

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

Una soluzione versatile e scalabile per la ricerca sul punteggio PRS multi-etnico

- Contenuto aggiornato che include 160.000 marcatori PRS per coprire un ampio gruppo di patologie
- Copertura esaustiva dell'intero genoma per una determinazione altamente accurata del punteggio PRS pan-etnico
- Compatibile con il modulo software Predict che semplifica il calcolo del punteggio PRS e la creazione di report

illumina®

Introduzione

L'accurata valutazione del rischio di patologia è un aspetto chiave della medicina di prevenzione e specializzata. Rimane tuttavia difficile la predizione del rischio per le patologie con eredità poligenica poiché su queste influiscono da centinaia a migliaia di varianti genetiche che agiscono assieme a fattori ambientali. I punteggi di rischio poligenico (PRS, Polygenic Risk Score), noti anche come punteggi poligenici o punteggi di rischio genetico, sono indicatori numerici che rappresentano la predisposizione genetica di un individuo a sviluppare un tratto o una patologia.¹ Il punteggio PRS è uno strumento utile per i ricercatori che studiano la stratificazione del rischio e per comprendere meglio chi, molto probabilmente, potrebbe trarre vantaggi da ulteriore monitoraggio o interventi di prevenzione precoce.¹⁻³

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 (Tabella 1) è una soluzione basata su microarray a elevate prestazioni che fornisce un flusso di lavoro flessibile e scalabile per supportare gli studi sul punteggio PRS. Il BeadChip a 8 campioni, costruito sul backbone del comprovato Infinium Global Diversity Array, dispone di un contenuto di genotipizzazione aggiornato e include 160.000 marker per un'accurata determinazione del punteggio PRS (Tabella 2). I dati ottenuti dal BeadChip vengono analizzati in modo efficace dal modulo software Predict. In questo modo la ricerca di patologie poligeniche ha a disposizione una soluzione completa a partire dal genotipo fino al punteggio di rischio.

Contenuto ottimizzato per la ricerca sul punteggio PRS

Il backbone dell'intero genoma dell'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 contiene circa 1,9 milioni di marker rilevanti dal punto di vista scientifico. Questo contenuto esaustivo dell'Infinium Global Diversity Array-8 v1.0, l'array di genotipizzazione utilizzato in [All of Us Research Program](#), è integrato con 160.000 marker PRS del [Polygenic Score \(PGS\) Catalog](#)⁴ in continua crescita per fornire una copertura ampliata su regioni genomiche di elevato valore (Tabella 2, Tabella 3). Questi marker, selezionati in collaborazione con [Allelica](#), forniscono una copertura eccellente dei sei disordini chiave PRS, inclusi diabete di tipo 1 e 2, cancro al seno, cancro al colon, cancro della prostata, coronaropatia e malattia di Alzheimer.

Per generare il contenuto aggiuntivo per PRS, i campioni degli array sono stati messi a confronto con i punteggi del gold standard [1000 Genomes Project](#) e i punteggi comuni del PGS Catalog (ad esempio, MAV313) per tutti i punteggi PRS di Allelica. Lo scopo di questo confronto era identificare le varianti che presentano gli errori più comuni nell'imputazione confrontando i risultati di imputazione con 30× di dati del sequenziamento dell'intero genoma (WGS, Whole-Genome Sequencing).

Per sviluppare l'array, sono state selezionate le varianti che influivano maggiormente sul punteggio PRS e con la qualità di imputazione più bassa. Queste varianti sono state incluse direttamente nell'array per essere genotipizzate invece di essere imputate. Il contenuto aggiuntivo PRS include le varianti con la dimensione dell'effetto (effect size) più alta specifica per ogni informazione ancestrale nei punteggi multi-ancestrali di Allelica per ottimizzare il valore dell'array per gli individui appartenenti a tutte le ascendenze e aumentare la concordanza per il punteggio PRS.

Tabella 1: Descrizione generale dell'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0

Caratteristica	Descrizione
Specie	Umana
Numero totale di marker ^a	2.028.571
Numero di campioni per BeadChip	8
Requisito di input di DNA	200 ng
Replicati SNP	15
Numero di SNP necessari per identificare CNV	10
Chimica del saggio	Infinium LCG
Supporto strumento	iScan System
Processività massima per campione con l'iScan System ^b	Circa 1.728 campioni/settimana
Durata scansione per campione ^b	3-5 minuti

a. Il contenuto include circa 1,9 milioni di marker dal backbone dell'intero genoma dell'Infinium Global Diversity Array-8 v1.0 oltre a 160.000 marker specifici per il punteggio PRS.

b. Valori approssimativi. Durata della scansione e processività massima variano in base alle configurazioni del laboratorio e del sistema.

Tabella 2: Contenuto di elevato valore dell'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0

Contenuto	N. di marker ^a	Applicazione/note di ricerca
Geni hg19 RefSeq	1.123.407	Tutti i geni noti
Hg19 RefSeq ± 10 kb	1.266.608	Regioni regolatorie ^b
Promotori RefSeq	48.249	2 kb a monte per includere le regioni del promotore
Esonico ADME	16.528	Assorbimento, distribuzione, metabolismo ed eliminazione del farmaco
Geni hg19 ADME	32.117	
Hg19 ADME ± 10 kb	37.468	Include le regioni regolatorie
Marker HLA	17.595	Difesa contro la patologia, rigetto del trapianto, malattie autoimmuni
Geni hg19 HLA	1.297	
Marker MHC	22.783	
Geni hg19 COSMIC	1.079.088	Mutazioni somatiche nel cancro
Geni hg19 GO	324.039	Annotazione Gene Ontology

- a. Il numero di marker per ogni categoria è soggetto a modifiche.
 b. Di tutti i geni noti.

Abbreviazioni: ADME, assorbimento, distribuzione, metabolismo ed eliminazione; COSMIC, catalogo delle mutazioni somatiche nel cancro; GO, database di ontologia genica; hg19, genoma umano versione 19; HLA, antigene leucocitario umano; MHC, complesso maggiore di istocompatibilità; RefSeq, database delle sequenze di riferimento del National Center for Biotechnology Information.

Tabella 3: Informazioni sui marker

Categoria di marker	N. di marker ^a
Marker esonici ^a	530.644
Marker nonsenso ^b	28.287
Marker missenso ^b	396.783
Marker sinonimi ^b	33.442
SNP silenti ^b	40.964
Marker di DNA mitocondriale ^b	1.346
Indel ^c	37.289
Marker crom. X ^c	62.103
Marker crom. Y ^c	6.449
Marker PAR/omologhi ^c	5.485

- a. RefSeq, NCBI Reference Sequence Database, ncbi.nlm.nih.gov/refseq.
 b. Confrontato rispetto al Genome Browser della UCSC, genome.ucsc.edu.
 c. NCBI Genome Reference Consortium, Version GRCh37 (hg19), ncbi.nlm.nih.gov/assembly/GCF_000001405.13.

Abbreviazioni: Indel: inserzione/delezione; NCBI: National Center for Biotechnology Information; PAR: regione pseudoautosomica; SNP: polimorfismo a singolo nucleotide; UCSC: University of California Santa Cruz.

Ampia copertura delle varianti associate alla patologia

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 fornisce la copertura delle varianti selezionate dal catalogo GWAS del National Human Genome Research Institute e dell'European Bioinformatics Institute (NHGRI-EBI)⁵ che rappresentano un'ampia gamma di classificazioni di fenotipi e patologie (Figura 1). La selezione delle varianti include anche una gamma di classificazioni di patologie basate sulle annotazioni del ClinVar⁶ e dell'American College of Medical Genetics (ACMG)⁷.

I database clinici come ClinVar sono in costante evoluzione con l'aggiunta di nuove varianti e varianti le cui designazioni sono state modificate in "Patogena" o "Probabilmente patogena". Il BeadChip fornisce la copertura aggiornata di molte di queste varianti di elevato valore contenute nei database annotati. Questo contenuto esaustivo fornisce efficaci opportunità per i ricercatori interessati a studiare diverse popolazioni per testare e convalidare le associazioni precedentemente individuate in popolazioni europee.

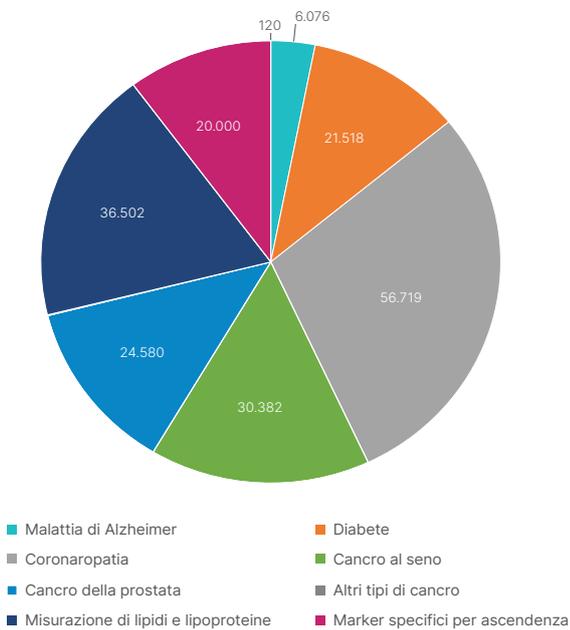


Figura 1: Il contenuto aggiuntivo PRS copre diversi fenotipi: il contenuto aggiuntivo PRS incluso nell'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 dispone di marker SNP su un'ampia gamma di categorie di patologia. Un sottogruppo di marker assicura prestazioni ottimali per determinate popolazioni ancestrali. Tenere presente che la somma di questi valori non è pari a 160.000 marker poiché molti marker sono condivisi sui fenotipi.

Eccellente contenuto esonico pan-etnico

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 include una migliore marcatura nelle regioni esoniche e una copertura arricchita per mappare i loci ottenuti da studi di associazione dell'intero genoma (GWAS, Genome-Wide Association Study) con associazioni a patologie o tratti precedentemente identificate. Più di 400.000 marker di esoni sono stati raccolti da 36.000 individui di diversi gruppi etnici, inclusi afro-americani, ispanici, abitanti di isole pacifiche, asiatici dell'est, europei e individui di discendenza mista. L'array fornisce inoltre contenuto esonico diversificato ottenuto dal database ExAC/gnomAD,⁸ inclusi i marker specifici per popolazione e tra popolazioni con funzionalità o forte prova di associazione (Tabella 4).

Tabella 4: Copertura esonica nelle popolazioni

Popolazione(i) ^{a,b}	N. di marker
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. Basato su gnomAD, gnomad.broadinstitute.org.

Abbreviazioni: NFE, europeo non finlandese; EAS, asiatico dell'est; AMR, americano misto; AFR, africano; SAS, asiatico del sud.

Marker QC per il monitoraggio dei campioni

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 include circa 10.000 marker di controllo qualità (QC, Quality Control). Questo contenuto di marker QC consente importanti funzioni di monitoraggio dei campioni, di determinazione ancestrale e di stratificazione per facilitare studi a elevata processività (Figura 2).



Figura 2: Marker di controllo qualità (QC) per categoria: le varianti di controllo qualità presenti sull'array consentono diverse funzionalità per il monitoraggio dei campioni come la determinazione del sesso, l'origine ancestrale continentale, l'identificazione umana e altro.

Flusso di lavoro a elevata processività

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 utilizza il BeadChip Infinium nel formato a 8 campioni che consente ai laboratori di scalare in modo efficiente. Il saggio Infinium LCG offre una processività flessibile grazie alla capacità di analizzare fino a 1.728 campioni alla settimana utilizzando un singolo iScan™ System. Il saggio Infinium fornisce un flusso di lavoro di 3 giorni che consente ai ricercatori di raccogliere e riportare rapidamente dati con interventi manuali minimi (Figura 3).

Saggio affidabile e di elevata qualità

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 utilizza l'affidabile chimica del saggio Infinium per fornire dati della stessa elevata qualità e riproducibilità (Tabella 5) che gli array di genotipizzazione di Illumina hanno fornito per più di dieci anni. Inoltre, l'elevato rapporto segnale-rumore delle singole identificazioni di genotipizzazione ottenute dal saggio Infinium consente di accedere alle identificazioni delle variazioni del numero di copie (CNV, Copy Number Variation) dell'intero genoma.

Tabella 5: Prestazioni dei dati e spaziatura

Prestazioni dei dati	Valore ^{a,b}	Specifica del prodotto ^c	
Percentuale di identificazione	99,78%	> 99% media	
Riproducibilità	99,99%	> 99,90% media	
Spaziatura			
Spaziatura (kb)	Media	Mediana	90° ^d
	1,439	0,619	3,680

a. I valori derivano dalla genotipizzazione di 2.228 campioni di riferimento HapMap.
 b. Valori previsti per progetti tipici usando protocolli standard di Illumina. Sono esclusi i campioni di tumore e i campioni preparati con protocolli non standard.
 c. Per i campioni femminili, esclude i marker del cromosoma Y.
 d. Basato sui risultati ottenuti dal set di campioni GenTrain.



Figura 3: Flusso di lavoro dell'Infinium LCG: il saggio Infinium utilizza un rapido flusso di lavoro di 3 giorni e richiede interventi manuali minimi.

Opzioni di contenuto flessibili

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 può essere personalizzato per incorporare tipi di microsfera in base alle determinate esigenze della ricerca PRS. Contenuto aggiuntivo può essere selezionato ed esportato dal catalogo PGS e incorporato nel BeadChip utilizzando il DesignStudio™ Microarray Assay Designer, per aumentare la copertura di varianti.

Pipeline di analisi efficace per la ricerca PRS

I dati generati dall'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 sono compatibili con l'analisi eseguita sul [BaseSpace™ Sequence Hub](#) basato sull'affidabile infrastruttura sul cloud [Illumina Connected Analytics \(ICA\)](#). Per i laboratori che sono alla ricerca di un flusso di lavoro completo per la predizione genotipo-rischio, l'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 può essere analizzato in modo ottimale con il modulo software Predict. Questo software di analisi di facile utilizzo è progettato per generare singoli valori PRS rispetto ai valori calcolati su una popolazione di riferimento con fenotipi noti e creare report PRS automatici contenenti la predizione di rischio. Il modulo software Predict può calcolare fino a 24 punteggi PRS su un massimo di 1.152 campioni in 15 minuti, per una determinazione del punteggio PRS rapida, accurata e con informazioni dal punto di vista ancestrale.

Riepilogo

L'Infinium Global Diversity Array with Polygenic Risk Score Content-8 v1.0 fa parte di una serie di strumenti completa che fornisce una soluzione ottimizzata per la predizione di rischio a partire dal genotipo per la ricerca PRS. La serie di strumenti di Illumina per PRS semplifica il processo di sviluppo del PRS facendo risparmiare mesi di analisi dei dati eseguite da bioinformatici dedicati. La soluzione è flessibile e compatibile con applicazioni sia a bassa processività sia ad alta processività e può essere personalizzata per un'ampia gamma di report PRS contenenti informazioni dal punto di vista ancestrale.

Maggiori informazioni

Punteggio di rischio poligenico, 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.

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

Informazioni per gli ordini

Prodotto	N. di catalogo
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

Bibliografia

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. Sito Web del Polygenic Score (PGS) Catalog. pgscatalog.org. Consultato il 18 agosto 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. Sito Web del ClinVar Database. ncbi.nlm.nih.gov/clinvar. Consultato il 18 agosto 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. Sito Web del Genome Aggregation Database (gnomAD) Browser. gnomad.broadinstitute.org. Consultato il 18 agosto 2022.

illumina®

Numero verde 1.800.809.4566 (U.S.A.) | Tel. +1.858.202.4566
techsupport@illumina.com | www.illumina.com

© 2023 Illumina, Inc. Tutti i diritti riservati. Tutti i marchi di fabbrica sono di proprietà di Illumina, Inc. o dei rispettivi proprietari. Per informazioni specifiche sui marchi di fabbrica, consultare la pagina Web www.illumina.com/company/legal.html.
M-GL-01187 ITA V1.0