DATA SHEET

Infinium™ Global Clinical Research Array with Enhanced PGx-24 v1.0

High-resolution genotyping array for pharmacogenomics studies and clinical research

- Genome-wide scaffold to detect common and low-frequency variants across a range of phenotypes
- Comprehensive coverage of ~1.2M annotated variants from public databases
- Additional ~35K markers supporting advanced pharmacogenomic research
- Automated workflow for medium- and high-throughput project needs

illumina®

Introduction

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 is a high-density BeadChip developed for pharmacogenomic (PGx) research, polygenic risk score development, and genetic disease research. Each BeadChip provides accurate assessment of ~1.2M backbone markers and 41,875 PGx-focused markers for up to 24 human samples (Figure 1, Table 1).

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 includes more than 35,900 absorption, distribution, metabolism, and excretion (ADME) markers spanning more than 2000 genes and exceptional coverage of priority level A and B Clinical Pharmacogenetics Implementation Consortium (CPIC) variants (Figure 2, Table 2).[1,2] The Infinium EX workflow with PGx also features a targeted gene amplification step that supports detection of high-impact PGx genes like *CYP2D6*, *CYP2C19*, and *DPYD* that have historically been challenging to assess.

Table 1: Product specifications

Feature	Description
Species	Human
Total number of markers[a]	1,185,155
Enhanced PGx markers	41,875
Number of samples per BeadChip	24
DNA input requirement	200 ng
Capacity for custom bead types	10K
Assay chemistry	Infinium EX
Instrument support[a]	iScan™ System Infinium Amplification System Infinium Automated Pipetting System with ILASS
Maximum iScan System sample throughput	~5760 samples/week
Scan time per BeadChip[b]	~30 minutes

a. The iScan System and Infinium Automated Pipetting System with ILASS are required for a fully supported solution. The Infinium Amplification System is optional.
b. Approximate values, scan times, and maximum throughput will vary depending on laboratory and system configurations.

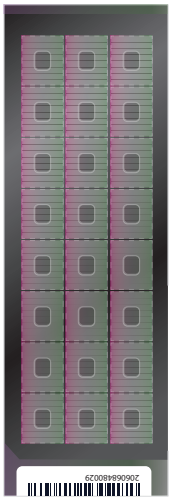

Figure 1: Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0— Features a multi-ethnic backbone with ~1.2M clinically relevant markers. The BeadChip uses Infinium EX chemistry for fast and accurate results.

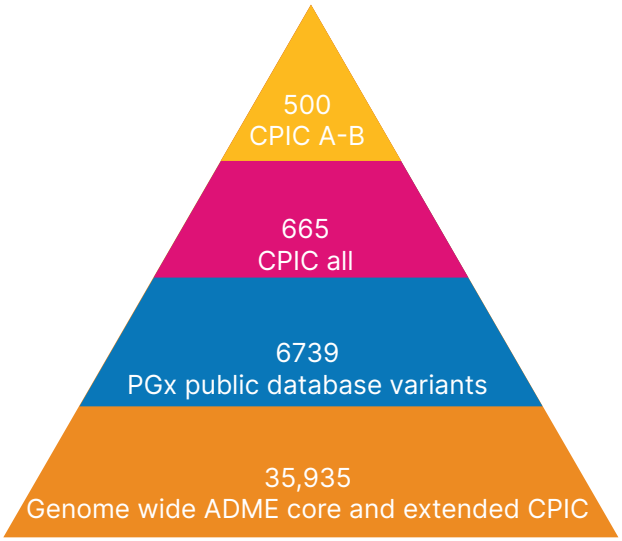

Figure 2: Broad spectrum of pharmacogenomics markers— Clinical research content developed from an extensive list of pharmacogenomics markers selected based on CPIC guidelines and the PharmGKB database.[1,2] Content includes variants annotated in PharmGKB, genome-wide PGx coverage, extended ADME genes,[3] CPIC level A genes, and targeted imputation tag SNPs.

Table 2: High-value content from key research databases

Content	No. of markers[a]	Research application/note	Content	No. of markers[a]	Research application/note
ACMG[4] 59 2016 gene coverage	31,996	Variants with known clinical significance identified from clinical WGS and WES samples	GO[9] CVS genes	204,889	Cardiovascular conditions
ACMG 59 all annotations	47,677		Database of Genomic Variants[10]	926,764	Genomic structural variation
ACMG 59 pathogenic	11,580		eQTLs[11]	4783	Genomic loci regulating mRNA expression levels
ACMG 59 likely pathogenic	4789		Fingerprint SNPs[12]	435	Human identification
ACMG 59 benign	3629		gnomAD[13] exome	134,595	WES and WGS results from unrelated individuals from various studies
ACMG 59 likely benign	10,072		HLA genes[14]	1286	Disease defense, transplant rejection, and autoimmune disorders
ACMG 59 VUS	12,213		Extended MHC[14 c]	13,711	Disease defense, transplant rejection, and autoimmune disorders
ADME[3] core and extended + CPIC genes	35,935	Drug absorption, distribution, metabolism, and excretion	KIR genes[5]	61	Autoimmune disorders and disease defense
ADME core and extended + CPIC genes +/- 10 kb	41,875	Includes regulatory regions	Neanderthal SNPs[15]	2775	Neanderthal ancestry and human population migration
AIMs[b]	2853	Ancestry-informative markers	Newborn/carrier screening gene coverage	48,837	Genes associated with childhood diseases included in the TruSight™ Inherited Disease Sequencing Panel[16]
APOE[5]	20	Cardiovascular disease, Alzheimer's disease, and cognition	NHGRI-EBI GWAS catalog[17]	48,904	Markers from published GWAS
Blood phenotype genes[6]	2337	Blood phenotypes	PharmGKB[1,4] all	5296	Human genetic variation associated with drug responses
ClinVar[7] variants	123,136	Relationships among variation, phenotypes, and human health	PharmGKB level 1A	46	
ClinVar pathogenic	33,695		PharmGKB level 1B	8	
ClinVar likely pathogenic	16,328		PharmGKB level 2A	59	
ClinVar benign	26,944		PharmGKB level 2B	51	
ClinVar likely benign	18,806		PharmGKB level 3	1955	
ClinVar VUS	21,297		PharmGKB level 4	489	
COSMIC[8] genes	553,745	Somatic mutations in cancer	RefSeq[18] 3' UTRs	24,302	3' untranslated regions[d]
CPIC[2] all	665	Variants with potential guidelines to optimize drug therapy	RefSeq 5' UTRs	11,663	5' untranslated regions[d]
CPIC-A	468		RefSeq All UTRs	34,928	Untranslated regions[d]
CPIC-A/B	500		RefSeq	615,027	All known genes
CPIC-B	29		RefSeq +/- 10 kb	715,588	Regulatory regions[d]
CPIC-C	45		RefSeq Promoters	26,878	2 kb upstream to include promoter regions[d]
CPIC-C/D	2		RefSeq Splice Regions	6993	Variants at splice sites[d]
CPIC-D	61				

a. The number of markers for each category may be subject to change.
b. Based on internal calculations.
c. Extended MHC is an 8 Mb region.
d. Of all known genes.

Abbreviations: ACMG: American College of Medical Genetics; ADME: absorption, distribution, metabolism, and excretion; AIM: ancestry-informative marker; APOE: apolipoprotein E; COSMIC: catalog of somatic mutations in cancer; CPIC: Clinical Pharmacogenetics Implementation Consortium; EBI: European Bioinformatics Institute; eQTL: expression quantitative trait loci; gnomAD: Genome Aggregation Database; GO CVS: gene ontology annotation of the cardiovascular system; GWAS: genome-wide association study; HLA: human leukocyte antigen; KIR: killer cell immunoglobulin-like receptor; MHC: major histocompatibility complex; NHGRI: national human genome research institute; PharmGKB: Pharmacogenomics Knowledgebase; RefSeq: NCBI Reference Sequence Database; UTR: untranslated region; VUS: variant of unknown significance; WES: whole-exome sequencing; WGS: whole-genome sequencing.

Infinium EX chemistry workflow

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 uses advanced Infinium EX chemistry for a rapid and accurate assay workflow. The Infinium EX chemistry workflow uses formamide-free reagents and is compatible with the Infinium Amplification System and the Infinium Automated Pipetting System with ILASS. The automated workflow substantially enhances scalability, decreases hands-on time, and reduces the potential for human error. The PGx workflow also features a targeted gene amplification step that increases performance for the detection of difficult targets, such as *CYP2D6*, *CYP2B6*, and *TPMT*.

Diverse backbone with enhanced exonic coverage

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 is built on a global high-density SNP backbone that is optimized for cross-population imputation coverage. The genome-wide content includes enhanced tagging in exonic regions and enriched coverage of loci from genome-wide association studies (GWAS) with known disease or trait associations (Figure 2, Table 3).

Table 3: Marker information

Marker categories			
Exonic markers[a]			161,451
Intronic markers[a]			492,186
Nonsense markers[b]			7088
Missense markers[b]			53,660
Synonymous markers[b]			14,806
Mitochondrial markers[b]			1038
Indels[c]			20,048
Sex chromosomes[c]	X	Y	PAR/homologous
	39,005	4435	824

a. RefSeq-NCBI Reference Sequence Database.[18]
b. Compared against the UCSC Genome Browser.[5]
c. NCBI Genome Reference Consortium, Version GRCh38.[19]

More than 130,000 exome markers were selected from individuals representing diverse ethnic backgrounds, including African Americans, Hispanics, Pacific Islanders, East Asians, Europeans, and individuals of mixed ancestry. The array also features exonic content from populations in the ExAC database, including cross-population and population-specific markers with functionality annotations or strong evidence for association (Table 4). The inclusive design allows for multiple applications, including polygenic risk scoring, nutrigenomics research, and clinical validation studies based on reported variants.

Table 4: Exonic coverage across populations

Population(s)[a,b]	No. of markers
EUR	102,826
EAS	52,568
AMR	78,427
AFR	72,275
SAS	70,708
EUR/EAS/AMR/AFR/SAS	36,289

a. www.internationalgenome.org/category/population.
b. Based on gnomAD, gnomad.broadinstitute.org/.

Abbreviations: EUR: Europe; EAS: East Asian; AMR: Ad Mixed American; AFR: African; SAS: South Asian.

Broad coverage of variants with known disease associations

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 is designed for high-value clinical research applications. It provides coverage of variants selected from the National Human Genome Research Institute genome-wide association studies (NHGRI-GWAS) catalog[17] representing an extensive range of phenotypes and disease classifications. This content provides powerful opportunities for researchers interested in studying diverse populations.

Clinical research content on the array enables validation of previously identified disease associations, risk profiling, predictive screening research, and PGx studies.

Variant selection includes a range of pathology classifications based on ClinVar and American College of Medical Genetics (ACMG) annotations.[3] The content covers an extensive range of phenotypes and disease classifications based on ClinVar and the NHGRI-GWAS catalog (Figure 3). Markers cover ACMG and ClinVar database variants with a range of phenotypes pathogenic, likely pathogenic, and variants of unknown significance (VUS), as well as benign variants (Figure 4).

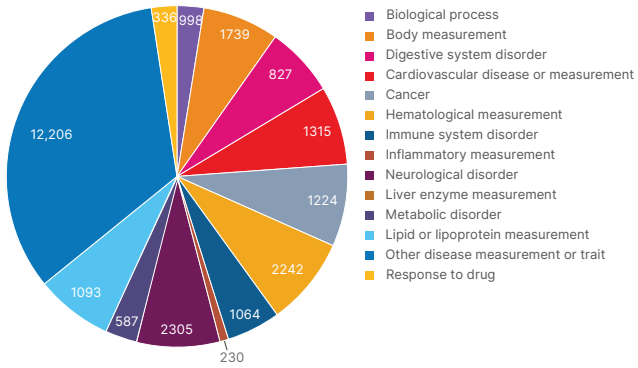

Figure 3: Disease research content covering diverse populations— The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 includes extensive coverage of phenotypes and disease classifications based on NHGRI GWAS database categories.

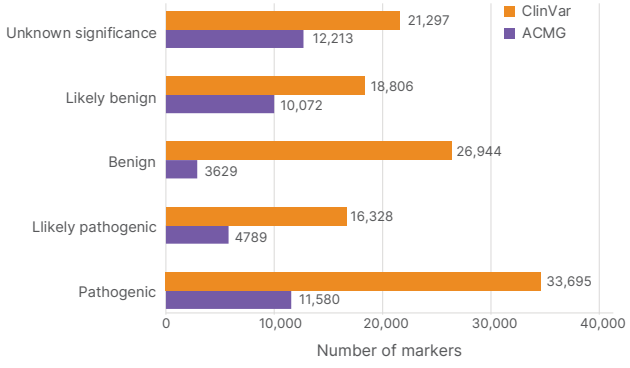

Figure 4: Distribution of variant pathology classifications according to ClinVar and ACMG annotations—Variants cover a range of pathogenic and nonpathogenic evidence.

Updated research content

Databases, such as ClinVar, are constantly evolving with the addition of new variants and as variants change designation to "pathogenic" or "likely pathogenic" categories. Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 provides updated coverage of many high-value variants contained within these annotated databases. Variants included on the array consist of markers with known disease association selected from ClinVar, PharmGKB, and the NHGRI-EBI database.[8,18] The array also provides imputation-based tag SNPs for HLA alleles, extended MHC region, the KIR gene, and exonic content from the gnomAD database[5,13,14] (Table 2, Figure 5).

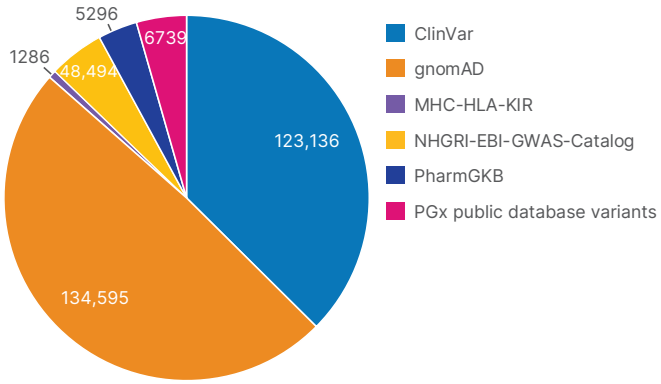

Figure 5: Clinical research content—Expertly selected clinical research content from key databases supports a broad range of applications.

QC markers for sample tracking and identification

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 includes ~9K quality control (QC) markers. QC markers on the array are selected to facilitate high-throughput studies and enable sample tracking functions, including ancestry determination and stratification (Figure 6).

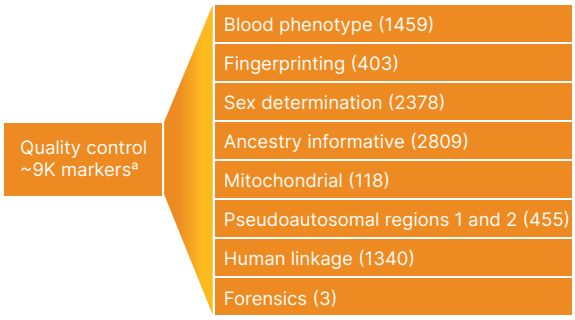

Figure 6: QC content by category—The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 contains ~9K markers enabling various sample tracking functions such as sex determination, continental ancestry, human linkage, and more.
a. Counts contain some markers that are represented in multiple QC categories.

High-performance assay

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 uses trusted Infinium EX chemistry to deliver a high-performance and accurate genotyping solution (Table 5). In addition, the high signal-to-noise ratio of the individual genotyping calls from the assay provides access to CNV calling and star allele annotation via DRAGEN™ Array secondary analysis. The Infinium EX workflow is also fast, providing results in ~2.5 days.

Table 5: Data performance and spacing

Data performance	Observed[a]	Product specification[b]	
Call rate	99.6%	> 99.0% avg	
Reproducibility	99.99%	> 99.9%	
Log R deviation	0.14[c]	< 0.30 avg[d]	
	Mean	Median	90th percentile[c]
Probe spacing	2.6 kb	1.3 kb	6 kb

a. Values are derived from genotyping 1394 HapMap reference samples.
b. Excludes Y chromosome markers for female samples.
c. Based on results from GenTrain sample set.
d. Value expected for typical projects using standard Illumina protocols.

Accurate, efficient secondary analysis

DRAGEN Array secondary analysis is recommended for analysis of the Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0. DRAGEN Array is a powerful bioinformatics software that uses cutting-edge data analysis tools to maximize genomic insights. The software includes SNP analysis, PGx star allele and variant coverage across 2400+ targets for over 50 genes, hybrid allele and allele-specific copy number detection, PGx CNV coverage on seven target genes across ten target regions, and more.

DRAGEN Array secondary analysis is fast, generating accurate results in multiple output file formats for easy downstream analysis, including the capability to generate VCF files from Infinium array-based assays in as little as 15 seconds per sample and full PGx analysis results in approximately 26 seconds per sample.

DRAGEN Array secondary analysis has two deployment options. A local analysis option provides a command-line interface for granular control. No specialized DRAGEN server or FPGA hardware is required for the local installation solution. A cloud-based package with an intuitive user interface is also available via BaseSpace™ Sequence Hub. This option offers easy access and additional functionality, including polygenic risk scoring for arrays.

Summary

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 provides an accurate, high-resolution genotyping assay optimized for PGx and clinical research applications. When combined with the Infinium Automated Pipetting System with ILASS and the Infinium Amplification System, the Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 offers a high-throughput option for labs looking to process large numbers of samples with minimal hands-on processing time for assays.

Learn more

The Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0

Infinium Automation Option Packages

DRAGEN Array secondary analysis

Ordering information

Infinium Global Clinical Research Array with Enhanced PGx-24 v1.0 kit	Catalog no.
24 samples	20065216
96 samples	20068339
1152 samples	20068340
DRAGEN Array secondary analysis	
DRAGEN Array Local - star allele annotation	20109885
DRAGEN Array Cloud - star allele annotation[a]	20109886

a. An Illumina Connected Analytics annual subscription is required for cloud analysis along with iCredits for data storage and analysis.

References

1. Whirl-Carrillo M, Huddart R, Gong L, et al. An Evidence-Based Framework for Evaluating Pharmacogenomics Knowledge for Personalized Medicine. *Clin Pharmacol Ther*. 2021;110(3):563-572. doi:10.1002/cpt.2350
2. Clinical Pharmacogenetics Implementation Consortium (CPIC). cpicpgx.org. Accessed September 23, 2024.
3. Gaedigk A, Casey ST, Whirl-Carrillo M, Miller NA, Klein TE. Pharmacogene Variation Consortium: A Global Resource and Repository for Pharmacogene Variation. *Clin Pharmacol Ther*. 2021;110(3):542-545. doi:10.1002/cpt.2321.
4. 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.
5. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996-1006. doi:10.1101/gr.229102
6. National Center for Biotechnology Information. Reference Sequence Blood Group Antigen Gene Mutation Database. https://ftp.ncbi.nlm.nih.gov/pub/mhc/rbc/Final%20Archive/. Accessed September 23, 2024.
7. National Center for Biotechnology Information. ClinVar Database. ncbi.nlm.nih.gov/clinvar. Accessed April 25, 2024.
8. Forbes SA, Bhamra G, Bamford S, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet*. 2008;Chapter 10:Unit-10.11. doi:10.1002/0471142905.hg1011s57
9. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25-29. doi:10.1038/75556.
10. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014;42(Database issue):D986-D992. doi:10.1093/nar/gkt958
11. National Center for Biotechnology Information. eQTL Database. www.ncbi.nlm.nih.gov/projects/gap/eqtl/index.cgi. Accessed January 23, 2020.
12. Rajeevan H, Osier MV, Cheung KH, et al. ALFRED: an allele frequency database for diverse populations and DNA polymorphisms--an update. *Nucleic Acids Res*. 2003;31(1):270-271. doi:10.1093/nar/gkg043.

13. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7
14. de Bakker PI, McVean G, Sabeti PC, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet*. 2006;38(10):1166-1172. doi:10.1038/ng1885
15. Green RE, Krause J, Briggs AW, et al. A draft sequence of the Neandertal genome. *Science*. 2010;328(5979):710-722. doi:10.1126/science.1188021.
16. Illumina (2017). TruSight Inherited Disease Sequencing Panel Data Sheet. support.illumina.com/content/dam/illumina-marketing/apac/china/documents/datasheet_trusight_inherited_disease.pdf. Accessed July 2016.
17. National Human Genome Research Institute. genome.gov. Accessed September 23, 2024.
18. RefSeq - NCBI Reference Sequence Database. ncbi.nlm.nih.gov/refseq. Accessed January 23, 2020.
19. NCBI Genome Reference Consortium. Version GRCh38. ncbi.nlm.nih.gov/grc/human. Accessed September 23, 2024.

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-00710 v1.0