

illumina TruPath^{MC} Genome

Flux de travail le plus simple pour
le séquençage du génome entier



Découvrez le flux de travail le plus simple des échantillons aux données avec seulement 10 minutes de préparation pratique

TAGCA
CGTAG
ATAGC

Résolvez les régions difficiles à cartographier du génome en utilisant des lectures courtes et des données de proximité avec la série NovaSeq^{MC} X



Obtenez des renseignements sur de longues distances grâce à une détection améliorée des variants structurels et à une mise en phase ultralongue

Accessibilité au génome qui change le paradigme

Le séquençage à lectures courtes de nouvelle génération (SNG) a été la méthode principale utilisée par les chercheurs pour tester la majeure partie du génome. Pourtant, certaines régions et certains types de variants restent difficiles à caractériser, notamment ceux présentant une répétitivité extrême, une forte homologie de séquence ou une complexité structurelle. Des études récentes ont démontré que ces régions abritent des variants susceptibles de jouer un rôle dans les maladies génétiques humaines^{1,2}, rendant leur déchiffrement d'une importance capitale lors de la recherche des causes sous-jacentes de ces maladies.

Des méthodes alternatives, comme le séquençage à lectures longues, offrent une résolution améliorée et permettent d'obtenir davantage de renseignements sur de nombreuses régions et variants difficiles à cartographier. Cependant, les méthodes à lectures longues ont tendance à présenter des flux de travail complexes et des problèmes de précision bien connus, ce qui entraîne des résultats variables.

Illumina TruPath Genome, associé à [la technologie de lecture à cartographie de proximité](#), change le paradigme du flux de travail du séquençage du génome entier (WGS, Whole Genome Sequencing). Ce test innovant capture de longs fragments d'ADN grâce à une étape de préparation pratique simple de 10 minutes. Le séquençage est réalisé sur la série NovaSeq X existante, avec l'analyse secondaire de pointe DRAGEN^{MC}.

Le résultat est un génome précis et complet, permettant aux chercheurs d'obtenir des renseignements couvrant des distances plus grandes que celles généralement accessibles avec les lectures produites par les méthodes à lectures longues.

Avec TruPath Genome, il est désormais possible de tirer parti de la simplicité et de la précision du séquençage à lectures courtes pour résoudre les régions du génome difficiles à cartographier, améliorer la détection des variants structurels (VS) et générer des lectures et des appels de variants mis en phase.

Simplicité de flux de travail sans précédent

TruPath Genome simplifie considérablement la mise en œuvre du WGS et réduit la durée de manipulation à 10 minutes ([figure 1](#)). La chimie de la transposase éprouvée d'Illumina supprime les étapes traditionnelles de préparation de librairie. Des algorithmes bioinformatiques avancés exploitent les renseignements de proximité des nanopuits en combinaison avec des données standard de lectures courtes appariées* pour résoudre les relations qui existent entre des variants séparés par des millions de bases.

* Les données standard de lectures courtes appariées désignent les données générées par séquençage du génome entier réalisé avec une préparation manuelle de la librairie et les méthodes classiques de séquençage par synthèse (SBS).

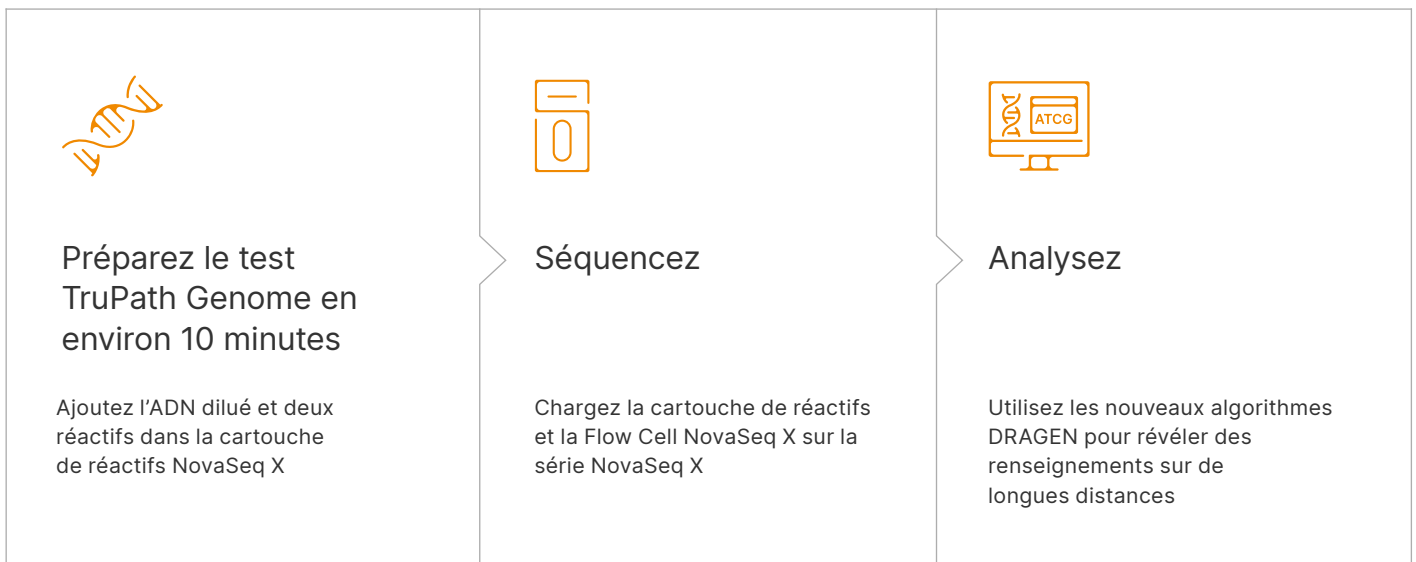


Figure 1 : Flux de travail TruPath Genome

L'ADN et un réactif TruPath Genome sont ajoutés à la barrette de tubes de librairie, qui est ensuite insérée dans la cartouche de réactifs NovaSeq X. Des réactifs TruPath Genome supplémentaires sont ensuite ajoutés à la cartouche. La cartouche et la Flow Cell sont chargées dans la série NovaSeq X et l'analyse est lancée.

Échantillons

L'ADN double brin (ADNdb) purifié est extrait à l'aide de trousseaux commerciaux standards ou de poids moléculaire élevé (HMW, High Molecular Weight). L'ADN, dont la quantité d'entrée recommandée est de 350 ng, est dilué dans le tampon de tagmentation TruPath puis ajouté à la barrette de tubes pour librairie. L'ADN d'entrée peut être isolé à partir de lignées cellulaires ou de sang total. L'ADN isolé à partir d'échantillons fixés au formol et inclus en paraffine (FFIP) ainsi que l'ADN acellulaire (ADNa) ne sont pas pris en charge. Pour des performances optimales, il est recommandé d'extraire l'ADN à l'aide d'une méthode qui préserve les fragments de grande taille, avec 40 % des fragments ayant une longueur supérieure à 60 kb.³

Pour obtenir des recommandations sur l'utilisation de types d'échantillons supplémentaires, les variations de qualité des échantillons et les quantités d'entrée optimales, consultez la note technique Performances de [TruPath Genome avec différents types d'échantillons, de qualités et quantités d'entrée](#).

Séquençage

La barrette de tubes de librairie contenant l'ADN dilué et deux réactifs TruPath Genome sont ajoutés à la cartouche de réactifs NovaSeq X, puis chargés dans la série NovaSeq X avec une Flow Cell NovaSeq X C2 ou C8, pour le séquençage de deux ou huit génomes humains, respectivement.

Les nanopuits à la surface de la Flow Cell sont amorcés avec des transposomes Illumina.[†] L'ADN est introduit dans la Flow Cell, où il est capturé par les transposomes dans plusieurs puits, formant un motif en constellation de nanopuits ensemencés ([figure 2A](#)). Les transposomes clivent l'ADN capturé et fixent les fragments à la surface de la Flow Cell dans un processus appelé tagmentation[‡] ([figure 2B](#)). Une étape de lavage permet d'éliminer l'ADN non lié et les transposomes.

À l'intérieur des nanopuits, l'ADN tagmenté est utilisé pour créer des amplifiats[§] destinés au séquençage. L'ADN provenant du même fragment long original formera des amplifiats dans des nanopuits voisins. Le séquençage est réalisé sur la série NovaSeq X en utilisant la chimie éprouvée de séquençage par synthèse (XLEAP-SBS^{MC}) Illumina XLEAP avec une analyse de 2 × 150 pb.

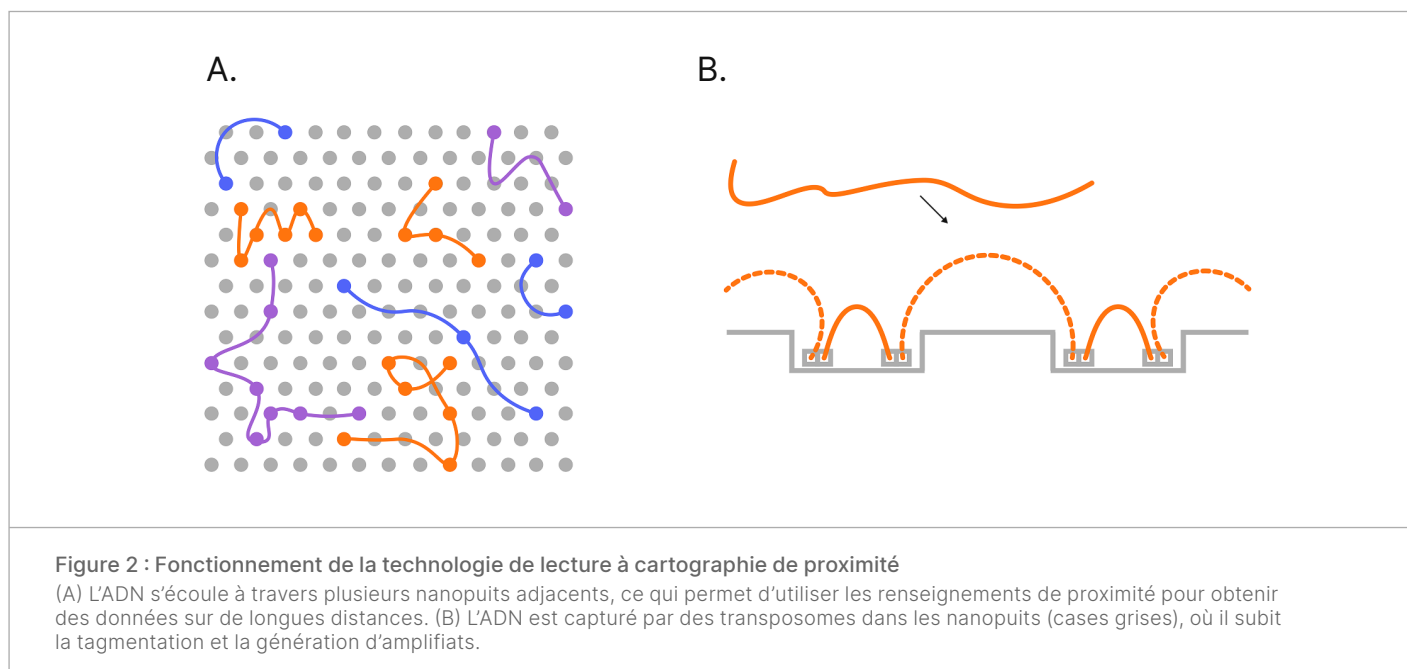
Analyse

Une fois le séquençage terminé, le pipeline DRAGEN^{MC} Germline utilise un nouvel algorithme qui combine les données de séquençage avec les renseignements de proximité des nanopuits pour aligner l'ADN sur un génome de référence. L'exploitation des renseignements de proximité permet l'attribution correcte des lectures qui, auparavant, ne se cartographiaient pas sur le génome de référence ou ne pouvaient être alignées qu'avec une faible confiance.

[†] Les transposomes sont des complexes de transposase d'ADN qui existent sous forme de dimère.

[‡] La tagmentation est le processus qui consiste à couper un fragment d'ADN et à y ajouter une séquence d'adaptateur (étiquetage) à l'aide d'un transposome.

[§] Un amplifiat est un point d'ADN amplifié sur une Flow Cell qui sera séquencée.



De cette manière, les lectures courtes standard peuvent être associées entre elles de manière probabiliste, permettant une détection complète des variants, qu'il s'agisse de petits variants (variants mononucléotidiques (SNV) et insertions/suppressions (indels)), de grands variants (> 50 pb), de régions homologues connues, ainsi que la mise en phase de petits variants pouvant être séparés par des millions de paires de bases.

Débloquez un génome complet avec des lectures courtes à haute précision

Le séquençage standard par synthèse (SBS) a démontré une détection très précise des petits variants (par ex., SNV, indels et variations du nombre de copies (VNC)), mais rencontre des difficultés avec certaines régions et certains types de variants, comme les variants structuraux. TruPath Genome introduit de nouvelles fonctionnalités au séquençage à lectures courtes, permettant d'obtenir un génome plus complet. Les lectures courtes peuvent désormais être utilisées pour analyser des régions du génome auparavant difficiles à traiter et offrir une précision accrue pour un plus grand nombre de types de variants.

Meilleure couverture dans les régions difficiles à cartographier

Les régions génomiques complexes, fortement polymorphes ou dupliquées peuvent être difficiles à assembler ou à aligner sur un génome de référence en raison d'une faible cartographiabilité, due soit à une couverture insuffisante, soit à une mauvaise qualité de cartographie (MapQ). Auparavant, la couverture de ces régions « sombres »⁴ entraînait de faibles scores MapQ, des lacunes dans les données de séquence du génome et, de manière générale, une résolution limitée pour l'appel des variants. Pour combler ces lacunes, TruPath Genome utilise les renseignements de proximité provenant des amplifiats adjacents afin d'attribuer l'emplacement génomique correct. Cela permet une cartographie à haute confiance de ces lectures initialement ambiguës (figure 3 et figure 4).

Appel des variants avec une précision accrue

TruPath Genome a été évalué sur le génome HG002 et l'ensemble vérifonctionnel T2T-Q100 v1.1 v0.019 du Consortium Telomere-to-Telomere (T2T), afin de démontrer ses performances en matière d'appel des petits variants, en détection des VS et en séquençage par mise en phase. La résolution de cartographie accrue permise par la technologie de lecture à cartographie

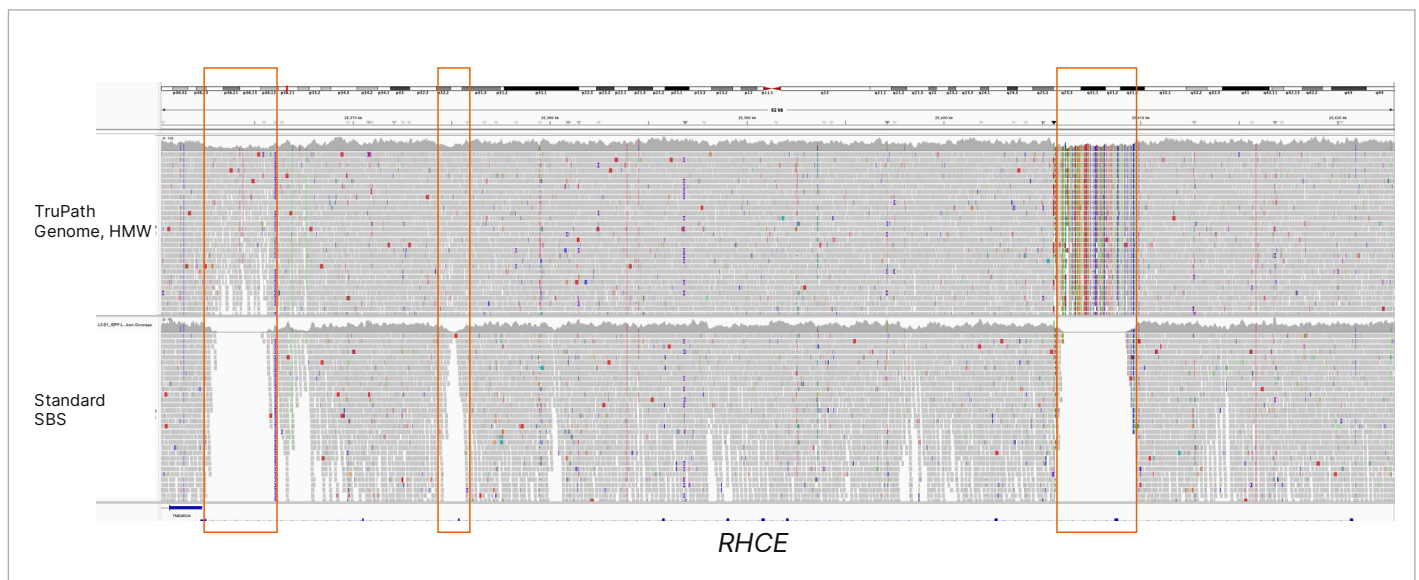


Figure 3 : Résolution efficace des gènes d'intérêt médical mais difficiles à cartographier avec TruPath Genome

Récupération de la couverture du gène *RHCE*, situé dans le locus du groupe sanguin Rh à proximité de *RHD*, et présentant une homologie de séquence très élevée avec des conversions géniques fréquentes et des réarrangements hybrides, rendant la cartographie par lectures courtes ambiguë par endroit. Un génotypage précis du gène *RHCE* est cliniquement important pour déterminer les antigènes Rh (C/c, E/e), car des incompatibilités peuvent accroître le risque de réactions hémolytiques transfusionnelles et de maladie hémolytique du fœtus/nouveau-né (HDFN, Hemolytic Disease of the Fetus/Newborn). Le SBS standard a été réalisé à l'aide d'Illumina DNA PCR-Free Prep.

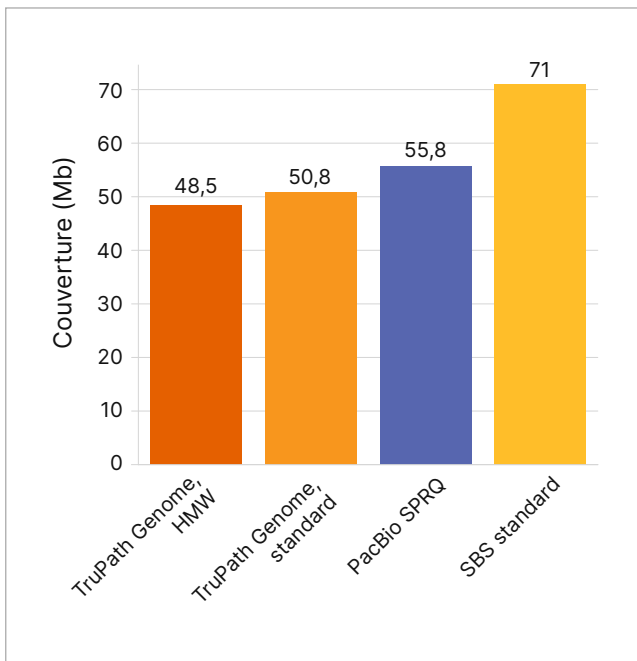


Figure 4 : TruPath Genome améliore la couverture des « régions sombres » du génome

TruPath Genome améliore la couverture des « régions sombres », définies comme les régions « sombres selon MAPQ » dans Ebbert et al¹, où 90 % des lectures couvrant la région présentent une qualité de cartographie (MAPQ) inférieure à 10. TruPath Genome a réduit la taille des régions sombres du génome grâce à une meilleure cartographie des lectures dans les zones génomiques difficiles à analyser. Un SBS standard a été réalisé en utilisant Illumina DNA PCR-Free Prep.

de proximité a amélioré la performance de l'appel des petits variants, en particulier dans les régions du génome historiquement difficiles à cartographier, telles que les régions homologues ou répétitives. Par rapport au séquençage à lectures courtes standard et à l'appel des variants, TruPath Genome présente une précision améliorée grâce à une réduction significative des appels de variants faux négatifs (FN) et faux positifs (FP) (figure 5).

TruPath Genome présente une amélioration spectaculaire de la performance d'appel des VS (> 50 pb), par rapport aux méthodes à lectures courtes standard. Le rappel des VS est passé de 86 % avec le SBS standard à 94 % grâce à la technologie de lecture à cartographie de proximité (figure 6).

Mise en phase ultralongue

Le séquençage par mise en phase fournit des renseignements sur les haplotypes permettant de distinguer les allèles des chromosomes maternels et paternels⁸. Ces renseignements aident les chercheurs à résoudre les variants hétérozygotes composés, ce qui est essentiel pour l'interprétation des maladies autosomiques récessives.

La technologie de lecture à cartographie de proximité est particulièrement adaptée pour la mise en phase du génome humain. L'ADN capturé sur la Flow Cell génère des blocs mis en phase s'étendant de centaines de kilobases à plusieurs mégabases (figure 7 et figure 8), offrant une meilleure compréhension des haplotypes et des variants hétérozygotes composés.

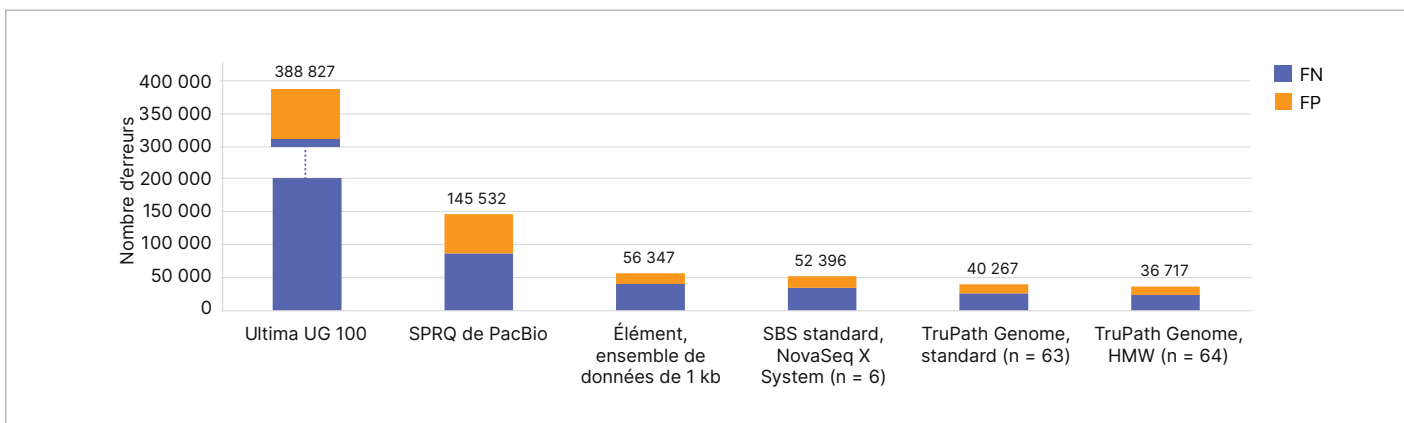


Figure 5 : Une plus grande précision pour l'appel des petits variants avec TruPath Genome

La performance de l'appel des petits variants sur différents systèmes et protocoles de SNG a été évaluée en référence à l'ensemble vérifonctionnel Genome in a Bottle NIST T2T-Q100 HG002 SV v1.1. Les données pour Ultima UG 100, PacBio avec chimie SPRQ et Element AVITI avec un jeu de données de 1 kb ont été obtenues à partir de sources publiées⁵⁻⁷. Les données SBS standard ont été générées avec des bibliothèques Illumina DNA PCR-Free Prep, séquencées sur NovaSeq X System avec le logiciel v1.4 et la Flow Cell 10B, puis analysées avec DRAGEN Germline v4.5.2 (six répliquats techniques). TruPath Genome : les données standard ont été générées à partir d'ADN extrait selon des méthodes standard, séquencées sur NovaSeq X System avec le logiciel v1.4 et la Flow Cell C8, puis analysées avec DRAGEN Germline v4.5.2 (63 répliquats techniques). Le TruPath Genome HMW a été généré à partir d'ADN extrait selon des méthodes HMW, séquencé sur NovaSeq X System avec le logiciel v1.4 et la Flow Cell C8, puis analysé avec DRAGEN Germline v4.5.2 (64 répliquats techniques). FN, appels des variants faux négatifs; FP, appels des variants faux positifs.

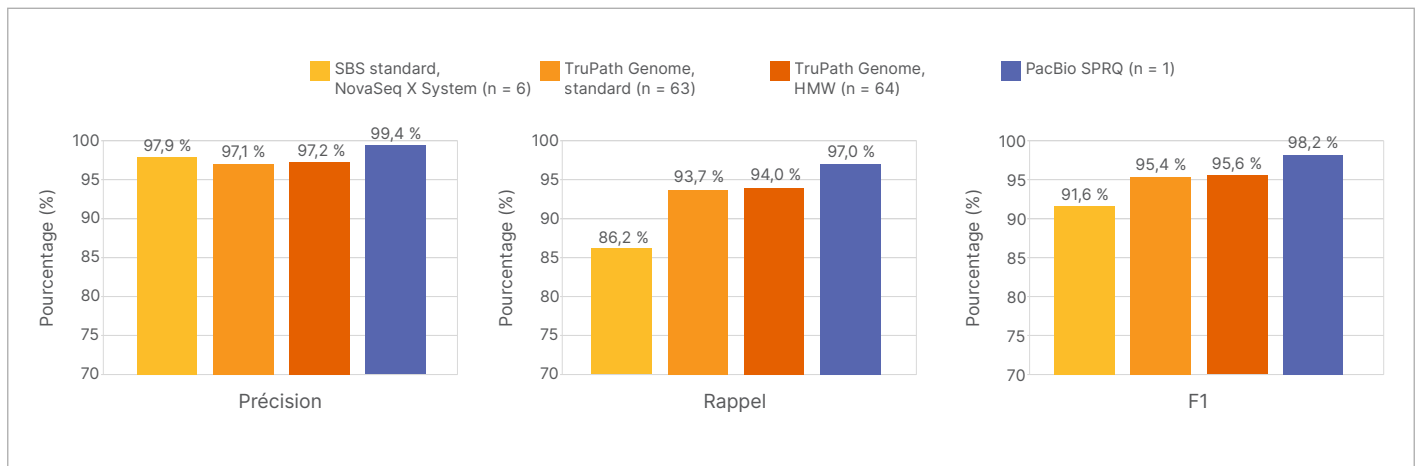


Figure 6 : Amélioration de l'appel des variants structuraux avec TruPath Genome

Performance d'appel des variants structuraux (VS) avec TruPath Genome comparée au SBS standard, utilisant DRAGEN Germline v4.5.2. L'analyse utilise l'ensemble vérifonctionnel Genome in a Bottle NIST T2T-Q100 HG002 SV v1.1 avec le fichier BED des variants structuraux à haute confiance. L'analyse comparative a été réalisée conformément aux [recommandations de Genome in a Bottle pour l'analyse comparative des variants structuraux](#) en utilisant les commandes « bench » et « refine » de Truvari v4.2.2. Les données SBS standard ont été générées à partir de bibliothèques Illumina DNA PCR-Free Prep, séquencées sur NovaSeq X System avec le logiciel v1.4 et la Flow Cell 10B (six réplicats techniques). TruPath Genome : les données standard ont été générées à partir d'ADN extrait selon des méthodes standard, séquencées sur NovaSeq X System avec le logiciel v1.4 et la Flow Cell C8 (63 réplicats techniques). Le TruPath Genome HMW a été généré à partir d'ADN extrait selon des méthodes HMW, séquencé sur NovaSeq X System avec le logiciel v1.4 et la Flow Cell C8 (64 réplicats techniques). PacBio avec chimie SPRQ provenant de sources publiées⁶. Les mesures de précision correspondent à la proportion de vrais positifs. Le rappel représente le pourcentage de vrais négatifs par rapport au nombre total de variants dans un échantillon. Le score F1 (%) est un calcul des résultats vrais positifs et vrais négatifs par rapport aux résultats totaux. Des scores plus élevés indiquent une meilleure précision par rapport aux données de référence.

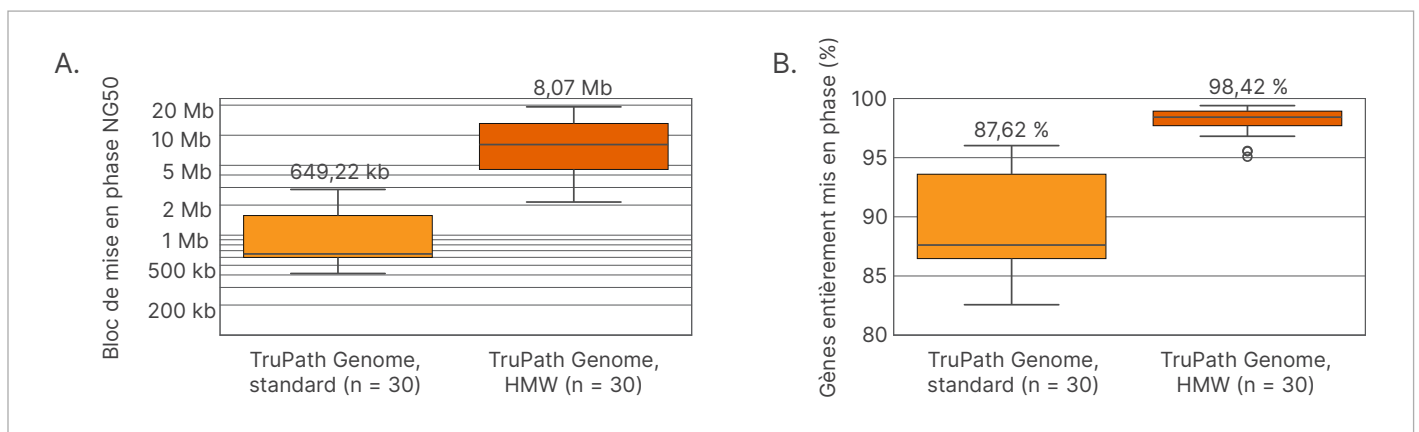


Figure 7 : Mise en phase ultralongue avec TruPath Genome sur 30 lignées cellulaires appariées, issues d'extractions selon la méthode HMW et la méthode standard

Le bloc de phase est mesuré sur les chromosomes 1 à 22 à l'aide de la mise en phase DRAGEN dans DRAGEN Germline v4.5.2. Les données TruPath Genome ont été obtenues à partir de 30 échantillons de lignées cellulaires Coriell, avec des extractions à la fois de poids moléculaire élevé (HMW) et standard. Le bloc de mise en phase NG50 correspond à la longueur du bloc de mise en phase lorsque 50 % de la région cible (génomique ou autre) a été mise en phase. Une technologie incapable de mettre en phase 50 % d'une région cible donnée aura un NG50 de zéro paire de bases. Le pourcentage de gènes entièrement mis en phase correspond à la proportion de régions géniques issues d'une liste de gènes spécifiée (Gencode v44 genes.gtf) qui sont entièrement contenues dans un seul bloc de mise en phase.

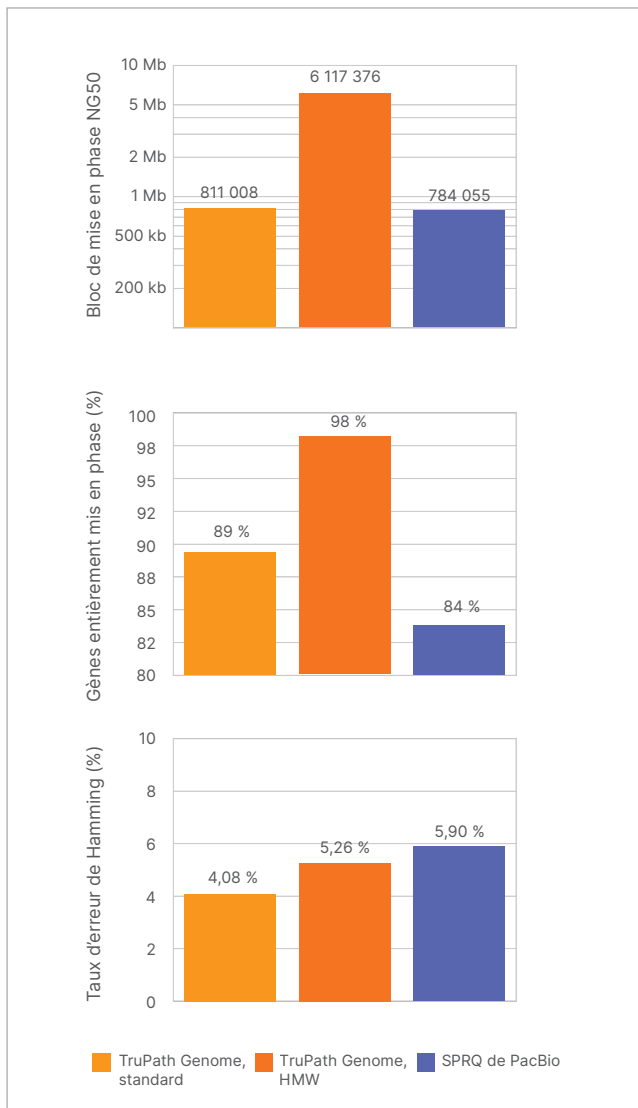


Figure 8 : Séquençage par mise en phase réalisé avec TruPath Genome, évalué à l'aide des données de l'ensemble vérifonctionnel HG002

Le bloc de phase est mesuré sur les chromosomes 1 à 22 à l'aide de la mise en phase DRAGEN dans DRAGEN Germline v4.5.2. Les données HiFi (PacBio SPRQ) avec VCF mis en phase ont été obtenues depuis https://downloads.pacbcloud.com/public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2/. Les données TruPath Genome ont été obtenues à partir de réplicats de séquençage des échantillons de la lignée cellulaire HG002, comprenant des extractions de poids moléculaire élevé (HMW) et des extractions standard. Le bloc de mise en phase NG50 correspond à la longueur du bloc de mise en phase lorsque 50 % de la région cible (génomique ou autre) a été mise en phase. Le pourcentage de gènes entièrement mis en phase correspond à la proportion de régions géniques issues d'une liste de gènes spécifiée (Gencode v44 genes.gtf) qui sont entièrement contenues dans un seul bloc de mise en phase. Le taux d'erreur de Hamming pour la mise en phase est évalué par rapport à l'ensemble vérifonctionnel T2T Q100⁹ téléchargé depuis https://ftp-trace.ncbi.nlm.nih.gov/ReferenceSamples/giab/data/AshkenazimTrio/analysis/NIST_HG002_DraftBenchmark_defrabbV0.020-20250117/GRCh38_HG2-T2TQ100-V1.1_smvar.vcf.gz.

TruPath Genome réalise la mise en phase complète d'une médiane d'environ 91 % de tous les gènes avec une extraction d'ADN standard et d'environ 97 % des gènes avec une extraction d'ADN de poids moléculaire élevé (HMW). De plus, TruPath Genome réalise la mise en phase d'environ 98 %[¶] de tous les SNV hétérozygotes, que l'extraction d'ADN soit standard ou HMW.

Les données TruPath Genome avec mise en phase DRAGEN pour l'ADN HMW réduisent les erreurs de Hamming** (~ 4 %) par rapport aux outils disponibles publiquement, comme HapCut2 (~ 6 %)[¶].

Détection des variants dans les régions à forte homologie

La détection précise des variants dans les régions paralogueues à l'aide de lectures courtes standard est difficile en raison de la forte homologie des séquences, ce qui entraîne une cartographie ambiguë des lectures et des erreurs dans la détection des variants. La résolution nécessite souvent des tests réflexes coûteux, tels que l'amplification par sondes dépendante de ligature multiplex (MLPA, Multiplex Ligation-dependent Probe Amplification) ou l'amplification en chaîne polymérase longs fragments (LR-PCR). TruPath Genome combine les renseignements de proximité des Flow Cell avec un nouvel algorithme de détection combinée multirégion (MRJD, Multiregion Joint Detection) pour fournir un appel fiable, *de novo*, des variants résolus par haplotype dans les gènes paralogueues (figure 9 et figure 10).

TruPath Genome cible onze gènes (tableau 1) connus pour posséder des copies en tandem ou non réparties dans l'ensemble du génome afin d'assurer une cartographie haute précision avec MRJD, et fournit des reconstructions résolues par haplotype ainsi que l'appel des petits variants (figure 10).

<i>CFHR</i>	<i>NCF1</i>	<i>SMN1/2</i>
<i>CYP2D6</i>	<i>PMS2</i>	<i>STRC</i>
<i>CYP11B1</i>	<i>RCCX (CYP21A2 et TNXB)</i>	<i>USP8</i>

[¶] Le pourcentage de variants hétérozygotes mis en phase correspond au pourcentage de SNV hétérozygotes mis en phase, calculé comme le nombre de SNV mis en phase divisé par le nombre total de SNV hétérozygotes.

** Les erreurs de Hamming surviennent lorsque des variants sont attribués au brin maternel ou paternel incorrect.

Résolution améliorée des répétitions en tandem courtes (STR)

Les STR sont des régions répétitives du génome pouvant s'étendre au-delà de leur longueur normale et provoquer des mutations associées à de nombreuses maladies génétiques, notamment le syndrome de l'X fragile, la sclérose latérale amyotrophique et la maladie de Huntington¹¹. L'estimation précise de la taille et la récupération des lectures de ces régions ont été difficiles, car les allèles élargis les plus pertinents dépassent souvent la longueur des lectures des données de séquençage à lectures courtes standard. Grâce à une combinaison de renseignements de proximité et de données de lectures courtes de haute qualité, TruPath Genome permet une récupération améliorée des lectures nécessaires pour un dimensionnement précis des STR et une meilleure résolution grâce au génotypage des deux copies à l'aide des renseignements de mise en phase (figure 11).

Renseignements sur de très longues distances pour les VS complexes

La capacité de TruPath Genome à résoudre de grands réarrangements structuraux inclut des fonctionnalités inédites qui vont au-delà des références de performance traditionnelles pour l'appel des variants. TruPath Genome capture les renseignements sur les lectures provenant d'amplifiats proximaux^{††} entre n'importe quelle paire de régions du génome, permettant ainsi de créer des représentations visuelles haute résolution des cartes de structure du génome, appelées « graphiques de colocalisation » (figure 12).

†† Les amplifiats proximaux sont des amplifiats physiquement proches les uns des autres sur la Flow Cell.

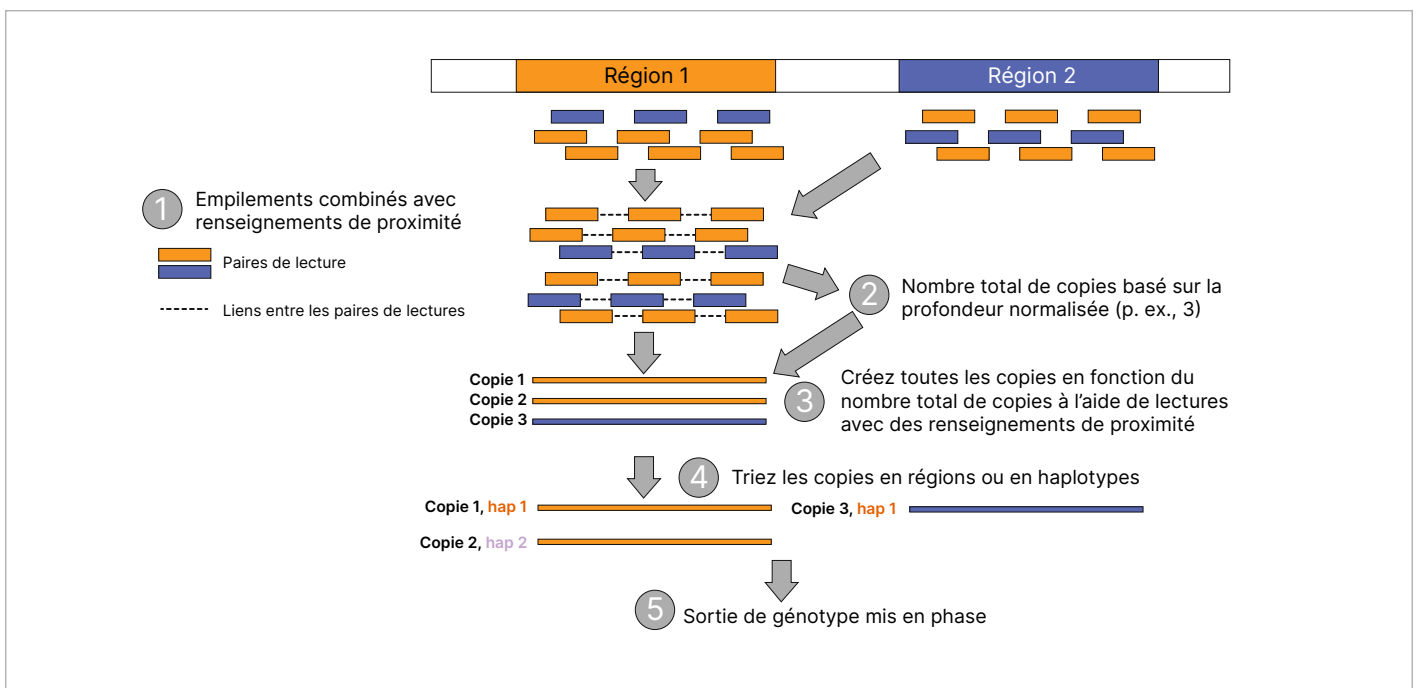
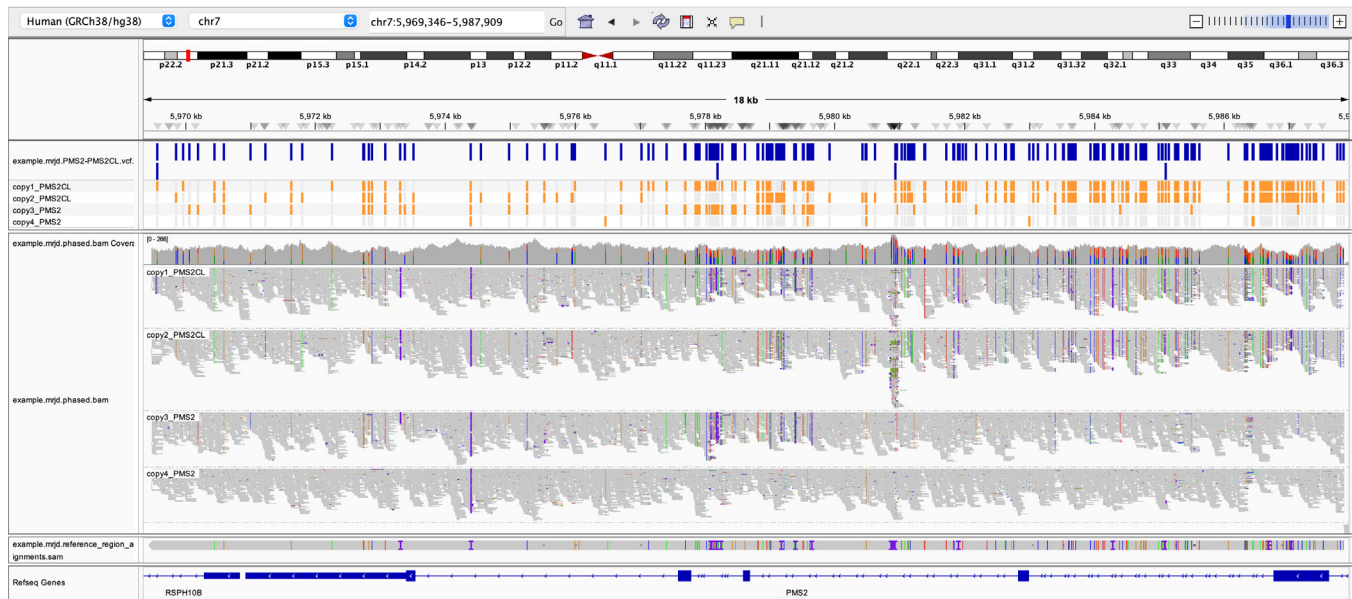


Figure 9 : Schéma de la méthode MRJD utilisée pour détecter les variants dans les régions de duplications segmentaires

1. Pour un ensemble donné de régions de duplications segmentaires, MRJD extrait toutes les lectures cartographiées sur l'une quelconque des copies de cet ensemble afin de les analyser conjointement. Les renseignements de proximité sont exploités pour établir des liens entre ces lectures et identifier les ensembles de lectures provenant de la même molécule d'origine; 2. Le nombre total de lectures extraites à l'étape 1 est ensuite normalisé, corrigé en fonction du contenu en GC, et utilisé pour estimer le nombre total de copies de la séquence dupliquée par segments dans l'échantillon d'entrée; 3. Une approche de génotypage conjoint, basée sur le nombre de copies de la région estimé à l'étape 2 et appliquée à l'ensemble des lectures liées par proximité, est utilisée pour identifier et mettre en phase tous les variants présents dans chaque copie de la séquence dupliquée par segments dans l'échantillon d'entrée. 4. Chaque copie entièrement mise en phase est assignée à un emplacement spécifique du génome de référence (lorsque les régions sont situées à des emplacements distaux dans la référence) ou mise en phase avec les autres copies pour former des haplotypes (ce qui n'est possible que lorsque les régions de duplication segmentaire sont en tandem ou proches dans la référence). 5. Les variants entièrement mis en phase sur toutes les copies, l'attribution des haplotypes pour chaque copie, l'attribution des copies à chaque emplacement du génome de référence, ainsi que les lectures assignées à chaque copie, sont exportés dans les fichiers VCF, JSON et BAM spécifiques au MRJD.

A. PMS2 gene



B. SMN gene

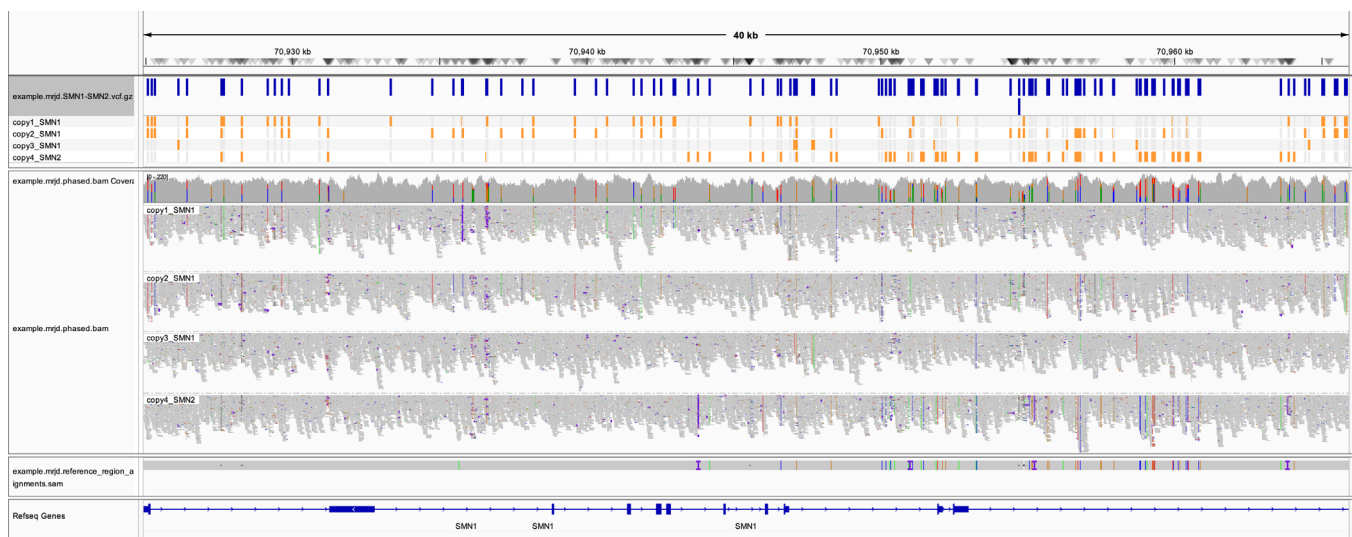


Figure 10 : Exemple de sortie des appels de variants et de l'attribution des lectures dans des gènes paralogues utilisant TruPath Genome et MJRD

(A) Appels des variants mis en phase et lectures de support associées couvrant toutes les copies de la paire gène-pseudogène *PMS2/PMS2CL*, représentées à l'emplacement du génome de référence pour *PMS2*. Deux des quatre copies sont identifiées comme des copies *PMS2CL* (pseudogène) et deux comme copies *PMS2* (gène actif), sur la base des liens de proximité avec les bords *PMS2/PMS2CL*. La copie3 (*PMS2*) présente ce qui semble être un événement de conversion génique chevauchant les exons 13 et 14, indiqué par la présence de variants correspondant à la séquence de référence *PMS2CL*. L'alignement entre la séquence de référence *PMS2CL* et la séquence de référence *PMS2* est représenté dans le troisième volet (de haut en bas). (B) Appels des variants mis en phase et lectures de support associées couvrant toutes les copies de la paire gène-pseudogène *SMN1/SMN2*, représentées à l'emplacement du génome de référence *SMN1*. Trois des quatre copies sont identifiées comme *SMN1* (gène actif) et deux comme *SMN2* (gène inactif), en fonction de la base présente à la position c.840 de *SMN1* dans chacune des copies mises en phase (les copies *SMN1* ont une base C et les copies *SMN2* ont une base T à cette position).

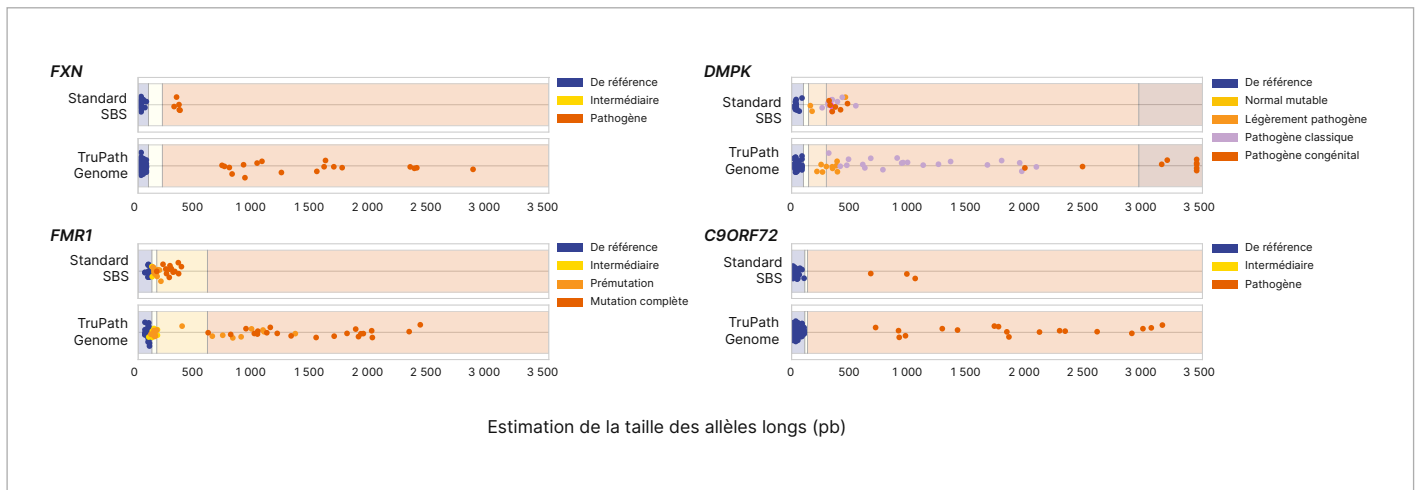


Figure 11 : Des mesures plus précises des expansions de STR peuvent être obtenues en utilisant TruPath Genome

Les classifications réelles des STR sont représentées par la couleur de chaque point. Les bandes colorées représentent les différentes plages de classification utilisées pour définir les différentes classes d'expansions à chaque locus. L'axe X indique l'estimation de la taille des STR obtenue avec DRAGEN STR, en utilisant comme données d'entrée soit TruPath Genome, soit un SBS standard (Illumina DNA PCR-Free Prep). Les classifications obtenues avec TruPath Genome sont nettement plus cohérentes avec les classifications réelles et couvrent une gamme plus large de longueurs de STR. Nombre total de réplcats pour le SBS standard : 78 échantillons biologiques (78 individus uniques). Nombre total de réplcats pour TruPath Genome : 135 échantillons biologiques (42 individus uniques).

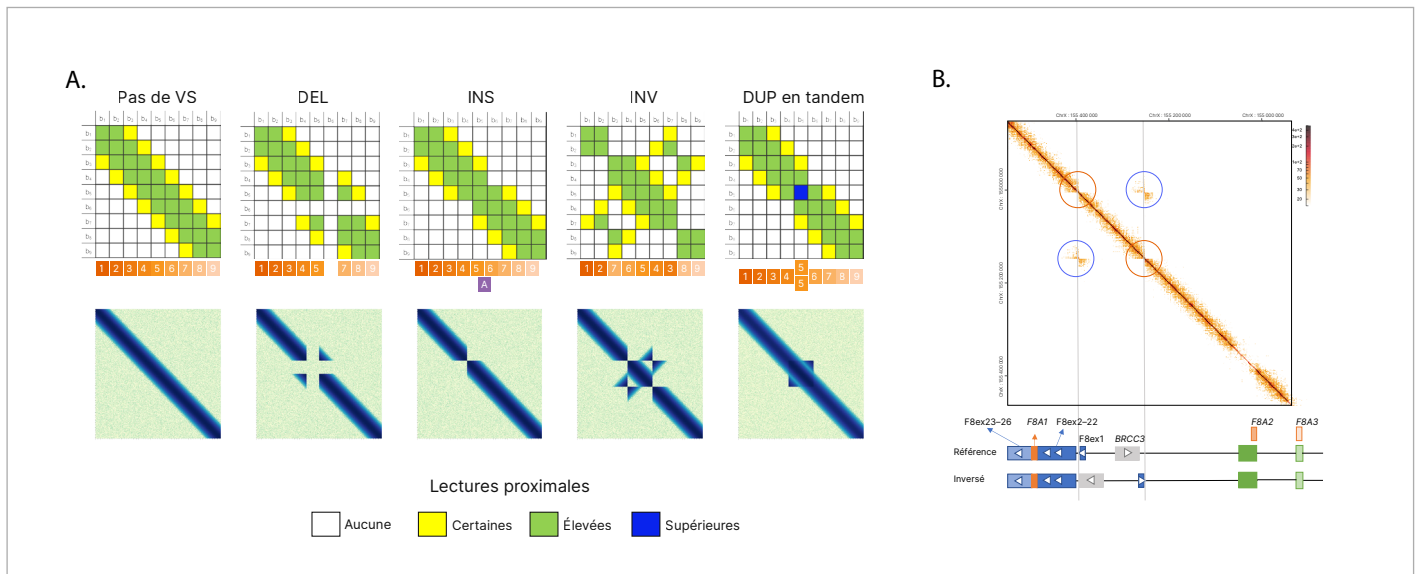


Figure 12 : Visualisation des événements de variants structuraux à l'aide de graphiques de colocalisation TruPath Genome

(A) Les nombres de lectures proximales (dans l'espace de la Flow Cell) cartographiées sur toutes les paires de compartiments génomiques sont évalués afin de générer une carte de colocalisation en deux dimensions. Les axes X et Y représentent les compartiments du génome de référence. Un grand nombre de connexions entre deux compartiments indique que ces compartiments sont proches les uns des autres dans l'espace génomique de l'échantillon d'entrée. Dans une région sans variants structuraux, les compartiments génomiques proches dans le génome de référence sont également proches dans l'échantillon analysé et apparaissent donc sous forme d'un grand nombre de connexions près de la diagonale du graphique de colocalisation. Lorsqu'un variant structural est présent, des compartiments génomiques qui sont proches dans le génome de référence peuvent ne plus être proches dans le génome de l'échantillon d'entrée, et des compartiments génomiques éloignés dans la référence peuvent se retrouver voisins dans l'échantillon. Ces changements de connectivité se reflètent sur les cartes de colocalisation sous forme de signaux hors diagonale de différentes formes, ainsi que par l'absence de connectivité diagonale. (B) Exemple de carte de colocalisation pour une inversion du gène dans l'intron 1 au niveau du locus du gène *FB8*, associée à une hémophilie A sévère. On remarque le signal en forme de sablier attendu dans la région hors diagonale (cercles bleus) et l'absence de signal diagonal (cercles orange) aux limites de l'événement. Une représentation schématique du locus dans le génome de référence et de la région génique inversée est présentée sous le graphique.

Le flux de travail le plus simple pour le séquençage du génome entier

Illumina TruPath Genome, propulsé par la technologie de lecture à cartographie de proximité, révolutionne le séquençage du génome humain. Le flux de travail innovant élimine la préparation de librairie traditionnelle pour une simplicité sans précédent, tandis que les renseignements de proximité conservent les données des brins d'ADN sur de longues distances avec des données de séquençage à lectures courtes standard hautement précises. Le résultat est un test qui combine les atouts de la qualité des données des lectures courtes, la détection très précise des petits variants, avec une couverture cartographique améliorée du génome humain et un outil puissant pour la compréhension des maladies génétiques.

En savoir plus →

[Illumina TruPath Genome](#)

[Technologie de lecture à cartographie de proximité](#)

[Série NovaSeq X](#)

Renseignements relatifs à la commande	
Produit	N° de référence
Illumina TruPath Genome (Flow Cell NovaSeq X C8)	20157405
Illumina TruPath Genome (Flow Cell NovaSeq X C2)	20157406

Références

1. Ebbert MTW, Jensen TD, Jansen-West K, et al. [Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight](#). *Genome Biol.* 2019;20(1):97. Publié le 20 mai 2019. doi :10.1186/s13059-019-1707-2
2. Ryan NM, Corvin A. [Investigating the dark-side of the genome: a barrier to human disease variant discovery?](#) *Biol Res.* 2023;56(1):42. Publié le 20 juillet 2023. doi :10.1186/s40659-023-00455-0
3. Illumina. TruPath Genome performance with samples of varying type and quality. [illumina.com/content/dam/illumina-marketing/documents/products/technotes/trupath-genome-sample-info-tech-note-m-gl-03932](https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/trupath-genome-sample-info-tech-note-m-gl-03932). Publié en février 2026. Consulté le XX février 2026.
4. Wadsworth ME, Page ML, Aguzzoli Heberle B, Miller JB, Steely C, Ebbert MTW. [Sequencing the gaps: dark genomic regions persist in CHM13 despite long-read advances](#). Prépublication. *bioRxiv.* 2025;2025.05.23.655776. Publié le 28 mai 2025. doi :10.1101/2025.05.23.655776
5. Ultima Genomics. Genome-in-a-Bottle (GIAB) samples HG001-HG005. <https://cdn.sanity.io/files/i7780ks7/production-2024/0a1b6a62a6da3e3fcfb81cad4c8ff2ffe85dd41.pdf>. Publié en février 2025. Consulté le 29 janvier 2026.
6. PacBio. Index of /public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2. https://downloads.pacbcloud.com/public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2/. Consulté le 29 janvier 2026.
7. Asri M, Chang PC, Mier JC, et al. [Pangenome-aware DeepVariant](#). Prépublication. *bioRxiv.* 2025;2025.06.05.657102. Publié le 6 juin 2025. doi :10.1101/2025.06.05.657102
8. Browning SR, Browning BL. [Haplotype phasing: existing methods and new developments](#). *Nat Rev Genet.* 2011; 12(10):703-714. doi: 10.1038/nrg3054
9. Hansen NF, Dwarshuis N, Ji HJ, et al. [A complete diploid human genome benchmark for personalized genomics](#). Prépublication. *bioRxiv.* 2025;2025.09.21.677443. Publié le 21 septembre 2025. doi :10.1101/2025.09.21.677443
10. Données internes. Illumina, Inc., 2026.
11. Steely CJ, Watkins WS, Baird L, Jorde LB. [The mutational dynamics of short tandem repeats in large, multigenerational families](#). *Genome Biol.* 2022;23(1):253. Publié le 12 décembre 2022. doi :10.1186/s13059-022-02818-4



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

© 2026 Illumina, Inc. Tous droits réservés. Toutes les marques de commerce sont la propriété d'Illumina, Inc. ou de leurs détenteurs respectifs. Pour obtenir des renseignements sur les marques de commerce, consultez la page www.illumina.com/company/legal.html.
M-GL-03931 FRA v1.0