

ILLUMINA TruPath™ Genome

El flujo de trabajo de secuenciación
del genoma completo más sencillo



Experimente el flujo de trabajo de muestra a datos más sencillo con un tiempo de preparación activa de 10 minutos

TAGCA
CGTAG
ATAGC

Resuelva regiones difíciles de asignar del genoma con lecturas cortas y datos de proximidad con NovaSeq™ X Series



Descubra información de larga distancia con una detección mejorada de variantes estructurales y fase de hebra retrasada ultralarga

Accesibilidad genómica que cambia el paradigma

La secuenciación de nueva generación (NGS, next-generation sequencing) de lecturas cortas ha sido el método principal utilizado por los investigadores para analizar la mayor parte del genoma. Sin embargo, algunas regiones y tipos de variantes siguen siendo difíciles de caracterizar, como los que tienen una repetibilidad extrema, una alta homología de secuencias o una complejidad estructural. Según estudios recientes, estas regiones albergan variantes que podrían desempeñar un papel en la enfermedad genética humana,^{1,2} lo que hace que descifrarlas sea de suma importancia a la hora de buscar causas subyacentes de enfermedades genéticas.

Los métodos alternativos, como la secuenciación de lecturas largas, ofrecen una resolución mejorada, lo que mejora la información sobre muchas de estas regiones y variantes difíciles de asignar. Sin embargo, los métodos de lectura larga tienden a tener flujos de trabajo complejos con problemas de precisión bien conocidos que dan lugar a resultados variables.

Illumina TruPath Genome con [tecnología de lectura asignada por proximidad](#) cambia el paradigma de los flujos de trabajo de secuenciación del genoma completo (WGS, whole-genome sequencing). Este innovador ensayo captura fragmentos de ADN largos con un sencillo paso de preparación activa de 10 minutos. La secuenciación se produce en NovaSeq X Series existente con el innovador análisis secundario DRAGEN™.

El resultado es un genoma preciso y completo que permite a los investigadores generar información que abarca distancias mayores que las longitudes de lectura generadas normalmente mediante métodos de lectura larga.

Ahora con TruPath Genome es posible aprovechar la facilidad y la precisión de la secuenciación de lectura corta para resolver regiones difíciles de asignar del genoma, mejorar la detección de variantes estructurales (SV) y generar lecturas y llamadas de variantes en fase de hebra retrasada.

Simplicidad sin precedentes del flujo de trabajo

TruPath Genome simplifica significativamente la implementación de la WGS y reduce el tiempo de participación activa a 10 minutos ([figura 1](#)). La fiable química de transposasa de Illumina elimina los pasos tradicionales de preparación de librerías. Los algoritmos bioinformáticos avanzados aprovechan la información sobre la proximidad entre nanopocillos con datos estándar de lectura corta "paired-end" para determinar las relaciones entre variantes separadas por hasta millones de bases.

* Los datos de lectura corta "paired-end" estándar hacen referencia a los datos generados mediante secuenciación del genoma completo realizada con métodos manuales de preparación de librerías y métodos estándar de secuenciación por síntesis (SBS, sequencing by synthesis).

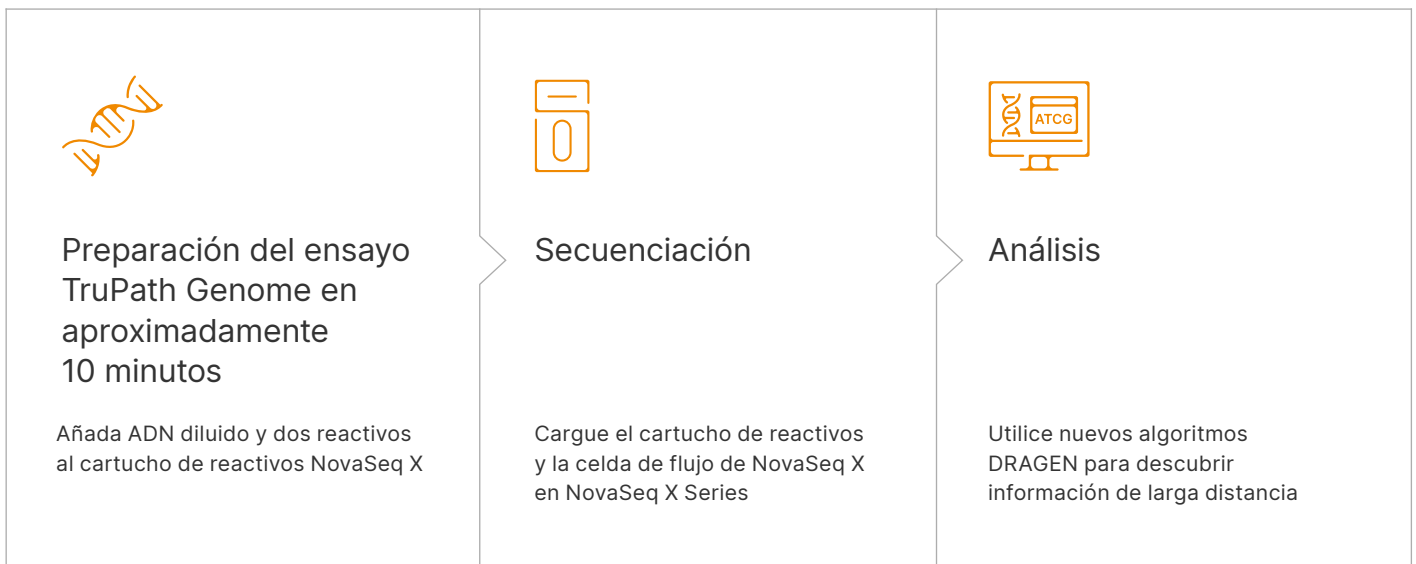


Figura 1: Flujo de trabajo de TruPath Genome

El ADN y un reactivo de TruPath Genome se añaden a la gradilla de tubos de la librería, que luego se inserta en el cartucho de reactivos NovaSeq X. A continuación, se añaden más reactivos de TruPath Genome al cartucho. El cartucho y la celda de flujo se cargan en NovaSeq X Series y se inicia el experimento.

Muestras

El ADN de cadena doble (ADNcd) purificado se extrae con kits comerciales estándar o de elevado peso molecular (EPM). El aporte recomendado de ADN de 350 ng se diluye con tampón de tagmentación TruPath y se añade a las gradillas de tubos de la librería. El ADN de aporte puede aislarse de estirpes celulares o sangre completa. No se admite ADN aislado de muestras de ADN fetal libre circulante (cfDNA, cell-free DNA) y ADN fijado en formol y embebido en parafina (FFPE, formalin fixed, paraffin-embedded). Para obtener un rendimiento óptimo, se recomienda extraer ADN mediante un método que conserve fragmentos grandes, de modo que el 40 % tenga una longitud superior a 60 kb.³

Para obtener orientación sobre el uso de tipos de muestras adicionales, la variabilidad en la calidad de la muestra y las cantidades de aporte óptimas, lea el [rendimiento de TruPath Genome con varios tipos de muestras, la calidad y la nota técnica de entrada](#).

Secuenciación

La gradilla de tubos de la librería con ADN diluido y dos reactivos TruPath Genome se añaden al cartucho de reactivos NovaSeq X y se cargan en NovaSeq X Series junto con una celda de flujo NovaSeq X C2 o C8 para la secuenciación de dos u ocho genomas humanos, respectivamente.

Los nanopocillos de la superficie de la celda de flujo se ceban con transposomas de Illumina.[†] El ADN se introduce en la celda de flujo, donde lo capturan los transposomas en varios pocillos, lo que crea un patrón similar a una constelación de nanopocillos sembrados ([figura 2A](#)). Los transposomas segmentan el ADN capturado y unen los fragmentos a la superficie de la celda de flujo en un proceso llamado tagmentación[‡] ([figura 2B](#)). Mediante un paso de lavado, se elimina el ADN no unido y los transposomas.

Dentro de los nanopocillos, el ADN fragmentado por tagmentación se utiliza para crear grupos[§] para la secuenciación. El ADN del mismo fragmento original de ADN largo formará grupos en nanopocillos cercanos. La secuenciación se produce en NovaSeq X Series mediante la química de secuenciación por síntesis (XLEAP-SBS™) Illumina XLEAP, de eficacia demostrada, y un experimento de 2 × 150 pb.

Análisis

Una vez finalizada la secuenciación, el proceso DRAGEN™ Germline utiliza un novedoso algoritmo que combina los datos de secuenciación con la información sobre la proximidad entre nanopocillos para alinear el ADN con un genoma de referencia. Aprovechar la información sobre la proximidad permite la asignación adecuada de las lecturas que anteriormente no se asignaban al genoma de referencia o que solo podían alinearse con poca confianza.

[†] Los transposomas son complejos de transposasa de ADN que existen como dímeros.
[‡] La tagmentación es el proceso de cortar un fragmento de ADN y añadir una secuencia de adaptadores (etiquetado) a través de un transposoma.
[§] Un grupo es un punto amplificado de ADN en una celda de flujo que se secuenciará.

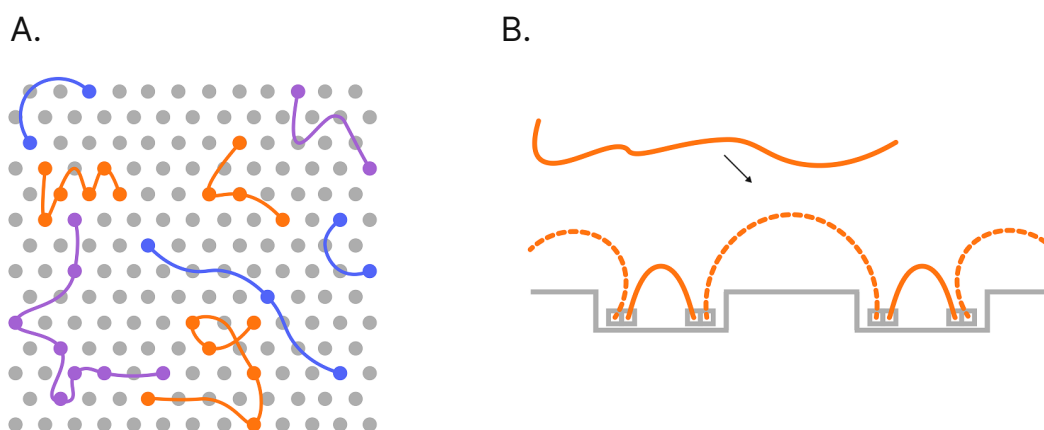


Figura 2: Cómo funciona la tecnología de lectura asignada por proximidad

(A) El ADN fluye a través de varios nanopocillos adyacentes, lo que permite el uso de información sobre proximidad para generar información de larga distancia. (B) El ADN se captura mediante transposomas en los nanopocillos (casillas grises), donde se somete a un proceso de tagmentación y generación de grupos.

De esta manera, las lecturas cortas estándar pueden asociarse entre sí de forma probabilística, lo que permite una detección exhaustiva de variantes pequeñas (variantes de nucleótido único [SNV] e inserciones/delecciones [indel]), variantes grandes (más de 50 pb), regiones homólogas conocidas y llamadas de variantes pequeñas de fase de hebra retrasada que pueden separarse por millones de pares de bases.

Descubra un genoma completo con lecturas cortas de alta precisión

La secuenciación por síntesis (SBS) estándar ha demostrado una detección de elevada precisión para variantes pequeñas (p. ej., SNV, indels y variantes en el número de copias [CNV, copy-number variants]), pero ha tenido dificultades con regiones y tipos de variantes específicos, como las variantes estructurales. TruPath Genome introduce nuevas capacidades para la secuenciación de lectura corta que permiten un genoma más completo. Ahora se pueden utilizar lecturas cortas para abordar regiones del genoma que antes eran difíciles y proporcionar una mayor precisión con tipos de variantes adicionales.

Cobertura mejorada en regiones difíciles de asignar

Las regiones genómicas complejas, altamente polimórficas o duplicadas pueden resultar difíciles de ensamblar o alinear con un genoma de referencia debido a su baja capacidad de asignación, ya sea por una cobertura insuficiente o por una calidad de asignación (MapQ, mapping quality) deficiente. Anteriormente, la cobertura de estas regiones "oscuras"¹⁴ ha dado lugar a puntuaciones bajas en la calidad de asignación (MapQ), lagunas en los datos de secuencias del genoma y resolución global limitada para la llamada de variantes. Para solucionar esto, TruPath Genome utiliza información sobre la proximidad entre grupos adyacentes para asignar la ubicación genómica correcta. De esta forma, se pueden efectuar asignaciones muy fiables de estas lecturas asignadas de forma ambigua (figura 3 y figura 4).

Llamada de variantes de mayor precisión

Se comparó TruPath Genome con el genoma HG002 y el conjunto de verdad T2T-Q100 v1.1 v0.019 del consorcio Telomere-to-Telomere (T2T) para demostrar el rendimiento en la llamada de variantes pequeñas, la detección de SV y la secuenciación en fase de hebra retrasada. El aumento de la resolución de la asignación gracias a la tecnología de lectura asignada por proximidad mejoró el rendimiento de las llamadas de variantes pequeñas, sobre todo en regiones del genoma tradicionalmente difíciles de asignar, como es el caso de las regiones homólogas o repetitivas.

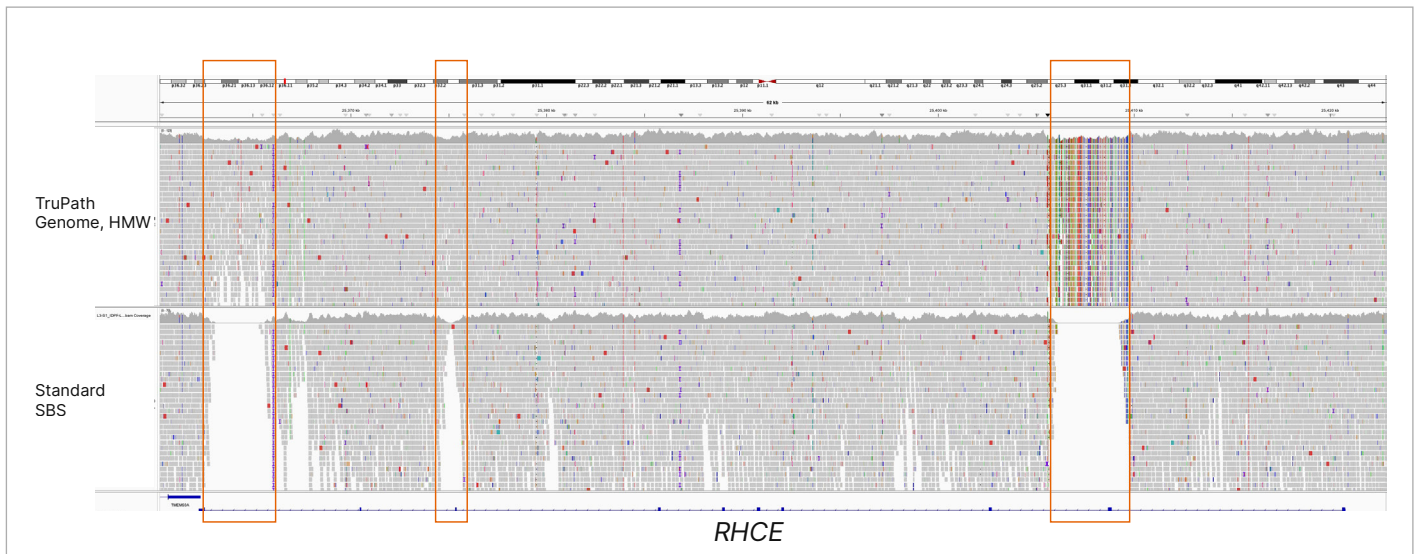


Figura 3: Resuelve eficazmente los genes de relevancia médica difíciles de asignar con TruPath Genome

Recuperación de la cobertura en el gen *RHCE*, que se encuentra en el locus del grupo sanguíneo Rh adyacente a *RHD* y comparte una homología de secuencia muy alta con la conversión frecuente de genes y reordenaciones híbridas, lo que hace que la asignación de lecturas cortas resulte ambigua en algunos puntos. Un genotipado preciso del *RHCE* es clínicamente importante para determinar los antígenos Rh (C/c, E/e), ya que las discrepancias pueden aumentar el riesgo de reacciones hemolíticas a las transfusiones y de enfermedad hemolítica en el feto/recién nacido (HDFN, hemolytic disease of the fetus/newborn). La SBS estándar se realizó con Illumina DNA PCR-Free Prep.

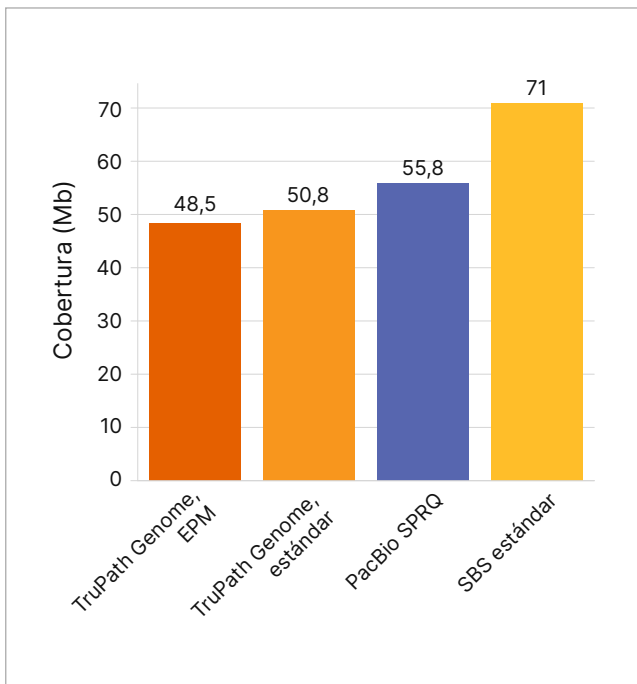


Figura 4: TruPath Genome mejora la cobertura de "regiones oscuras" del genoma

TruPath Genome mejora la cobertura de "regiones oscuras" definidas como las regiones "oscuras por MAPQ" en Ebbert et al.,¹ en las que el 90 % de las lecturas que cubren la región tienen una calidad de asignación (MAPQ) inferior a 10. TruPath Genome redujo el tamaño del genoma oscuro debido a una mejor asignación de lecturas en regiones complejas del genoma. La SBS estándar se realizó con Illumina DNA PCR-Free Prep.

En comparación con la secuenciación de lecturas cortas estándar y las llamadas de variantes, TruPath Genome muestra una mayor precisión mediante una reducción sustancial de las llamadas de variantes de falsos negativos (FN) y falsos positivos (FP). (Figura 5).

TruPath Genome muestra un aumento drástico en el rendimiento de las llamadas de SV (>50 pb) en comparación con los métodos de lectura corta estándar. La retirada de SV aumentó del 86 % con SBS estándar al 94 % con tecnología de lectura asignada por proximidad (figura 6).

Fase de hebra retrasada ultralarga

La **secuenciación en fase de hebra retrasada** proporciona información sobre los haplotipos que permite distinguir entre los alelos de los cromosomas maternos y paternos.⁸ Esta información puede ayudar a los investigadores a resolver variantes heterocigóticas compuestas, importante para interpretar las afecciones autosómicas recesivas.

La tecnología de lectura asignada por proximidad es especialmente adecuada para la fase de hebra retrasada del genoma humano. El ADN capturado en la celda de flujo produce bloques ordenados que abarcan desde cientos de kilobases hasta varias megabases (figura 7 y figura 8), lo que contribuye a una mayor información sobre los haplotipos y heterocigotos compuestos.

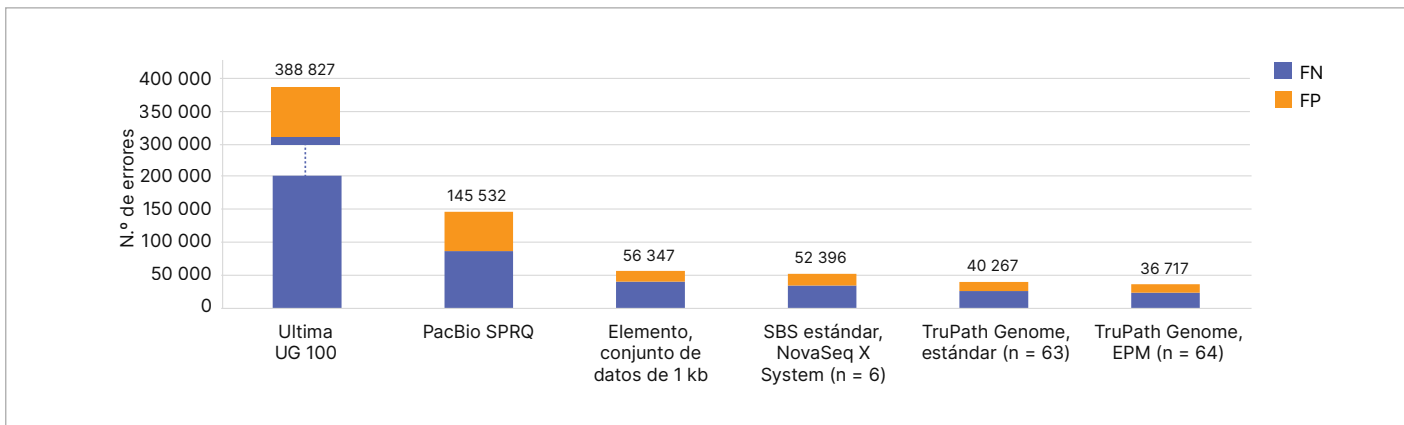


Figura 5: Mayor precisión para la llamada de variantes pequeñas con TruPath Genome

El rendimiento de las llamadas de variantes pequeñas en varios sistemas y ensayos de NGS se comparó con el conjunto de verdad Genome in a Bottle NIST T2T-Q100 HG002 SV v1.1. Los datos de Ultima UG 100, PacBio con química SPRQ y Elemento AVITI con un conjunto de datos de 1 kb se obtuvieron de materiales publicados.⁵⁻⁷ Los datos de SBS estándar se generaron con librerías Illumina DNA PCR-Free Prep secuenciadas en un NovaSeq X System con software v1.4 y la celda de flujo 10B y se analizaron con DRAGEN Germline v4.5.2 (seis réplicas técnicas). Los datos estándar de TruPath Genome se generaron con ADN extraído mediante métodos estándar y se secuenciaron en un NovaSeq X System con software v1.4 y la celda de flujo C8 y se analizaron con DRAGEN Germline v4.5.2 (63 réplicas técnicas). Los datos de EPM de TruPath Genome se generaron con ADN extraído mediante métodos de EPM y se secuenciaron en un NovaSeq X System con software v1.4 y la celda de flujo C8, y se analizaron con DRAGEN Germline v4.5.2 (64 réplicas técnicas). FN: llamadas de variantes con falsos negativos; FP: llamadas de variantes con falsos positivos.

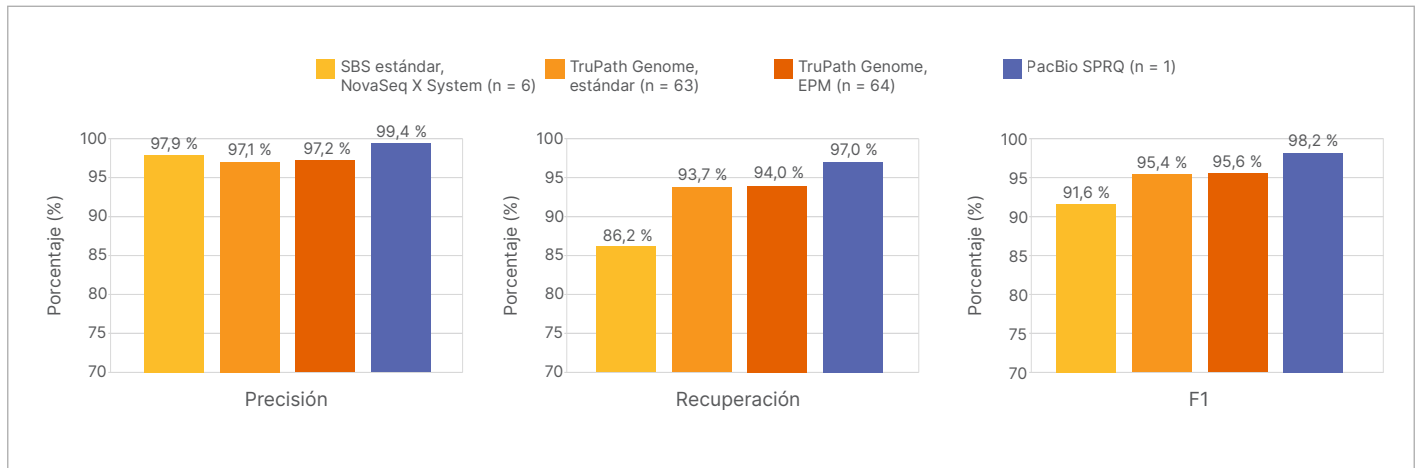


Figura 6: Llamada de variantes estructurales mejorada con TruPath Genome

Rendimiento de la variante estructural (SV) de TruPath Genome en comparación con la SBS estándar utilizando DRAGEN Germline v4.5.2. El análisis utiliza el conjunto de verdad Genome in a Bottle NIST T2T-Q100 HG002 SV v1.1 con el archivo BED seguro de SV. El análisis comparativo se realizó de acuerdo con las [directrices de Genome in a Bottle para el análisis comparativo de variantes estructurales](#) utilizando comandos "bench" y "refine" de Truvari v4.2.2. Los datos de SBS estándar se generaron con librerías Illumina DNA PCR-Free Prep secuenciadas en un NovaSeq X System con software v1.4 y la celda de flujo 10B (seis réplicas técnicas). Los datos estándar de TruPath Genome se generaron con ADN extraído mediante métodos estándar y se secuenciaron en un NovaSeq X System con software v1.4 y la celda de flujo C8 (63 réplicas técnicas). Los datos de EPM de TruPath Genome se generaron con ADN extraído mediante métodos de EPM y se secuenciaron en un NovaSeq X System con software v1.4 y la celda de flujo C8 (64 réplicas técnicas). Datos de PacBio con química de SPRQ obtenidos de material publicado.⁶ La precisión mide la proporción de verdaderos positivos. La retirada representa el porcentaje de verdaderos negativos en comparación con el número total de variantes en una muestra. La puntuación F1 (%) es un cálculo de resultados positivos verdaderos y negativos verdaderos, representado como una proporción de los resultados totales. Las puntuaciones más altas indican una mayor precisión en función de los datos de referencia.

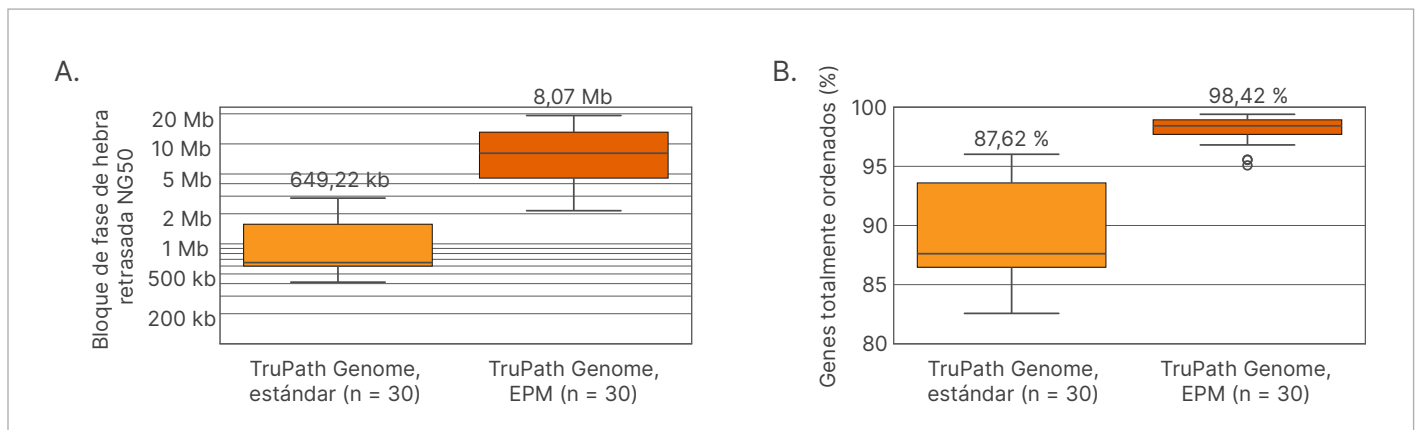


Figura 7: Fase de hebra retrasada ultralarga con TruPath Genome en 30 estirpes celulares de extracción estándar y de EPM emparejadas

El bloque de fase se mide sobre los cromosomas 1-22 mediante la fase de hebra retrasada de DRAGEN en DRAGEN Germline v4.5.2. Los datos de TruPath Genome se obtuvieron utilizando 30 muestras de Coriell de estirpes celulares diferentes con extracciones tanto de elevado peso molecular (EPM) como estándar. El bloque de fase NG50 es la longitud del bloque ordenado una vez ordenado el 50 % de la región objetivo (el genoma u otra). Una tecnología que no pueda secuenciar el 50 % de una región objetivo determinada tendrá un NG50 de pares de bases cero. El porcentaje de genes completamente ordenado es el porcentaje de regiones génicas de una lista de genes especificada (Gencode v44 genes.gtf) que están completamente contenidas en un único bloque de fase de hebra retrasada.

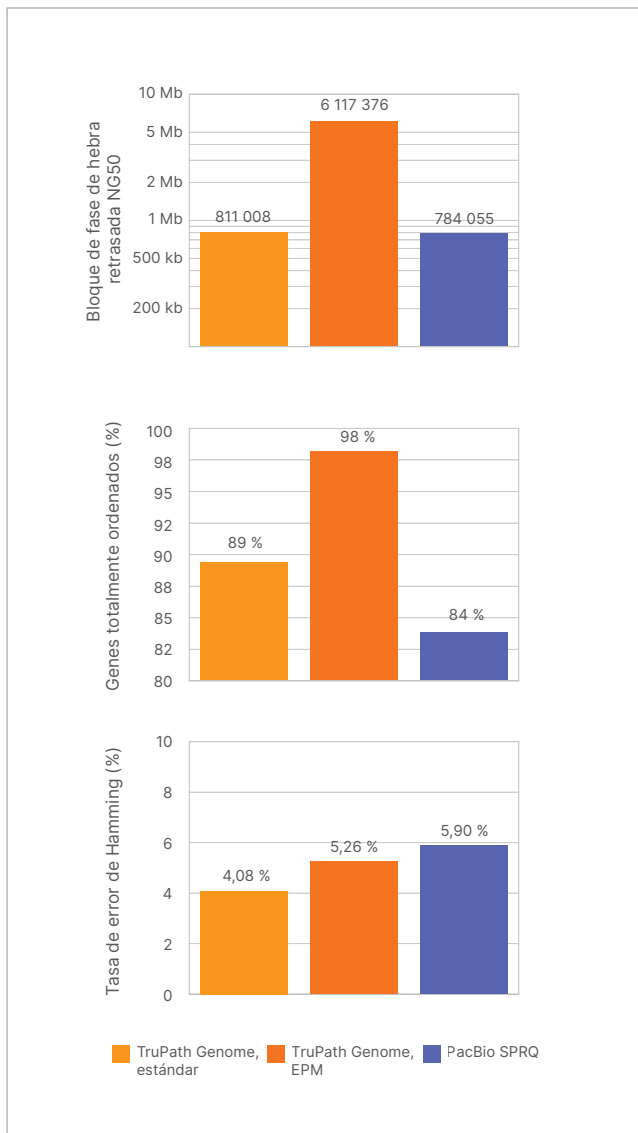


Figura 8: Secuenciación en fase de hebra retrasada lograda con TruPath Genome comparada con datos del conjunto de verdad HG002

El bloque de fase se mide sobre los cromosomas 1-22 mediante la fase de hebra retrasada de DRAGEN en DRAGEN Germline v4.5.2. El VCF ordenado de los datos de HiFi (PacBio SPRQ) se obtuvo de https://downloads.pacbcloud.com/public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2/. Los datos de TruPath Genome se obtuvieron mediante réplicas de secuenciación de muestras de estirpes celulares HG002 con muestras de extracción estándar y de EPM. El bloque ordenado NG50 es la longitud del bloque ordenado una vez ordenado el 50 % de la región objetivo (el genoma u otra). El porcentaje de genes completamente ordenados es el porcentaje de regiones génicas de una lista de genes especificada (Gencode v44 genes.gtf) que están completamente contenidas en un único bloque de fase de hebra retrasada. La tasa de error de Hamming de fase de hebra retrasada se compara con el conjunto de verdad T2T Q100⁹ descargado de 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 permite ordenar por completo una mediana de alrededor del 91 % de todos los genes con la extracción de ADN estándar y alrededor del 97 % de los genes con la extracción de ADN de elevado peso molecular (EPM). Además, TruPath Genome secuencia aproximadamente el 98 %[¶] de todas las SNV heterocigóticas tanto en extracciones de ADN estándar como de EPM.

Los datos de TruPath Genome con la fase de hebra retrasada de DRAGEN para ADN de EPM reducen los errores de Hamming** (~4 %) en comparación con herramientas disponibles públicamente como HapCut2 (~6 %).¹⁰

Detección de variantes en regiones de alta homología

La detección de variantes precisa en regiones parálogas mediante lecturas cortas estándar es compleja debido a la elevada homología de secuencias, lo que da lugar a errores ambiguos en la asignación de lecturas y la detección de variantes. La resolución suele exigir costosos ensayos secundarios, como la amplificación de sonda dependiente de ligadura múltiple (MLPA, multiplex ligation-dependent probe amplification) o la reacción en cadena de la polimerasa de largo alcance (LR-PCR, long-range polymerase chain reaction). TruPath Genome combina la información sobre la proximidad entre celdas de flujo con un nuevo algoritmo de detección de articulaciones multirregión (MRJD, Multiregion Joint Detection) para proporcionar una llamada de variantes fiable, *de novo* y resuelta por haplotipos en genes parálogos (figura 9 y figura 10).

TruPath Genome se centra de forma selectiva en once genes (tabla 1) de los que se sabe que tienen copias en tándem o no en tándem distribuidas por todo el genoma para realizar una asignación de alta precisión con MRJD y genera reconstrucciones resueltas por haplotipos y llamadas de variantes pequeñas (figura 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 y TNXB)</i>	<i>USP8</i>

¶ El porcentaje de variantes heterocigóticas ordenadas es el porcentaje de SNV heterocigóticas ordenadas, calculado como el número de SNV ordenadas dividido por el número de SNV heterocigóticas.

** Los errores de Hamming se producen cuando se asignan variantes a la cadena materna o paterna incorrecta.

Resolución mejorada para repeticiones cortas en tándem (STR)

Las repeticiones cortas en tándem (STR, short tandem repeats) son regiones repetitivas del genoma que pueden expandirse hasta alcanzar longitudes superiores al rango normal y provocar mutaciones asociadas a numerosas enfermedades genéticas, entre ellas el síndrome del cromosoma X frágil, la esclerosis lateral amiotrófica y la enfermedad de Huntington.¹¹ La estimación precisa del tamaño y la recuperación de la lectura de estas regiones ha sido difícil debido a que los alelos ampliados más relevantes suelen superar la longitud de lectura de los datos de secuenciación estándar de lectura corta. Mediante una combinación de información sobre la proximidad y datos de lectura corta de alta calidad, TruPath Genome mejora la recuperación de las lecturas necesarias para determinar con precisión el tamaño de los STR y ofrece una mayor resolución mediante el genotipado de las dos copias utilizando información ordenada (figura 11).

Información de distancia ultralarga para SV complejas

La capacidad de TruPath Genome para detectar grandes reordenaciones estructurales incluye nuevas funciones que van más allá de los parámetros de referencia tradicionales en cuanto al rendimiento de la llamada de variantes. TruPath Genome captura información sobre lecturas de grupos proximales^{††} entre cualquier par de regiones del genoma, lo que permite crear representaciones visuales de alta resolución de mapas de estructura del genoma, denominados "gráficos de coubicación" (figura 12).

†† Los grupos proximales son grupos que se encuentran físicamente cerca el uno del otro en la celda de flujo.

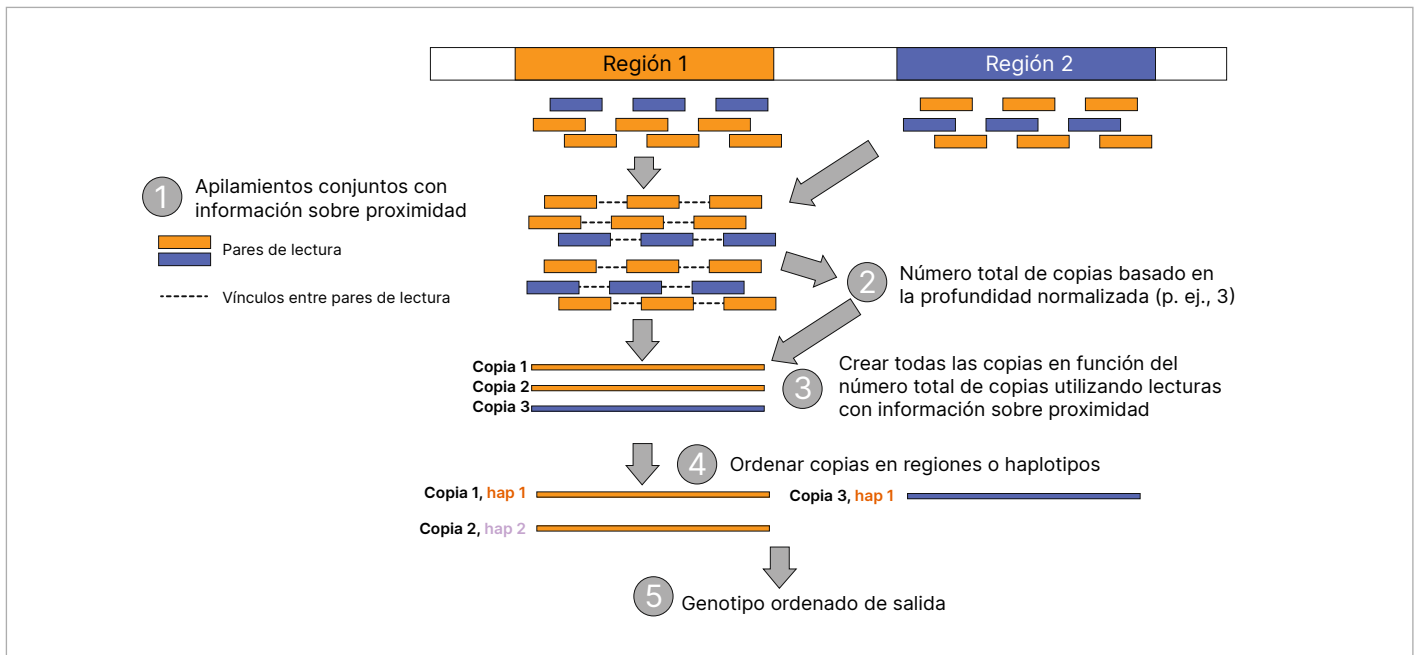
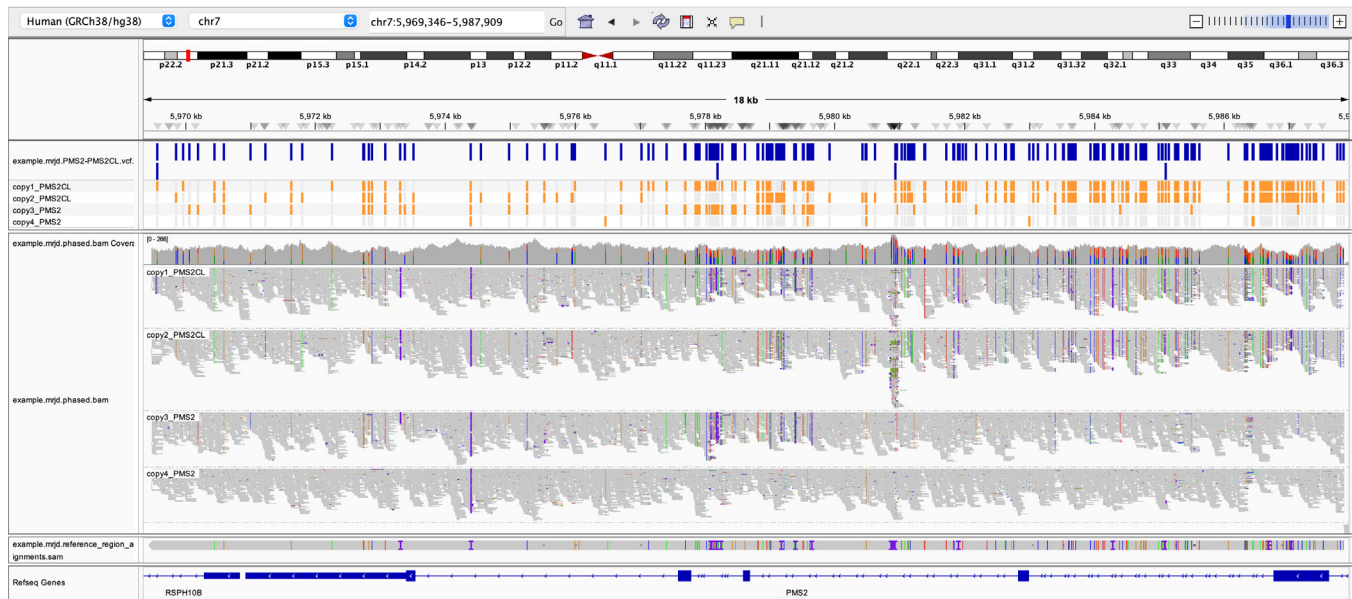


Figura 9: Esquema del método de MRJD utilizado para detectar variantes en regiones de duplicación segmental

1. Para un conjunto determinado de regiones duplicadas segmentariamente, MRJD extrae todas las lecturas asignadas a cualquiera de las copias de ese conjunto de regiones para su análisis conjunto. La información sobre proximidad se aprovecha para establecer vínculos entre dichas lecturas e identificar conjuntos de lecturas que se originen a partir de la misma molécula original. 2. El recuento de todas las lecturas extraídas en el paso 1 se normaliza, se corrige mediante GC y se utiliza para estimar el número total de copias de la secuencia duplicada segmentariamente en la muestra de entrada. 3. Se utiliza un enfoque de genotipado conjunto basado en el número de copias de la región estimada en el paso 2 y aplicado al conjunto completo de lecturas vinculadas por proximidad para identificar y ordenar todas las variantes de cada copia de la secuencia duplicada segmentariamente en la muestra de entrada. 4. Cada copia completamente ordenada se asigna a una ubicación específica del genoma de referencia (cuando las regiones están en ubicaciones de referencia distales) o se ordena con las otras copias en haplotipos (esto solo es posible cuando las regiones de duplicación segmental están en tándem/próximas en el control). 5. Las variantes completamente ordenadas en todas las copias, la asignación de haplotipos de cada copia, la asignación de copias a cada ubicación del genoma de referencia y las lecturas asignadas a cada copia se generan en los archivos VCF, JSON y BAM específicos de MRJD.

A. *PMS2* gene



B. *SMN* gene

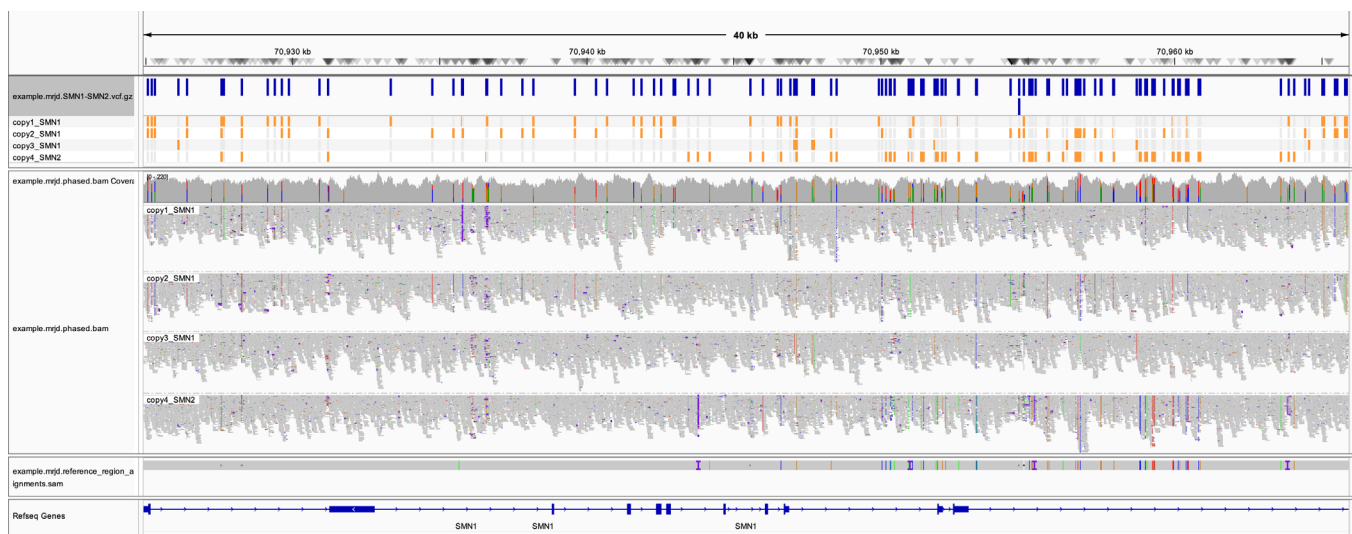


Figura 10: Resultado de ejemplo de llamadas de variantes y asignación de lecturas en genes parálogos mediante TruPath Genome y MJRD

(A) Llamadas de variantes en fase de hebra adelantada y lecturas de respaldo asociadas en todas las copias del par de gen-pseudogén *PMS2/PMS2CL* representadas en la ubicación del genoma de referencia de *PMS2*. Dos de las cuatro copias se identifican como copias del *PMS2CL* (pseudogén) y dos se identifican como copias de *PMS2* (gen activo) en función de los vínculos de proximidad a los flancos de *PMS2/PMS2CL*. Copy3 (*PMS2*) contiene lo que se parece a un evento de conversión génica que se solapa con los exones 13 y 14, indicado por la presencia de variantes que coinciden con la secuencia de referencia del *PMS2CL*. La alineación entre la secuencia de referencia del *PMS2CL* y la secuencia de referencia del *PMS2* se representa en el tercer panel (de arriba abajo). (B) Llamadas de variantes en fase de hebra adelantada y lecturas de respaldo asociadas en todas las copias del par de gen-pseudogén *SMN1/SMN2* representadas en la ubicación del genoma de referencia de *SMN1*. Tres de las cuatro copias se identifican como *SMN1* (gen activo) y dos se identifican como *SMN2* (gen inactivo) de acuerdo con la base en la posición c.840 de *SMN1* en cada una de las copias ordenadas (las copias de *SMN1* tienen la base C y las copias de *SMN2* tienen la base T en esa posición).

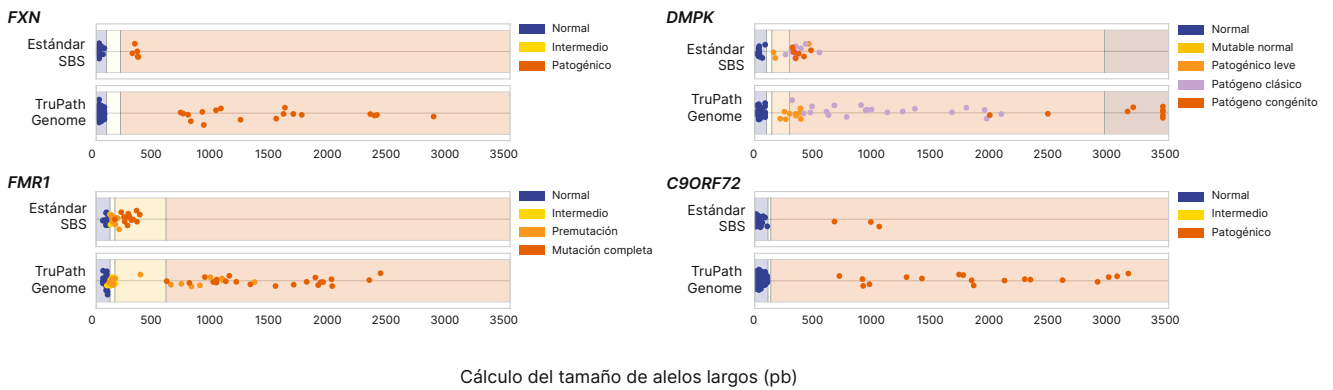


Figura 11: Se pueden obtener mediciones más precisas de las expansiones de STR con TruPath Genome

Las clasificaciones verdaderas de STR se representan por el color de cada punto. Las bandas de colores representan los diferentes rangos de clasificación utilizados para definir las diferentes clases de expansiones en cada locus. El eje X indica la estimación del tamaño de STR de DRAGEN STR utilizando TruPath Genome o SBS estándar (Illumina DNA PCR-Free Prep) como entradas. Las clasificaciones de TruPath Genome son significativamente más coherentes con las clasificaciones verdaderas y abarcan un intervalo más amplio de longitudes de STR. Total de réplicas para SBS estándar: 78 muestras biológicas (78 individuales únicas). Total de réplicas para TruPath Genome: 135 muestras biológicas (42 individuales únicas).

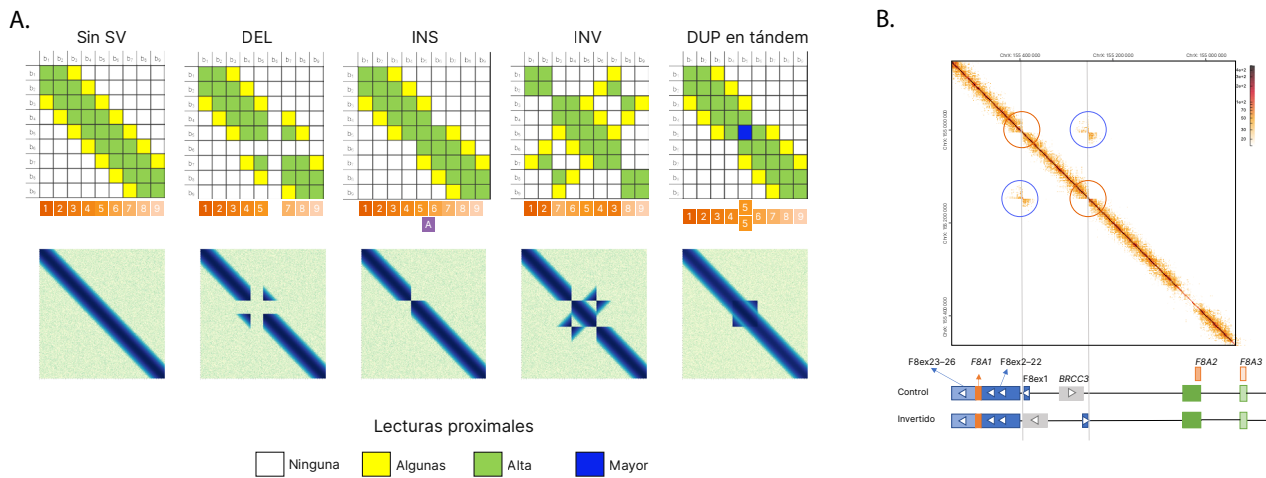


Figura 12: Visualización de eventos de variantes estructurales mediante gráficos de coubicación de TruPath Genome

(A) Los recuentos de lecturas proximales (espacio de la celda de flujo) que se asignan a todos los pares de grupos genómicos se evalúan para generar un mapa de coubicación bidimensional. Los ejes X e Y representan grupos del genoma de referencia. Un elevado número de conexiones entre dos grupos indica que estos se encuentran próximos entre sí en el espacio genómico de la muestra de entrada. En una región sin variantes estructurales, los grupos genómicos cercanos en el genoma de referencia también están cercanos en la muestra de entrada y, por lo tanto, aparecen como un gran número de conexiones cerca de la diagonal del gráfico de coubicación. Cuando hay presente una variante estructural, los grupos genómicos que se encuentran cerca en el genoma de referencia pueden no estar más cerca en el genoma de la muestra de entrada y los grupos genómicos que están lejos en la referencia pueden estar próximos en la muestra de entrada. Dichos cambios en la conectividad se reflejan en los mapas de coubicación como señales fuera de la diagonal de diferentes formas, así como en la ausencia de conectividad diagonal. (B) Un ejemplo de un mapa de coubicación para una inversión del gen del intrón 1 en el locus del gen *F8*, asociado a hemofilia A grave. Obsérvese la señal esperada en forma de reloj de arena en la región fuera de la diagonal (círculos azules) y la ausencia de señal diagonal (círculos naranjas) en los límites del evento. Debajo del gráfico se muestra una representación esquemática del locus en el genoma de referencia y la región genética invertida.

El flujo de trabajo de secuenciación del genoma completo más sencillo

Illumina TruPath Genome, con la tecnología fundamental de lectura asignada por proximidad fundacional, transforma la secuenciación del genoma humano. El innovador flujo de trabajo elimina la preparación de librerías tradicionales para lograr una simplicidad sin precedentes, mientras que la información sobre proximidad conserva la información de cadena de ADN de larga distancia con datos de secuenciación de lectura corta estándar de alta precisión. El resultado es un ensayo que combina las fortalezas de calidad de los datos de lecturas cortas, la detección de alta precisión de variantes pequeñas, mejor cobertura cartográfica del genoma humano y una potente herramienta para comprender la enfermedad genética.

Más información →

[Illumina TruPath Genome](#)

[Tecnología de lectura asignada por proximidad](#)

[NovaSeq X Series](#)

Datos para realizar pedidos	
Producto	N.º de catálogo
Illumina TruPath Genome (celda de flujo C8 de NovaSeq X)	20157405
Illumina TruPath Genome (celda de flujo C2 de NovaSeq X)	20157406

Bibliografía

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. Fecha de publicación: 20 de mayo de 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. Fecha de publicación: 20 de julio de 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). Fecha de publicación: febrero de 2026. Fecha de consulta: XX de febrero de 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](#). Preprint. *bioRxiv.* 2025;2025.05.23.655776. Fecha de publicación: 28 de mayo de 2025. doi:10.1101/2025.05.23.655776
5. Ultima Genomics. Genome-in-a-Bottle (GIAB) samples HG001-HG005. <https://cdn.sanity.io/files/l7780ks7/production-2024/0a1b6a62a6da3e3fcafb81cad4c8ff2ffe85dd41.pdf>. Fecha de publicación: febrero de 2025. Fecha de consulta: 29 de enero de 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/. Fecha de consulta: 29 de enero de 2026.
7. Asri M, Chang PC, Mier JC, et al. [Pangenome-aware DeepVariant](#). Preprint. *bioRxiv.* 2025;2025.06.05.657102. Fecha de publicación: 6 de junio de 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](#). Preprint. *bioRxiv.* 2025;2025.09.21.677443. Fecha de publicación: 21 de septiembre de 2025. doi:10.1101/2025.09.21.677443
10. Datos recopilados. 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. Fecha de publicación: 12 de diciembre de 2022. doi:10.1186/s13059-022-02818-4



1 800 809 4566 (llamada gratuita, EE. UU.) | tel.: +1 858 202 4566
 techsupport@illumina.com | www.illumina.com

© 2026 Illumina, Inc. Todos los derechos reservados. Todas las marcas comerciales pertenecen a Illumina, Inc. o a sus respectivos propietarios. Si desea consultar información específica sobre las marcas comerciales, consulte www.illumina.com/company/legal.html.
 M-GL-03931 ESP v1.0