

Infinium™ Global Diversity Array with Polygenic Risk Score Content-8 v1.0

Una solución versátil y
flexible para la investigación
multiétnica de la PRS

- El contenido actualizado incluye 160 000 marcadores de la PRS para una amplia cobertura de trastornos
- Una amplia cobertura genómica permite determinar con gran precisión la PRS panétnica
- La compatibilidad con el módulo de software Predict simplifica el cálculo de la PRS y la generación de informes

illumina®

Introducción

La evaluación precisa del riesgo de enfermedad es un aspecto clave de la medicina preventiva y personalizada. Sin embargo, en el caso de las enfermedades con herencia poligénica, en las que influyen cientos o miles de variantes genéticas que actúan en combinación con factores ambientales, la predicción del riesgo sigue siendo un reto. Las puntuaciones de riesgo poligénico (PRS, Polygenic Risk Scores), también conocidas como puntuaciones poligénicas o puntuaciones de riesgo genético, son indicadores numéricos que representan la tendencia genética de una persona a desarrollar un rasgo o una enfermedad.¹ Las PRS son una herramienta útil para que los investigadores estudien la estratificación del riesgo y comprendan mejor quién tendría más probabilidades de beneficiarse de una supervisión adicional o de intervenciones preventivas tempranas.¹⁻³

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 (Tabla 1) es una solución de microarrays de alto rendimiento que proporciona un flujo de trabajo versátil y flexible para respaldar los estudios de la PRS. El BeadChip de ocho muestras, construido sobre la base de Infinium Global Diversity Array de eficacia probada, presenta un contenido de genotipado actualizado que incluye 160 000 marcadores para la determinación precisa de la PRS (Tabla 2). Los datos del BeadChip se analizan de forma óptima mediante el módulo de software Predict, que ofrece una solución completa de genotipo a riesgo para la investigación de enfermedades poligénicas.

Contenido optimizado para la investigación de la PRS

La base del genoma completo de Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 contiene aproximadamente 1,9 millones de marcadores científicamente relevantes. Este amplio contenido de Infinium Global Diversity Array-8 v1.0, la array de genotipado usada en el programa de investigación [All of Us Research Program](#), se complementa con 160 000 marcadores de la PRS del amplio Catálogo de Puntuación Poligénica (PGS, PolyGenic Score)⁴ para proporcionar una cobertura ampliada a través de regiones genómicas de alto valor (Tabla 2, Tabla 3). Estos marcadores, seleccionados en colaboración con [Allelica](#), proporcionan una cobertura excepcional para seis trastornos clave de la PRS, como la diabetes de tipo 1 y 2, el cáncer de mama, el cáncer de colon, el cáncer de próstata, la arteriopatía coronaria y la enfermedad de Alzheimer.

A fin de generar el contenido de refuerzo de la PRS, las puntuaciones de las muestras de array se compararon con las de referencia del estándar de oro de [1000 Genomes Project](#) y las puntuaciones comunes del Catálogo de PGS (por ejemplo, MAV313) para todas las PRS de Allelica. El objetivo de esta evaluación comparativa era identificar las variantes con los errores más frecuentes en la imputación comparando los resultados de la asignación con datos de secuenciación del genoma completo (WGS, Whole-Genome Sequencing) 30x.

A fin de desarrollar la array, se seleccionaron las variantes con el mayor tamaño del efecto en la PRS y la menor calidad de imputación. Estas variantes se incluyeron directamente en la array para ser genotiparse, en lugar de imputarse. El contenido de refuerzo de la PRS incluye variantes con el mayor tamaño del efecto específico para cada ascendencia en las puntuaciones de múltiple ascendencia de Allelica para optimizar el valor de la array para personas de todas las ascendencias y aumentar la concordancia para la PRS.

Tabla 1: Descripción general de Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0

Característica	Descripción
Especie	Humana
Número total de marcadores ^a	2 028 571
Número de muestras por BeadChip	8
Cantidad necesaria de aporte de ADN	200 ng
Duplicados de SNP	15
Número de SNP que se necesitan para hacer llamadas de CNV	10
Proceso químico del ensayo	Infinium LCG
Compatibilidad con instrumentos	iScan System
Productividad de muestras máxima de iScan System ^b	~1728 muestras/semana
Tiempo de lectura por muestra ^b	3-5 minutos

a. El contenido incluye aproximadamente 1,9 millones de marcadores del genoma completo de Infinium Global Diversity Array-8 v1.0 más 160 000 marcadores específicos de la PRS.

b. Los valores son aproximados. Los tiempos de lectura y la productividad máxima variarán en función del laboratorio y de las configuraciones del sistema.

Tabla 2: Contenido de alto valor de Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0

Contenido	N.º de marcadores ^a	Aplicación en investigación/Nota
Genes hg19 de RefSeq	1 123 407	Todos los genes conocidos
hg19 de RefSeq ± 10 kb	1 266 608	Regiones de regulación ^b
Promotores de RefSeq	48 249	2 kb secuencia arriba para abarcar las regiones promotoras
ADME exónica	16 528	Absorción, distribución, metabolismo y eliminación del fármaco
Genes hg19 de ADME	32 117	
hg19 de ADME ± 10 kb	37 468	Incluye regiones de regulación
Marcadores de HLA	17 595	Defensa frente a las enfermedades, rechazos de trasplantes, trastornos autoinmunitarios
Genes hg19 de HLA	1297	
Marcadores de MHC	22 783	
Genes hg19 de COSMIC	1 079 088	Mutaciones somáticas del cáncer
Genes hg19 de GO	324 039	Anotación de ontología genética

- a. El número de marcadores de cada categoría puede cambiar.
 b. De todos los genes conocidos.

Abreviaturas: ADME, absorción, distribución, metabolismo y excreción; COSMIC, catálogo de mutaciones somáticas en cáncer; GO, base de datos de ontología genética; hg19, versión 19 del genoma humano; HLA, antígeno leucocitario humano; MHC, complejo mayor de histocompatibilidad; RefSeq, base de datos de secuencias de referencia del National Center for Biotechnology Information (Centro Nacional para la Información Biotecnológica).

Tabla 3: Información sobre los marcadores

Categorías de marcadores	N.º de marcadores ^a
Marcadores exónicos ^a	530 644
Marcadores sin sentido ^b	28 287
Marcadores con cambio de sentido ^b	396 783
Marcadores sinónimos ^b	33 442
SNP ocultas ^b	40 964
Marcadores de ADN mitocondrial ^b	1346
Indels ^c	37 289
Marcadores del cromosoma X ^c	62 103
Marcadores del cromosoma Y ^c	6449
Marcadores de PAR/homólogos ^c	5485

- a. Base de datos RefSeq, NCBI Reference Sequence Database, ncbi.nlm.nih.gov/refseq.
 b. En comparación con el navegador Genome Browser, genome.ucsc.edu.
 c. Versión consensuada NCBI Genome Reference Consortium, Version GRCh37 (hg19), ncbi.nlm.nih.gov/assembly/GCF_000001405.13.

Abreviaturas: indel, inserción/delección; NCBI, National Center for Biotechnology Information; PAR, región pseudoautosómica (PseudoAutosomal Region); SNP, polimorfismos de un solo nucleótido (Single Nucleotide Polymorphisms); UCSC, Universidad de California Santa Cruz.

Amplia cobertura de variantes asociadas a enfermedades

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 proporciona cobertura de variantes seleccionadas del catálogo GWAS del National Human Genome Research Institute and European Bioinformatics Institute (NHGRI-EBI [Instituto Nacional de Investigación del Genoma Humano y el Instituto Europeo de Bioinformática])⁵ que representan una amplia variedad de fenotipos y clasificaciones de enfermedades (Figura 1). La selección de variantes incluye también una variedad de clasificaciones de enfermedades obtenidas de las anotaciones de ClinVar⁶ y American College of Medical Genetics (ACMG [Colegio Americano de Genética Médica y Genómica])⁷.

Las bases de datos clínicas, como ClinVar, evolucionan constantemente a medida que se añaden nuevas variantes y que las designaciones de estas cambian a "Patógena" o "Probablemente patógena". El BeadChip proporciona una cobertura actualizada para muchas de estas variantes de alto valor que se incluyen en bases de datos anotadas. Este amplio contenido es una gran oportunidad para los investigadores interesados en estudiar poblaciones diversas para probar y validar asociaciones previamente reconocidas en poblaciones europeas.

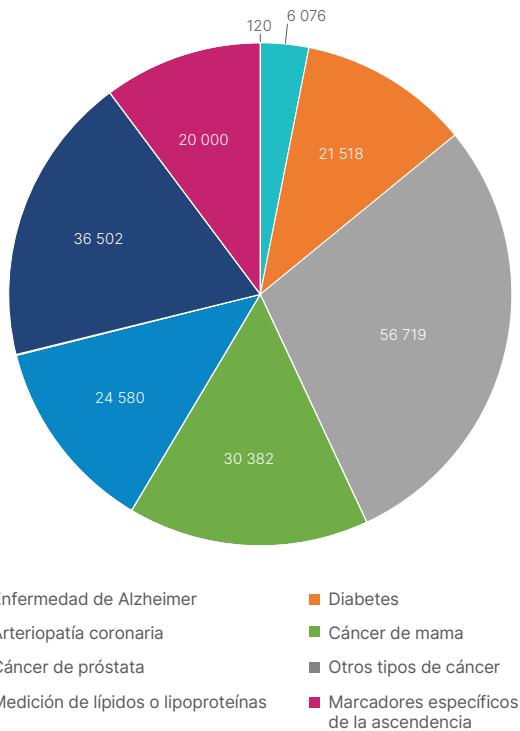


Figura 1: El contenido de refuerzo de la PRS abarca múltiples fenotipos. El contenido adicional de la PRS incluido en Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 incorpora marcadores de SNP en una amplia variedad de categorías de enfermedades. Un subconjunto de marcadores garantiza un rendimiento óptimo para poblaciones ascendentes específicas. Obsérvese que, dado que se comparten numerosos marcadores entre los fenotipos, la suma de estos valores no es igual a 160 000 marcadores.

Contenido exónico panétnico excepcional

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 incluye un marcado mejorado en las regiones exónicas y una mejor cobertura para asignar con precisión los locus de estudios de asociación del genoma completo (GWAS, Genome-Wide Association Studies) con asociaciones de enfermedades o rasgos previamente identificados. Se han recopilado más de 400 000 marcadores de exomas de 36 000 personas de grupos étnicos diversos, incluidos afroestadounidenses, hispanos, isleños del Pacífico, asiáticos orientales, europeos y personas de ascendencia mixta. La array también incorpora contenido exónico diverso de la base de datos ExAC/gnomAD,⁸ incluidos marcadores interpoblacionales y específicos de poblaciones con funcionalidad o pruebas sólidas de asociación (Tabla 4).

Tabla 4: Cobertura exónica entre poblaciones

Poblaciones ^{a,b}	N.º de marcadores
NFE	346 340
EAS	146 281
AMR	272 178
AFR	257 690
SAS	224 431
NFE/EAS/AMR/AFR/SAS	69 432

a. internationalgenome.org/category/population.
 b. Basado en gnomAD, gnomad.broadinstitute.org.

Abreviaturas: NFE, europeos no finlandeses; EAS, asiáticos orientales; AMR, estadounidenses de origen mixto; AFR, africanos; SAS, sudasiáticos.

Marcadores de CC para el seguimiento de muestras

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 incluye aproximadamente 10 000 marcadores de control de calidad (CC). Este contenido de los marcadores de CC posibilita importantes funciones de seguimiento de las muestras, la determinación de la ascendencia y la estratificación para facilitar la realización de estudios de mayor productividad (Figura 2).

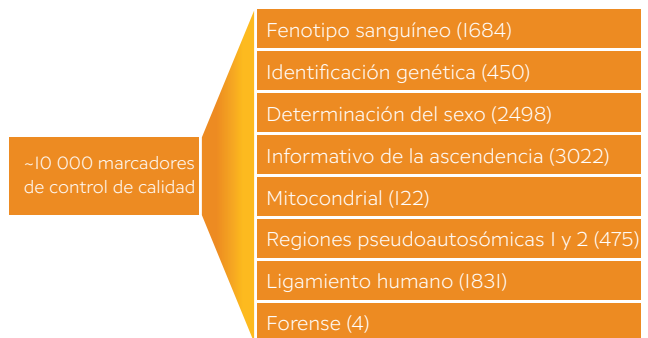


Figura 2: Marcadores de CC según la categoría. Las variantes de CC en la array hacen posibles distintas formas de llevar un seguimiento de las muestras, por ejemplo, por determinación del sexo, por ascendencia continental, por identificación humana y otras.

Flujo de trabajo de alta productividad

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 usa el formato de BeadChip de ocho muestras de Infinium, que permite que los laboratorios redimensionen eficientemente su trabajo en función de sus necesidades. Para lograr un procesamiento flexible con una gran productividad, el ensayo de LCG permite procesar hasta 1728 muestras por semana con un iScan™ System único. El ensayo Infinium ofrece un flujo de trabajo de tres días que permite a los investigadores recopilar datos y comunicarlos rápidamente con un tiempo de participación activa mínimo (Figura 3).

Ensayo de confianza y de gran calidad

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 emplea el proceso químico de confianza del ensayo Infinium para ofrecer los mismos datos reproducibles y de gran calidad (Tabla 5) que llevan proporcionando las arrays de genotipado de Illumina durante más de una década. Además, la elevada relación señal/ruido de las llamadas de genotipado individuales del ensayo Infinium permite acceder a llamadas de variantes en el número de copias (CNV, Copy Number Variation) del genoma completo.

Tabla 5: Rendimiento de los datos y separación

Rendimiento de los datos	Valor ^{a,b}	Especificaciones del producto ^c	
Índice de llamada	99,78 %	>99 % de media	
Reproducibilidad	99,99 %	>99,90 % de media	
Separación			
Separación (kb)	Media	Mediana	90° % ^d
	1,439	0,619	3,680

- a. Valores derivados del genotipado de 2228 muestras de referencia de HapMap.
- b. Valores previstos para proyectos típicos con protocolos estándares de Illumina. Se excluyen las muestras tumorales y muestras preparadas mediante protocolos no estándares.
- c. Se excluyen los marcadores del cromosoma Y en muestras femeninas.
- d. Basado en los resultados del conjunto de muestras de GenTrain.



Figura 3: Flujo de trabajo de LCG de Infinium. El ensayo Infinium usa un flujo de trabajo rápido de tres días que requiere un tiempo de participación activa mínimo.

Opciones de contenido flexibles

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 se puede personalizar para incorporar tipos de bolas que satisfagan las necesidades específicas de investigación de la PRS. El contenido complementario se puede seleccionar y exportar desde el catálogo de PGS e incorporarse al BeadChip usando DesignStudio™ Microarray Assay Designer, a fin de aumentar la cobertura de variantes.

Potente solución de análisis para la investigación de la PRS

Los datos generados por Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 son compatibles con el análisis mediante [BaseSpace™ Sequence Hub](#) construido sobre la fiable infraestructura basada en la nube [Illumina Connected Analytics \(ICA\)](#). En el caso de los laboratorios que buscan un flujo de trabajo completo de predicción de genotipo-riesgo, Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 se analiza de forma óptima con el módulo de software Predict. Este software de análisis fácil de usar está diseñado para emitir valores de la PRS individuales en relación con los calculados en una población de referencia con fenotipos conocidos para generar informes de la PRS automatizados con predicción de riesgo. El módulo de software Predict puede calcular hasta 24 PRS en un máximo de 1152 muestras en 15 minutos, para una determinación de la PRS rápida y precisa basada en la ascendencia.

Resumen

Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 forma parte de un completo conjunto de herramientas que proporciona una solución optimizada de predicción de genotipo-riesgo para la investigación de la SRP. El conjunto de herramientas de la PRS de Illumina simplifica el proceso de desarrollo de la PRS, lo que ahorra meses de análisis de datos que suelen realizar bioinformáticos especializados. Esta solución versátil es compatible con aplicaciones de baja y alta productividad y se puede personalizar para la generación de informes de la PRS amplios y basados en la ascendencia.

Información adicional

Puntuaciones de riesgo poligénico, illumina.com/areas-of-interest/complex-disease-genomics/polygenic-risk-scores

Infinium Global Diversity Array, illumina.com/products/by-type/microarray-kits/infinium-global-diversity

Módulo de software Predict, illumina.com/products/by-type/informatics-products/polygenic-risk-score-software

Datos para realizar pedidos

Producto	N.º de catálogo
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 (16 samples)	20090683
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 (48 samples)	20090684
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 (96 samples)	20090685
Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 (384 samples)	20090686

Bibliografía

1. Khera AV, Chaffin M, Aragam KG, et al. [Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations](#). Nat Genet. 2018; 50(9): 1219-1224. doi:10.1038/s41588-018-0183-z
2. Lambert SA, Abraham G, Inouye M. [Towards clinical utility of polygenic risk scores](#). Hum Mol Genet. 2019; 28(R2): R133-R142. doi:10.1093/hmg/ddz187
3. Lewis CM, Vassos E. [Polygenic risk scores: from research tools to clinical instruments](#). Genome Med. 2020; 12(1): 44. doi:10.1186/s13073-020-00742-5
4. Sitio web de The Polygenic Score (PGS) Catalog. [pgscatalog.org](#). Fecha de consulta: 18 de agosto de 2022.
5. Buniello A, MacArthur JAL, Cerezo M, et al. [The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019](#). Nucleic Acids Res. 2019; 47(D1): D1005-D1012. doi:10.1093/nar/gky1120
6. NCBI. Sitio web de la ClinVar Database. [ncbi.nlm.nih.gov/clinvar](#). Fecha de consulta: 18 de agosto de 2022.
7. Green RC, Berg JS, Grody WW, et al. [ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing](#). Genet Med. 2013; 15(7): 565-574. doi:10.1038/gim.2013.73
8. Broad Institute. Sitio web de Genome Aggregation Database (gnomAD) Browser. [gnomad.broadinstitute.org](#). Fecha de consulta: 18 de agosto de 2022.

illumina®

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

© 2023 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-01187 ESP V1.0