

Software Emedgene™

Mejore la interpretación de variantes definida por el usuario para enfermedades raras y otras aplicaciones de investigación genética

- Optimiza el análisis terciario para aplicaciones germinales hasta en un 75 % mediante XAI y una automatización personalizable.
- Se integra con LIMS, EMR y otros sistemas informáticos para simplificar y proteger todo el flujo de trabajo de los ensayos.
- Impulsa el crecimiento de los usuarios para satisfacer las demandas en constante evolución de la ciencia, la tecnología y la escala para NGS y ensayos con microarray.



Introducción

Las aplicaciones de la secuenciación de nueva generación (NGS, next-generation sequencing) para investigar enfermedades genéticas raras y de otro tipo, farmacogenómica, citogenética, otras afecciones germinales o estudios de cribado de portadores son fundamentales para avanzar en nuestra comprensión de la genética humana y traducirla en medicina de precisión. La secuenciación del genoma completo (WGS, whole-genome sequencing), la secuenciación del exoma completo (WES, whole-exome sequencing), los paneles selectivos y virtuales proporcionan una visión imparcial de alta resolución del genoma completo, las regiones codificantes o las regiones génicas selectivas, respectivamente. Sin embargo, el análisis y la interpretación de las enormes cantidades de datos producidos por estos métodos representan un importante impedimento para los laboratorios. En particular, la interpretación de variantes es un proceso manual que requiere mucho tiempo y que requiere numerosos pasos repetitivos que pueden durar hasta siete horas por genoma.¹ Para aplicaciones de paneles de alta productividad, como proyectos de cribado de poblaciones sanas o portadoras, los laboratorios suelen intentar mejorarla eficiencia y la precisión mediante herramientas de automatización, elección de variantes y generación eficiente de informes de investigación.

Para que los laboratorios puedan afrontar estas dificultades, Illumina ofrece el software Emedgene, un completo software de investigación diseñado por genetistas para optimizar los flujos de trabajo de interpretación y reducir el trabajo de selección de variantes. El software Emedgene cuenta con inteligencia artificial explicable (XAI, explainable artificial intelligence) y automatización para flujos de trabajo de interpretación de variantes definidos por el usuario de alta productividad, lo que permite un aumento de 2 a 5 veces en la eficiencia y una reducción del 50 al 75 % en el tiempo total del flujo de trabajo por sujeto en comparación con la interpretación manual.

Al integrarse con las plataformas de análisis secundario y almacenamiento de datos de Illumina, incluidos el análisis secundario de DRAGEN™, BaseSpace™ Sequence Hub e Illumina Connected Analytics, el software Emedgene ofrece opciones modulares en una solución de proveedor único, desde el procesamiento de muestras hasta la generación de informes de investigación. El software admite la entrada de datos de WGS, WES, paneles selectivos y microarrays. Se admite una gama completa de tipos de variantes, incluidas variantes de nucleótido único (SNV, single nucleotide variants), inserciones/delecciones (indel), variantes de número de copias (CNV, copy number variants), variantes de ADN mitocondrial (ADNmt), variantes estructurales (SV, structural variants) y repeticiones cortas en tándem (STR, short tandem repeats). El software Emedgene integra las exhaustivas

capacidades de llamada de variantes del análisis secundario de DRAGEN para garantizar que incluso los genes y regiones de interés más difíciles, como las variantes de mosaico, las regiones de duplicación segmentaria y más, se identifiquen con precisión como entrada para la interpretación posterior por parte del usuario final del laboratorio.

El software Emedgene también ofrece la opción de definir “paneles virtuales” mediante procesos bioinformáticos a partir de una cadena principal más completa, como la WES o la WGS, lo que permite a los laboratorios de investigación desarrollar un ensayo de laboratorio y seleccionar el panel de genes más relevante para su aplicación, lo que hace que las actualizaciones del panel de ensayos sean más eficientes a medida que se ponen de manifiesto nuevas asociaciones entre genes y enfermedades.

Análisis terciario completo mediante XAI en todos los tipos de variantes

Los algoritmos exclusivos de aprendizaje automático proporcionan información basada en pruebas y ayudan a los usuarios a ampliar sus flujos de trabajo de interpretación. La XAI del software Emedgene destaca las variantes priorizadas para su revisión, basándose en pruebas seleccionadas que se relacionan automáticamente dentro del software. Este enfoque permite evitar tareas manuales y repetitivas que suelen ser necesarias para conectar las pruebas con los conocimientos. La XAI cubre una amplia gama de tipos de variantes, que incluye SNV, indels, CNV, inserciones de SV, STR y ADNmt.

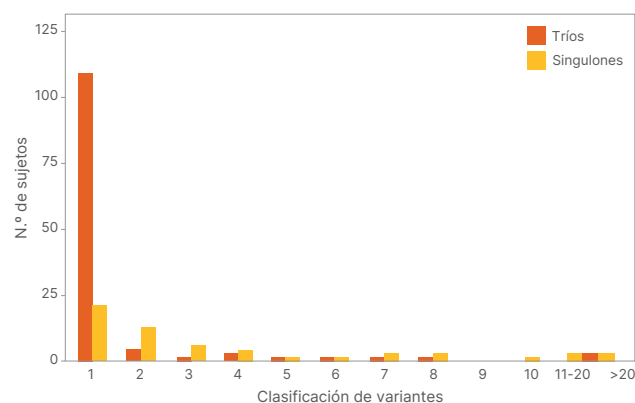


Figura 1: clasificación de variantes basada en XAI. En un estudio de validación realizado por investigadores de Baylor Genetics, la interpretación de variantes basada en XAI clasificó correctamente las variantes causales entre las 10 primeras en el 96,6 % de los sujetos de investigación evaluados, incluidos los tríos (color naranja) y los singlones (color amarillo).²

En un estudio de validación realizado por investigadores de Baylor Genetics, Houston, TX (EE. UU.), se evaluó la interpretación de variantes con XAI con una cohorte de 180 conjuntos de datos de WES procedentes de sujetos de investigación, que previamente se habían resuelto mediante revisión manual. En el 96,6 % de las muestras evaluadas, se identificó la variante causal y se clasificó entre las 10 primeras para su posterior revisión (figura 1).²

Selección de variantes automatizada y explicable

Como parte de una experiencia para el usuario optimizada, el software Emedgene conecta pruebas para cada variante preseleccionada por la XAI. El gráfico de pruebas imita el trabajo realizado por un científico, mostrando asociaciones enfermedad-gen junto con patrones hereditarios y fenotipos asociados al sujeto. Cada prueba está vinculada a fuentes externas, ya sea en la bibliografía científica o en una base de datos (figura 2). El gráfico de pruebas permite revisar y evaluar rápidamente los resultados de la plataforma de XAI.

Las actualizaciones mensuales de fuentes de datos públicas y de propiedad exclusiva garantizan que el software Emedgene se mantenga actualizado con los descubrimientos científicos. El software agrega e integra datos procedentes de las bases de datos externas preferidas, como el catálogo Online Mendelian Inheritance in Man (OMIM), ClinVar y otros en un

gráfico de conocimiento que impulsa los modelos de XAI e incluye la selección automática de la bibliografía científica mediante procesamiento del lenguaje natural (NLP, Natural Language Processing).

Clasificaciones automáticas de ACMG

Con el uso de la automatización de las clasificaciones del American College of Medical Genetics (ACMG) para variantes de SNV, indel, CNV, SV y ADNmt, Emedgene optimiza la revisión de las selecciones de etiquetas de clasificación automatizadas por parte de un selector de variantes (figura 3). Esto reduce el tiempo de interpretación de las etiquetas cuantitativas y puede aumentar la concordancia entre selectores.

Optimización de flujos de trabajo

El software Emedgene incluye varias funciones que optimizan y agilizan los flujos de trabajo de interpretación definidos por el usuario (figura 4). El software Emedgene es independiente del ensayo y compatible con una amplia gama de entradas de datos, incluidos los formatos de archivo FASTQ y VCF. El software se integra con la solución de almacenamiento del usuario para lograr una experiencia “sin transferencia de datos”.

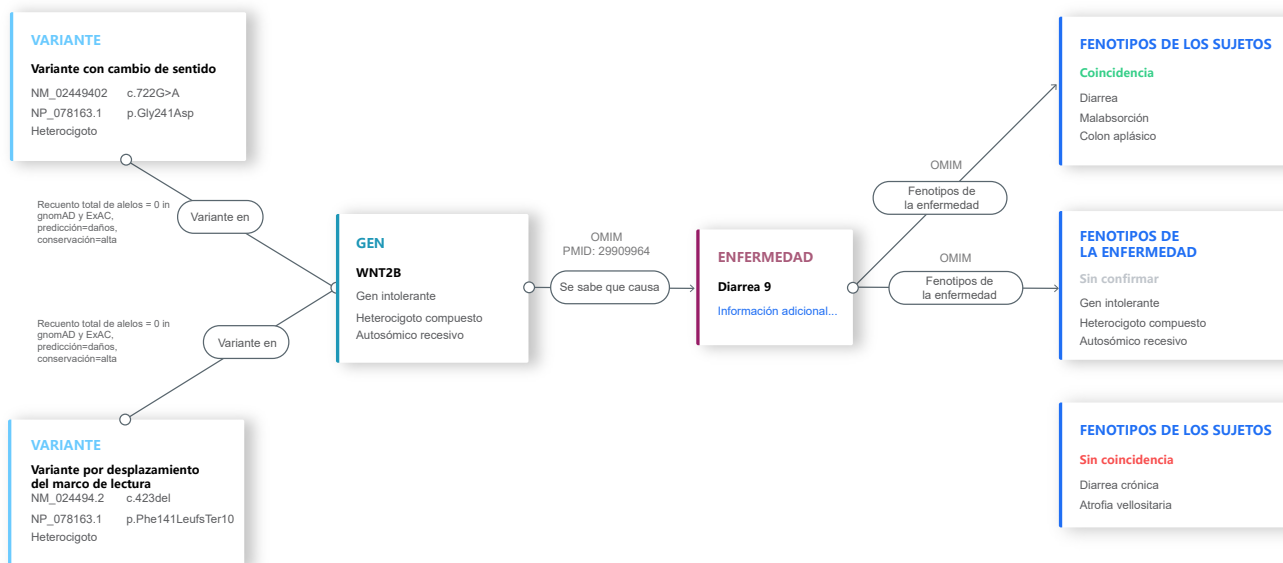


Figura 2: gráfico de pruebas del software Emedgene. Las variantes altamente clasificadas se muestran con un gráfico de pruebas seleccionado automáticamente que recopila pruebas para una revisión rápida y sencilla por parte de un científico de investigación de interpretación de variantes, incluidos enlaces a la bibliografía y las bases de datos relevantes.

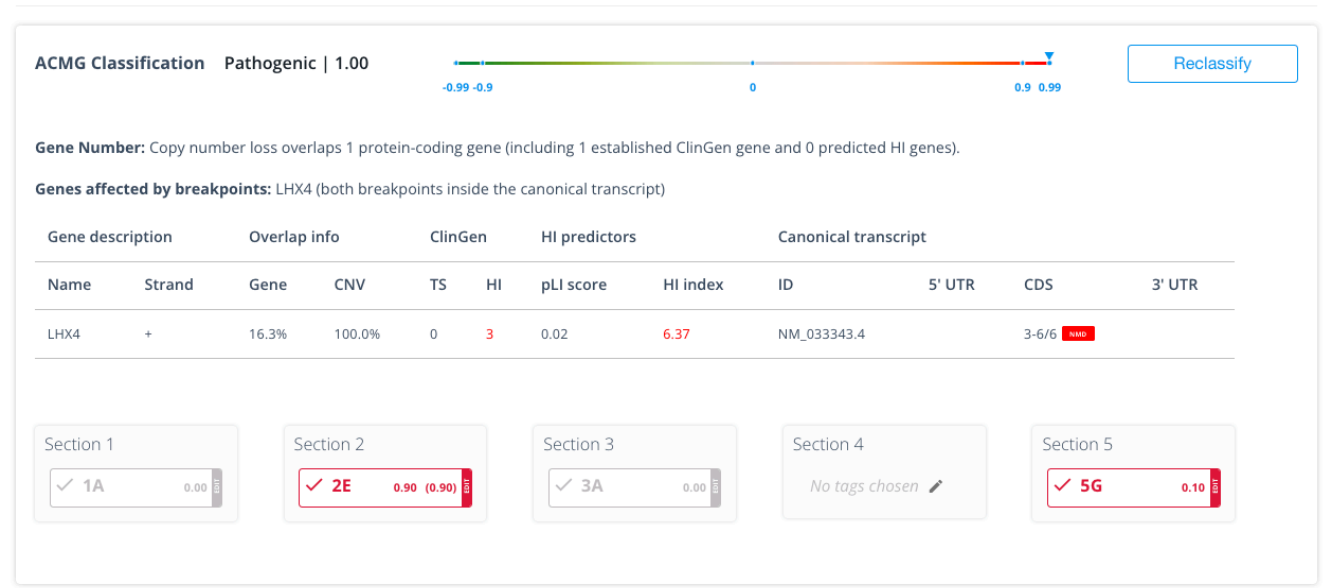


Figura 3: clasificaciones automatizadas de ACMG. El software genera automáticamente clasificaciones de ACMG para variantes, incluidas variantes de SNV, indels, CNV, SV y ADNmt, lo que permite una revisión optimizada por parte de un científico de investigación de interpretación de variantes.

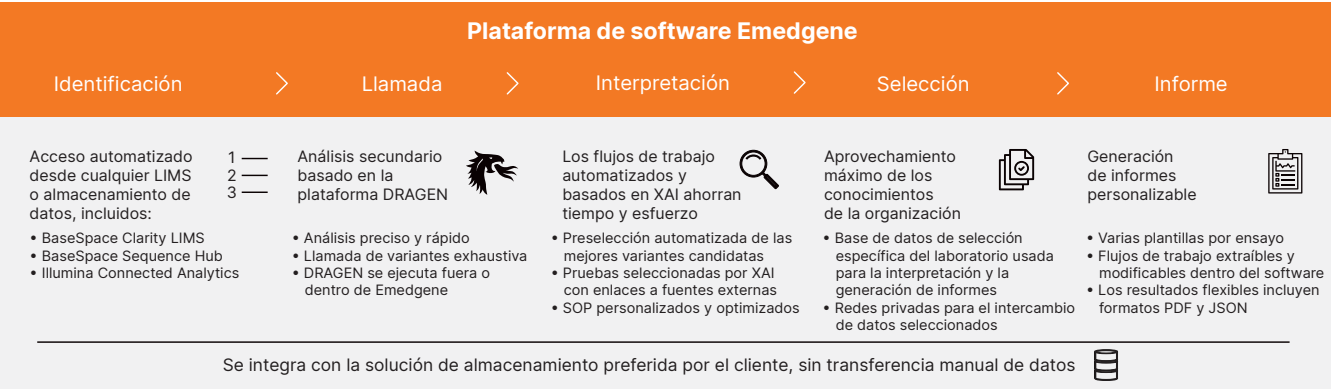


Figura 4: el software Emedgene cuenta con numerosas funciones para optimizar y agilizar los flujos de trabajo de interpretación definidos por el usuario.

Procedimientos operativos estándar (SOP) personalizables

El software Emedgene ofrece una plataforma flexible y personalizable que permite a los laboratorios de investigación potenciar la eficiencia y aumentar sus flujos de trabajo de interpretación específicos. Se pueden implantar procedimientos operativos estándar (SOP, Standard Operating Procedures) configurables mediante la creación de conjuntos de filtros personalizados predefinidos, flujos de trabajo, generación de informes y la inclusión de entornos de validación y producción.

Selección específica del laboratorio

El software Emedgene cuenta con un repositorio para los datos seleccionados de un laboratorio, que puede incluir selecciones de variantes y genes que se utilizan automáticamente para anotar cómodamente casos y rellenar automáticamente plantillas de interpretación. El mantenimiento y el crecimiento de los datos seleccionados da como resultado aumentos importantes de eficiencia. La base de datos seleccionada se puede buscar por variante y gen, para facilitar el acceso a todos los casos evaluados previamente.

Intercambio de datos seleccionados

Gracias al poder de la colaboración, los laboratorios pueden ampliar conocimientos intercambiando datos seleccionados en una red privada de laboratorios. Los colaboradores de la red y los permisos para intercambiar datos son de carácter exclusivo y los determinan los laboratorios participantes. Esta función de datos compartidos de suscripción permite a los socios compartir conocimientos más profundos con socios de confianza.

Flujo de trabajo basado en la visualización para la interpretación de CNV

Las visualizaciones mejoradas del genoma y del cromosoma completo aceleran la interpretación de las CNV con un cómodo acceso con un solo clic a los datos a nivel de variantes. Los completos seguimientos de visualización de poblaciones y bases de conocimientos se combinan con una fácil visualización de datos organizativos, históricos y seleccionados para proporcionar una solución optimizada para la interpretación citogenética a partir de datos de NGS o microarray.

Integración del análisis secundario de DRAGEN

El software de análisis secundario de DRAGEN es un conjunto de aplicaciones que procesa datos de NGS y realiza llamadas de variantes precisas, completas y eficientes. Como parte de PrecisionFDA Truth

Challenge V2, el análisis secundario de DRAGEN fue declarado como la solución más precisa (puntuación F1) y más completa entre las variantes evaluadas en todas las regiones de referencia en comparación con otras herramientas participantes.³ El rendimiento demostrado de esta solución aborda, además, los retos comunes en el análisis genómico, como la precisión óptima de las llamadas de variantes, la compatibilidad con una amplia gama de ensayos y tipos de variantes, y el procesamiento eficiente de volúmenes masivos de datos genómicos. La referencia multigenómica (gráfico) refleja la diversidad genética, reduce el sesgo étnico y mejora la precisión de SNV en comparación con otras aplicaciones. El análisis secundario de DRAGEN también llama de forma exhaustiva a las variantes en regiones de alta homología.

El software Emedgene proporciona la integración más completa posible con el análisis secundario de DRAGEN para generar una solución de análisis secundario a través de la solución de análisis terciaria. Se puede acceder a los procesos de DRAGEN a través de una instancia local, BaseSpace Sequence Hub, Illumina Connected Analytics (ICA), o se puede configurar y acceder a los procesos directamente dentro del software Emedgene para una solución lista para usar fácilmente integrada. Para lograr la máxima flexibilidad, el software Emedgene también es compatible con resultados en formato de llamada de variantes genómicas (VCF, genomic variant call format) de prácticamente cualquier llamador de variantes, lo que permite a los laboratorios beneficiarse de una solución de un único proveedor para su flujo de trabajo de NGS.

Generación de informes personalizable

Los usuarios pueden personalizar, editar y rellenar automáticamente los informes siguiendo un flujo de trabajo definido por el usuario utilizando una cómoda herramienta de edición similar a Microsoft Word. Gracias a la compatibilidad con flujos de trabajo modificados y revisados, los informes se pueden enviar para su revisión adicional dentro del software. Para aumentar la flexibilidad, los informes se pueden descargar en formato PDF o JSON, facilitando así el intercambio de datos.

Integración con otras plataformas y sistemas

Mantenerse al día de las nuevas tecnologías de genómica, los nuevos instrumentos y las nuevas metodologías e integrarlos puede ser complicado. El software Emedgene simplifica el proceso, integrándose con NovaSeq™ X Series, NovaSeq 6000, NextSeq™ 2000 u otros sistemas de secuenciación de Illumina para optimizar los análisis de NGS. El software Emedgene también representa una pieza fundamental en el flujo de trabajo de investigación de enfermedades genéticas al integrar una solución “desde el ADN hasta el informe” con las tecnologías de preparación y secuenciación de librerías de Illumina.

El software Emedgene ofrece sólidas interfaces de programación de aplicaciones (API, Application Programming Interface) que permiten la integración con otros sistemas de gestión de información de laboratorio (LIMS, Laboratory Information Management System) institucionales. Las integraciones de almacenamiento sin transferencia de datos están disponibles para BaseSpace Sequence Hub, Illumina Connected Analytics, Amazon S3 for Amazon Web Services, Azure Blob y más, para una integración flexible en la infraestructura informática existente.

Red privada y segura

El software Emedgene cumple los requisitos de la Ley de Portabilidad y Responsabilidad del Seguro Médico (HIPAA, Health Insurance Portability and Accountability Act) y sigue los principios del Reglamento General de Protección de Datos (RGPD) de la Unión Europea. Illumina implementa la protección de datos según el RGPD, que requiere que las organizaciones implementen medidas técnicas y organizativas para la seguridad de los datos e integren salvaguardas en las actividades de tratamiento de datos para proteger los derechos individuales, por diseño y por defecto. Además, Illumina ofrece varias funciones para cumplir los requisitos de la HIPAA, incluidas protecciones administrativas, físicas y técnicas.

El software Emedgene también cumple con los requisitos de privacidad de datos globales y locales, incluida la ISO 27001, y cumple con SOC 1 y 2 tipo II (control de la organización del servicio, en inglés: Service Organization Control). Las opciones de seguridad avanzadas personalizadas incluyen la integración con la política de inicio de sesión único de la organización. También hay disponible una opción para que utilice su propia clave y para el doble cifrado de los datos.

Resumen

El software Emedgene ofrece una solución basada en XAI y habilitada para la automatización para análisis terciario de enfermedades raras y datos de investigación genética. Aumenta la eficacia de la interpretación de variantes y permite a los laboratorios ampliar sus operaciones. Gracias a la integración de las soluciones de secuenciación, microarray y software de Illumina, incluida la plataforma DRAGEN para el análisis secundario preciso, exhaustivo y eficiente de llamadas de variantes, el software Emedgene proporciona a los laboratorios de investigación una potente herramienta para un crecimiento continuo.

Más información

[Software Emedgene](#)

Datos para realizar pedidos

Producto	N.º de catálogo
Emedgene software—Genome Equivalent Sample (FASTQ)	20073772
Emedgene software—Genome Equivalent Sample (VCF)	20073773
Emedgene software—Annual Support and Maintenance (Basic)	20073776
Emedgene software—Annual Support and Maintenance (Basic)	20073777
Emedgene software—Annual Support and Maintenance (Enterprise)	20073778
Emedgene software—Training at Customer Site	20073774
Custom report (Annual)	20073778
Cualquier elemento adicional necesario para el laboratorio se incluirá en el presupuesto preparado.	

Para consultas de cualificación, Illumina ofrece una experiencia de evaluación asistida, que permite a los clientes trabajar con conjuntos de datos de investigación de ejemplo disponibles en el software Emedgene o cargar y evaluar sus propios conjuntos de datos dentro del software. Póngase en contacto con el representante de ventas de Illumina para obtener más información.

Bibliografía

1. Austin-Tse CA, Jobanputra V, Perry DL, et al. [Best practices for the interpretation and reporting of clinical whole genome sequencing](#). *NPJ Genom Med*. 2022;7:27. doi.org/10.1038/s41525-022-00295-z.
2. Meng L, Attali R, Talmy T, et al. [Evaluation of an automated genome interpretation model for rare disease routinely used in a clinical genetic laboratory](#). *Genet Med*. 2023;25(6):100830. doi:10.1016/j.gim.2023.100830
3. Olson ND, Wagner J, McDaniel J, et al. [PrecisionFDA Truth Challenge V2: Calling variants from short and long reads in difficult-to-map regions](#). *Cell Genomics*. 2022; 2(5): 100129. doi.org/10.1016/j.xgen.2022.100129.



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

© 2024 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-01057 ESP v2.0