

DRAGEN[™] Bio-IT Platform integrada en NextSeq[™] 1000 y NextSeq 2000 Systems

Solución integrada precisa
y eficaz para un análisis
de datos de NGS

- Proporciona un flujo de trabajo optimizado para un análisis secundario de gran precisión con una interfaz gráfica fácil de usar
- Genera llamadas de variantes de alta calidad directamente en integradas en los sistemas NextSeq 1000 y NextSeq 2000
- Reduce la necesidad de una infraestructura bioinformática adicional y requiere menos puntos de contacto que otros servicios de software de DRAGEN

illumina[®]

Introducción

El desarrollo del potencial del genoma es un objetivo compartido por muchos laboratorios de todo el mundo. Los recientes avances tecnológicos están mejorando el acceso a las plataformas y métodos de secuenciación de nueva generación (NGS, Next-Generation Sequencing) que pueden ayudar a desvelar los secretos del genoma. NextSeq 1000 y NextSeq 2000 Systems admiten aplicaciones de pequeña e intermedia productividad, como la secuenciación de exomas, transcriptomas y genomas pequeños; el enriquecimiento de objetivos; la secuenciación unicelular y la metagenómica de escopeta.

Una de las principales ventajas de NextSeq 1000 y NextSeq 2000 Systems es la inclusión de una implementación integrada de DRAGEN Bio-IT Platform (Dynamic Read Analysis for GENomics). El software DRAGEN integrado proporciona a los laboratorios un análisis secundario de NGS excepcionalmente rápido, preciso y completo, lo que reduce la necesidad de una infraestructura informática y recursos bioinformáticos adicionales. Gracias al software DRAGEN integrado, el análisis secundario se puede configurar durante la configuración del experimento de secuenciación, lo que permite un flujo de trabajo optimizado con menos puntos de contacto para el usuario y un tiempo de respuesta más rápido. Para los laboratorios que buscan un alto grado de configurabilidad, Illumina ofrece opciones adicionales de implementación del software DRAGEN (Figura 1).

El software DRAGEN integrado es el complemento perfecto para NextSeq 1000 y NextSeq 2000 Systems versátiles y flexibles. Al incluir varios de los mejores algoritmos de proceso, el software DRAGEN integrado permite a los investigadores superar los cuellos de botella en el análisis de datos y generar rápidamente resultados precisos para una amplia gama de aplicaciones (Tabla 1).

Eficiencia

El software DRAGEN integrado en NextSeq 1000 y NextSeq 2000 Systems permite a los laboratorios integrar a la perfección las funciones de secuenciación y análisis en la configuración de sus ensayos. El software DRAGEN integrado permite a los laboratorios obtener archivos FASTQ y archivos de formato de llamada de variantes (VCF, Variant Call Format) para las aplicaciones sucesivas directamente desde el instrumento, a medida que se completa el experimento, ahorrando tiempo de transferencia de archivos y de análisis. Los algoritmos del proceso de DRAGEN integrado ayudan a los usuarios principiantes y expertos a completar las funciones de análisis comunes y reducir la dependencia de expertos informáticos externos. Los archivos de datos de gran tamaño se gestionan fácilmente mediante el almacenamiento de archivos DRAGEN Original Read Archive (ORA), que consigue una compresión promedio de 4x sin pérdidas de los archivos FASTQ y usa un 80 % menos de energía para un lugar de trabajo más sostenible.



Figura 1: Opciones de implementación del proceso de DRAGEN Bio-IT para adaptarse a las necesidades de análisis de NGS de cada laboratorio.

Tabla 1: Aplicaciones de software DRAGEN integradas en NextSeq 1000 y NextSeq 2000 Systems

Aplicación ^a	Descripción
Conversión de BCL	Convierte los archivos BCL producidos por los sistemas de secuenciación de Illumina en archivos FASTQ
Compresión DRAGEN ORA	Produce una compresión basada en referencias sin pérdidas de los archivos FASTQ
DRAGEN FASTQ + MultiQC	Realiza criterios de medición de FastQC acelerados por hardware sin un tiempo de experimento adicional
Genoma completo	Realiza la asignación, la alineación y la llamada de variantes pequeñas del genoma humano. Solo germinal
Enriquecimiento (incluido el exoma)	Realiza llamadas de variantes pequeñas en muestras germinales o llamadas de variantes de baja frecuencia en muestras somáticas
Amplicón de ADN	Analiza la variación genética en regiones genómicas específicas. Usa el proceso de ADN de DRAGEN con un paso adicional para fijar levemente los cebadores y reescribir las alineaciones, garantizando que las secuencias de los cebadores no contribuyan a las llamadas de variantes
ARN	Ofrece un alineador de ARN-Seq (que tiene en cuenta el corte y el empalme) con filtrado opcional de ARNr durante la alineación, lo que reduce la duración del experimento y el tamaño del archivo
ARN unicelular	Procesa una amplia gama de conjuntos de datos de ARN-Seq unicelular, desde lecturas hasta matrices de expresión de célula por gen
Expresión diferencial	Ejecuta el algoritmo DESeq2 en los datos de cuantificación de ARN producidos por el proceso de ARN de DRAGEN. Produce genes y transcritos que se expresan de forma diferencial entre dos grupos de muestras
NanoString GeoMx NGS	Optimiza el análisis para los clientes que usan instrumentos GeoMx NGS e Illumina para los flujos de trabajo de genómica espacial
Metilación	Maneja datos de metilación de bisulfito y de secuenciación de piridina asistida por tet (TAPS, Tet-Assisted Pyridine Sequencing)

a. Hay disponibles otros procesos de aplicación en BaseSpace Sequence Hub, Illumina Connected Analytics, DRAGEN Multi-Cloud y el servidor DRAGEN en las instalaciones

Precisión

Los procesos de análisis de DRAGEN generan resultados excepcionalmente precisos. En Precision FDA Truth Challenge V2 (PrecisionFDA V2) organizado por la FDA en 2020, DRAGEN v3.7 fue galardonado con el premio a los resultados más precisos en todas las regiones de referencia y las regiones difíciles de trazar para los datos de secuenciación de Illumina. Otras innovaciones en Graph Genomes y en Illumina Machine Learning con DRAGEN v3.10 establecieron un nuevo estándar de precisión de datos en todas las tecnologías de secuenciación en el conjunto de datos de todas las regiones de referencia, logrando una puntuación F1 del 99,83 %, (un cálculo de resultados verdaderos positivos y verdaderos negativos como proporción de los resultados totales) (Figura 2). DRAGEN 3.10, Graph, e Illumina Machine Learning también recibieron la mejor puntuación F1 por la mayor precisión de las llamadas en comparación con todas las presentaciones de PrecisionFDA V2 en la región del complejo principal de histocompatibilidad (MHC, Major Histocompatibility Complex).^{1,2}

DRAGEN Bio-IT Platform excepcionalmente eficiente consigue tiempos de análisis rápidos gracias al uso de la tecnología de array de puertas lógicas programables en campo (FPGA, Field-Programmable Gate Array). La tarjeta de FPGA integrada, exclusiva de NextSeq 1000 y NextSeq 2000 Systems, permite realizar procesos informáticos de análisis secundarios acelerados por hardware para una gama de aplicaciones. Estos procesos de DRAGEN se mejoran continuamente y se crean procesos adicionales para ofrecer la mejor funcionalidad, precisión y velocidad posibles. La plataforma está diseñada en torno a las necesidades de análisis de los investigadores, lo que les permite dedicar menos tiempo y esfuerzo a la ejecución de procesos a nivel de producción para que se puedan centrar más en los resultados.

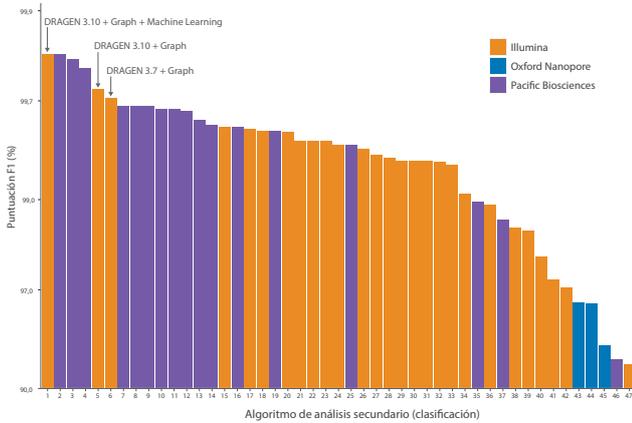


Figura 2: La precisión de la llamada de variantes pequeñas lograda con DRAGEN 3.10, Graph y Machine Learning en comparación con los datos presentados en PrecisionFDA Truth Challenge v2 en el conjunto de datos para todas las regiones de referencia: el método de análisis de DRAGEN 3.10 + Graph + Machine Learning ocupa el primer lugar, al mismo nivel que la lectura presentada por Pacific Biosciences de mayor rendimiento. DRAGEN 3.10 + Graph muestra una mejora respecto a DRAGEN 3.7 + Graph gracias a las mejoras en el gráfico y en la manipulación referencias/cóntigo ALT.

Amplia selección de aplicaciones

El software DRAGEN integrado en NextSeq 1000 y NextSeq 2000 Systems cuenta con procesos de análisis versátiles y puede crear diversos archivos de rendimiento en diferentes etapas del flujo de trabajo

(Figura 3). La selección incluida de procesos de análisis es compatible con una amplia variedad de tipos de análisis, incluidos, entre otros, el análisis unicelular, del exoma y del ARN (Tabla 1).

Procesos de ARN unicelular de DRAGEN integrados

Los procesos de ARN unicelular de DRAGEN son soluciones rápidas y flexibles que pueden procesar varios conjuntos de datos de ARN-Seq de células únicas con formatos de resultados fáciles de usar, incluyendo gráficos de agrupación de tipos de células basados en actividades de expresión de ARN (Figura 4). Además, cuenta con útiles extensiones para el procesamiento de conjuntos de datos multiplexados que consisten en varias muestras (por ejemplo, usando el demultiplexado de genotipos o el algoritmo de resumen de células) y el recuento de la expresión de proteínas de la superficie celular. Asimismo, admite una serie de tipos de preparación de librerías de entrada para la compatibilidad con las herramientas de análisis sucesivos.

El CC integrado de las librerías de expresión unicelular y el proceso de análisis de células únicas se ejecutan en menos tiempo en comparación con las soluciones basadas en la nube y sin perder precisión. El proceso unicelular va desde la configuración del experimento hasta la expresión cuantificada por célula en NextSeq 1000 y NextSeq 2000 Systems partiendo de un único punto de contacto, reduciendo la necesidad de recursos informáticos adicionales.

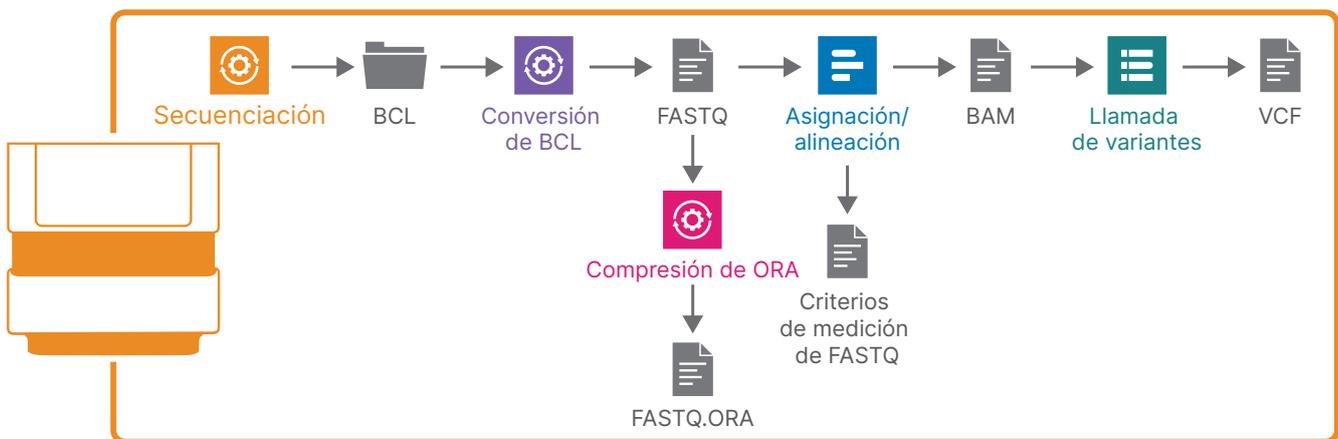


Figura 3: El software DRAGEN integrado en NextSeq 1000 y NextSeq 2000 Systems proporciona un flujo de trabajo de procesamiento de datos y un análisis secundario optimizados: la solución de software DRAGEN integrado proporciona una herramienta altamente eficiente para producir los resultados en los formatos que los laboratorios necesitan para las aplicaciones sucesivas.

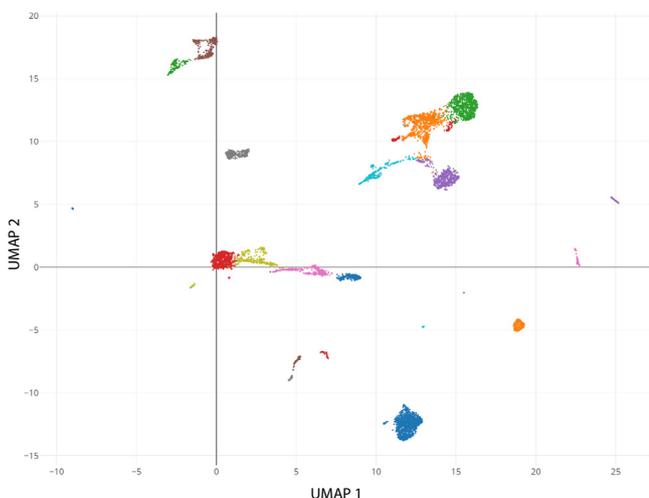


Figura 4: Ejemplo de gráfico de agrupación de tipos de células a partir del proceso de ARN de células únicas de DRAGEN: el análisis de aproximación y proyección uniforme de la matriz (UMAP, Uniform Manifold Approximation and Projection) permite la visualización de células individuales por tipo o estados funcionales basándose en perfiles de expresión complejos y multidimensionales.⁴

Proceso de DRAGEN Enrichment integrado

El proceso de DRAGEN Enrichment integrado proporciona un análisis rápido, incluyendo la calibración avanzada del modelo de error para una mayor precisión de los paneles de exoma basados en el enriquecimiento y la hibridación, y criterios de medición basados en el objetivo de entrada. El software se puede ejecutar en modo germinal o somático, realizando llamadas de variantes pequeñas germinales o llamadas de variantes somáticas (de baja frecuencia). La llamada de la variación en el número de copias (CNV, Copy Number Variation) y la llamada de la variante estructural se pueden activar en el modo germinal. El proceso de DRAGEN Enrichment es más rápido y más preciso que el análisis con el alineador Burrows-Wheeler (BWA, Burrows-Wheeler Aligner) y el llamador de variantes de Genome Analysis Toolkit (GATK), particularmente en la llamada de indels (Tabla 2, Figura 5). Una celda de flujo en NextSeq 1000 y NextSeq 2000 Systems se puede procesar normalmente en menos de dos horas.

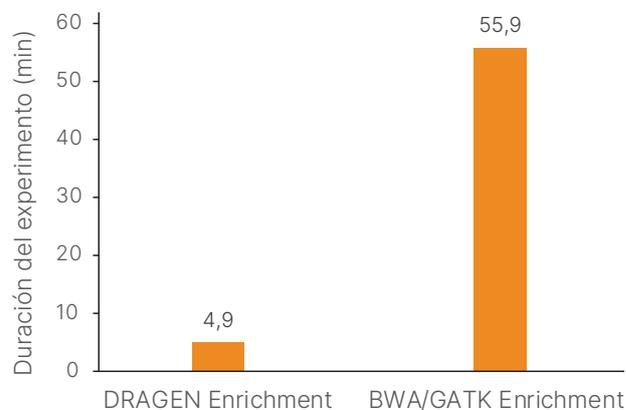


Figura 5: Tiempo promedio de experimento por muestra para el proceso de DRAGEN Enrichment y la aplicación BWA Enrichment en BaseSpace™ Sequence Hub: el análisis de enriquecimiento germinal de 24 duplicados de NGS muestra tiempos de análisis por muestra más rápidos para el proceso de DRAGEN Enrichment que el análisis similar que usa el enriquecimiento de BWA, seguido de llamadas de variantes con el llamador de variantes de GATK.

Proceso de ARN de DRAGEN integrado

El proceso de ARN de DRAGEN integrado realiza el análisis secundario de los transcritos de ARN. El proceso de ARN ofrece múltiples modos de funcionamiento, incluyendo la alineación solo de referencia y la alineación asistida por anotación con detección de fusión génica. Se puede activar un filtro opcional de ARN ribosómico (ARNr) para reducir la duración del experimento de análisis y el tamaño del archivo. El módulo de fusión génica aprovecha el alineador de ARN cortado y empalmado de DRAGEN para realizar un análisis de lectura dividida en alineaciones complementarias (quiméricas) con el fin de detectar posibles puntos de ruptura, al tiempo que añade un tiempo de procesamiento mínimo al análisis general. La detección de la fusión génica, la cuantificación de los transcritos y la llamada de variantes se pueden activar usando las funciones del software DRAGEN integrado.

Tabla 2: Resultados de precisión del proceso de DRAGEN Enrichment y de BWA/GATK Enrichment

	SNV			Indel		
	Precisión (%)	Exhaustividad (%)	F1	Precisión (%)	Exhaustividad (%)	F1
DRAGEN Enrichment	99,7	94,64	97,11	96,54	83,83	89,74
BWA/GATK Enrichment	99,82	92,77	96,17	96,71	71,32	82,1

Resumen

El software DRAGEN Bio-IT integrado en NextSeq 1000 y NextSeq 2000 Systems ofrece una bioinformática fácil de configurar y una experiencia de usuario simplificada en una implementación local. Ofrece un análisis secundario de los datos de NGS generados en el instrumento que es más rápido que los procesos de DRAGEN en BaseSpace™ Sequence Hub. Los procesos integrados, como los unicelulares, de enriquecimiento y de ARN, ofrecen muchas ventajas de velocidad y precisión con respecto a las ofertas actuales.

Datos para realizar pedidos

Producto	N.º de catálogo
NextSeq 1000 Sequencing System	20038898
NextSeq 2000 Sequencing System	20038897
Actualización de NextSeq 1000 a NextSeq 2000	20047256

Información adicional

NextSeq 1000 y NextSeq 2000 Systems, illumina.com/systems/sequencing-platforms/nextseq-1000-2000.html

DRAGEN Bio-IT Platform, illumina.com/products/by-type/informatics-products/dragen-bio-it-platform.html

Servicio de asistencia de DRAGEN, support.illumina.com/sequencing/sequencing_software/dragen-bio-it-platform.html

illumina®

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

© 2022 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, visite www.illumina.com/company/legal.html.
M-GL-00729 v1.0 ESP

Bibliografía

1. Food and Drug Administration. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Acceso: 14 de marzo de 2022.
2. Illumina. DRAGEN Sets New Standard for Data Accuracy in PrecisionFDA Benchmark Data. Optimizing Variant Calling Performance with Illumina Machine Learning and DRAGEN Graph. illumina.com/science/genomics-research/articles/dragen-shines-again-precisionfda-truth-challenge-v2.html. Acceso: 14 de marzo de 2022.
3. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498. doi:10.1038/ng.806
4. McInnes, L, Healy J y Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. [arXiv preprint arXiv:1802.03426](https://arxiv.org/abs/1802.03426). Revisión: 18 de septiembre de 2020. Acceso: 5 de mayo de 2022. doi.org/10.48550/arXiv.1802.03426