Brevets et marques de commerce

Ce document et son contenu sont exclusifs à Illumina, Inc. et à ses sociétés affiliées (« Illumina »); ils sont exclusivement destinés à l'usage contractuel de son client dans le cadre de l'utilisation du ou des produits décrits dans les présentes et ne peuvent servir à aucune autre fin. Ce document et son contenu ne seront utilisés ou distribués à aucune autre fin ni communiqués, divulgués ou reproduits d'aucune façon sans le consentement écrit préalable d'Illumina. Illumina ne cède aucune licence en vertu de son brevet, de sa marque de commerce, de ses droits d'auteur ou de ses droits traditionnels ni des droits similaires d'un tiers quelconque par ce document.

Les instructions contenues dans ce document doivent être suivies strictement et explicitement par un personnel qualifié et adéquatement formé de façon à assurer l'utilisation correcte et sûre du ou des produits décrits dans les présentes. Le contenu intégral de ce document doit être lu et compris avant l'utilisation de ce ou ces produits.

SI UN UTILISATEUR NE LIT PAS COMPLÈTEMENT ET NE SUIT PAS EXPLICITEMENT TOUTES LES INSTRUCTIONS CONTENUES DANS LES PRÉSENTES, IL RISQUE DE CAUSER DES DOMMAGES AU(X) PRODUIT(S), DES BLESSURES, NOTAMMENT AUX UTILISATEURS ET À D'AUTRES PERSONNES, AINSI QUE D'AUTRES DOMMAGES MATÉRIELS, ANNULANT AUSSI TOUTE GARANTIE S'APPLIQUANT AU(X) PRODUIT(S).

ILLUMINA DÉCLINE TOUTE RESPONSABILITÉ DÉCOULANT DE L'UTILISATION INAPPROPRIÉE DU OU DES PRODUITS DÉCRITS DANS LES PRÉSENTES (Y COMPRIS LEURS COMPOSANTES ET LE LOGICIEL).

© 2017 Illumina, Inc. Tous droits réservés.

Toutes les marques de commerce sont la propriété d'Illumina, Inc. ou de leurs détenteurs respectifs. Pour obtenir des renseignements sur les marques de commerce, consultez la page www.illumina.com/company/legal.html.

Coordonnées



Illumina

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



Étiquette du produit

Pour voir la liste complète des symboles qui peuvent apparaître sur l'emballage et l'étiquetage du produit, reportez-vous à la légende des symboles, sur le site support.illumina.com, à l'onglet *Documentation & Literature* (Documentation) propre à votre trousse.