DATA SHEET

Infinium™ Global Screening Array-48 v4.0

High-throughput genotyping solution for population genetics and precision medicine research

- Comprehensive coverage of over 650K annotated variants from public research databases
- Advanced, high-density bead types support additional throughput of 48 samples on a single BeadChip
- High-performance EX chemistry for high-throughput and scalable workflows

illumina®

Introduction

The Infinium Global Screening Array-48 v4.0 is a high-throughput BeadChip for population studies, variant screening, and precision medicine research (Figure 1). The BeadChip features a multiethnic genome-wide backbone with ~650K clinical research markers, quality control (QC) markers for sample tracking and classification, and up to 50K custom markers. The BeadChip uses Infinium EX chemistry for an accurate high-density assay that delivers results in under three days (Table 1). Content includes variants with established disease associations selected from key databases, including ClinVar, NHGRI, PharmGKB, and ExAC (Figure 2, Table 2).[1–4]

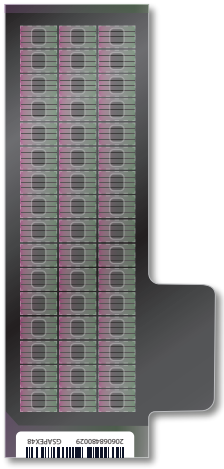

Figure 1: Infinium Global Screening Array-48 v4.0— The versatile BeadChip features a multiethnic backbone with 650,321 clinically relevant markers. The high-density Infinium EX chemistry supports rapid, high-throughput analysis.

Table 1: Product information

Feature	Description
Species	Human
Total number of markers[a]	650,321
Number of samples per BeadChip	48
DNA input requirement	100 ng
Assay chemistry	Infinium EX
Instrument support[a]	iScan System Infinium Amplification System Infinium Automated Pipetting System with ILASS[b]
Maximum iScan System sample throughput	~11,520 samples/week
Scan time per BeadChip[c]	~30 min

a. The Infinium EX chemistry workflow specifies the use of Infinium automation solutions.
b. ILASS: Illumina Lab Automation Software Solution.
c. Approximate values, scan times, and maximum throughput will vary depending on laboratory and system configurations.

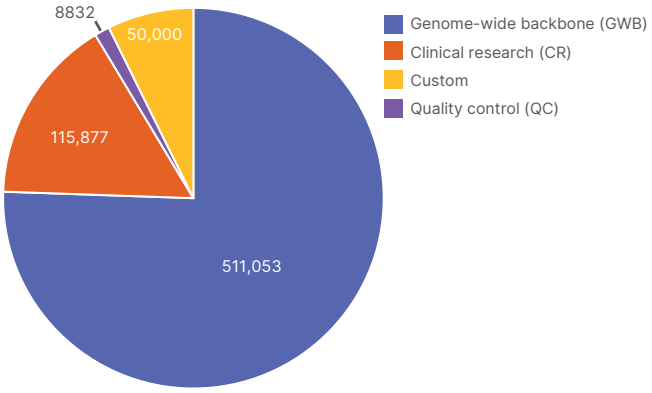

Figure 2: Annotated research database content summary— Distribution of unique markers for genome-wide coverage, clinical research, quality control (QC), and potential custom marker content.

Table 2: High-value content from key research databases

Content	No. of markers[a]	Research application/note	Content	No. of markers	Research application/note
ACMG[5] 59 2016 gene coverage	18,446	Variants with known clinical significance identified from clinical WGS and WES samples	GO[11] CVS genes	99,142	Cardiovascular conditions
ACMG 59 all annotations	4436		Database of Genomic Variants[12]	492,527	Genomic structural variation
ACMG 59 pathogenic	2588		eQTLs[13]	2648	Genomic loci regulating mRNA expression levels
ACMG 59 likely pathogenic	1111		Fingerprint SNPs[14]	429	Human identification
ACMG 59 benign	145		gnomAD[15] exome	73,179	WES and WGS results from unrelated individuals from various studies
ACMG 59 likely benign	203		HLA genes[16]	446	Disease defense, transplant rejection, and autoimmune disorders
ACMG 59 VUS	488		Extended MHC[16,c]	8156	Disease defense, transplant rejection, and autoimmune disorders
ADME[6] core and extended + CPIC genes	15,906	Drug absorption, distribution, metabolism, and excretion	KIR genes[7]	24	Autoimmune disorders and disease defense
ADME core and extended + CPIC genes +/- 10 kb	18,366	Includes regulatory regions	Neanderthal SNPs[17]	1498	Neanderthal ancestry and human population migration
AIMs[b]	2794	Ancestry-informative markers	Newborn/carrier screening gene coverage	24,051	Genes associated with childhood diseases included in the TruSight™ Inherited Disease Sequencing Panel[18]
APOE[7]	11	Cardiovascular disease, Alzheimer's disease, and cognition	NHGRI-EBI GWAS catalog[2]	30,999	Markers published from GWAS
Blood phenotype genes[8]	1786	Blood phenotypes	PharmGKB[3,19] all	4684	Human genetic variation associated with drug responses
ClinVar[1] variants	20,059	Relationships among variation, phenotypes, and human health	PharmGKB level 1A	218	
ClinVar pathogenic	3032		PharmGKB level 1B	7	
ClinVar likely pathogenic	1351		PharmGKB level 2A	33	
ClinVar benign	8072		PharmGKB level 2B	44	
ClinVar likely benign	2820		PharmGKB level 3	1720	
ClinVar VUS	4199		PharmGKB level 4	406	
COSMIC[9] genes	290,137	Somatic mutations in cancer	RefSeq[20] 3' UTRs	13,249	3' untranslated regions[d]
CPIC[10] all	480	Variants with potential guidelines to optimize drug therapy	RefSeq 5' UTRs	6000	5' untranslated regions[d]
CPIC-A	331		RefSeq All UTRs	18,677	Untranslated regions[d]
CPIC-A/B	3		RefSeq +/- 10 kb	378,074	Regulatory regions[d]
CPIC-B	17		RefSeq Promoters	14,181	2 kb upstream to include promoter regions[d]
CPIC-C	42		RefSeq Splice Regions	3178	Variants at splice sites[d]
CPIC-C/D	1				
CPIC-D	58				

a. The number of markers for each category may be subject to change based on internal calculations
b. The extended MHC is an 8 Mb region.
c. 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 Screening Array-48 v4.0 BeadChip uses advanced Infinium EX chemistry for a rapid and accurate assay workflow. The Infinium EX chemistry workflow is optimized for automation, 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 Infinium EX workflow is also fast, providing results in two to three days.

Diverse backbone with enhanced exonic coverage

The Infinium Global Screening Array-48 v4.0 BeadChip is built on a high-density SNP backbone that is optimized for cross-population imputation coverage. The updated 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			No. of markers
Exonic markers[a]			81,168
Intronic markers[a]			257,722
Nonsense markers[b]			5269
Missense markers[b]			45,829
Synonymous markers[b]			8476
Mitochondrial markers[b]			1089
Indels[c]			8471
Sex chromosomes[c]	X	Y	PAR/homologous
	27,036	3887	823

a. RefSeq-NCBI Reference Sequence Database.[19]
b. Compared against the UCSC Genome Browser.[6]
c. NCBI Genome Reference Consortium, Version GRCh37.[21]

More than 73,000 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).[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
NFE	27,941
EAS	32,336
AMR	46,554
AFR	43,304
SAS	41,270
NFE/EAS/AMR/AFR/SAS	22,769

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

Abbreviations: NFE: non-Finnish European; EAS: East Asian; AMR: Ad Mixed American; AFR: African; SAS: South Asian.

Broad coverage of variants with known disease associations

Content on the Infinium Global Screening Array-48 v4.0 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[2] representing an extensive range of phenotypes and disease classifications. The selection of markers provides extensive opportunities for researchers interested in studying diverse populations.

The clinical research content on the BeadChip enables validation of previously identified disease associations, risk profiling, predictive screening research, and pharmacogenomic studies.

Selected variants represent an extensive range of phenotypes and disease classifications based on ClinVar and the NHGRI-GWAS catalog (Figure 3).[1,2] The content also covers 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).

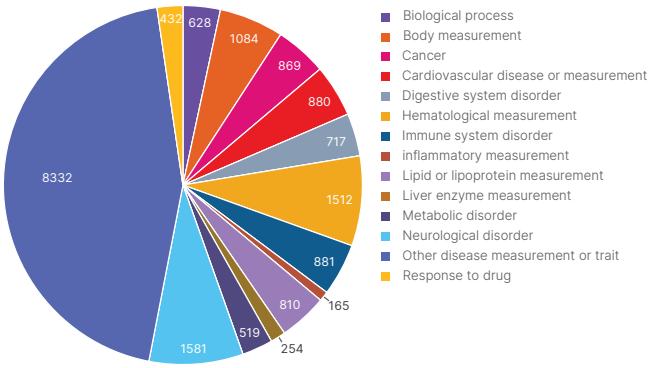

Figure 3: Disease research content covering diverse populations—The Infinium Global Screening Array-48 v4.0 includes extensive coverage of phenotypes and disease classifications based on NHGRI–GWAS database categories.

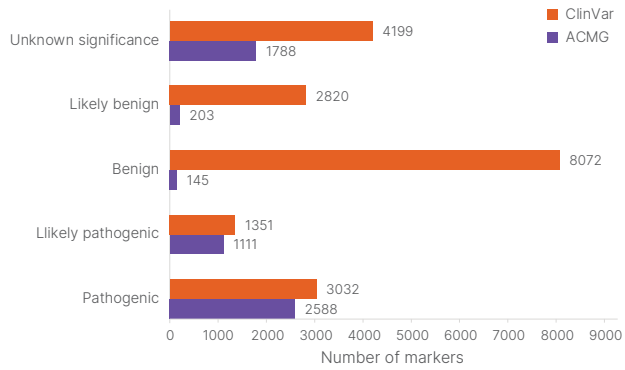

Figure 4: Distribution of variant pathology classifications according to ClinVar and ACMG annotations—Variants cover a range of pathogenic and nonpathogenic evidence.

Updated and 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. The Infinium Global Screening Array-48 v4.0 provides updated information for 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.[1–3] The BeadChip also provides imputation-based tag SNPs for HLA alleles, extended MHC region, the KIR gene, and exonic content from the gnomAD database (Table 2, Figure 5).[7,15,16]

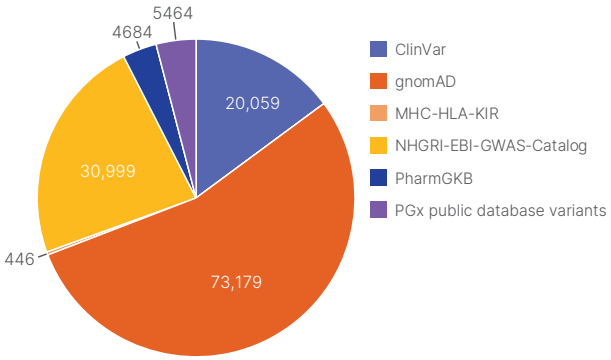

Figure 5: Clinical research content—Expertly selected clinical research content from key databases supports a broad range of applications.

QC markers

The Infinium Global Screening Array-48 v4.0 includes ~8.8K quality control (QC) markers. QC markers on the BeadChip are selected to facilitate high-throughput studies and enable sample tracking functions, including ancestry determination, and stratification (Figure 6).

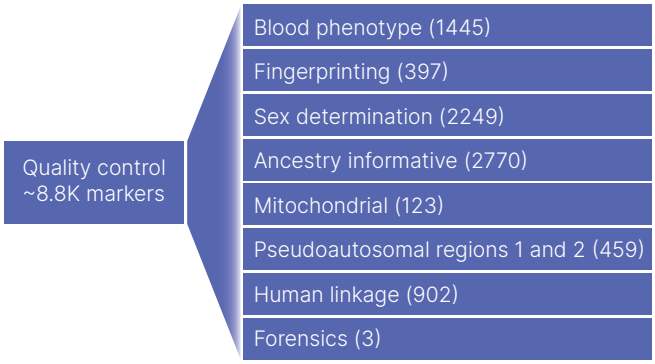

Figure 6: QC content by category—The BeadChip contains ~8.8K markers enabling various sample tracking functions such as sex determination, continental ancestry, human linkage, and more.

High-performance assay

The Infinium Global Screening Array-48 v4.0 BeadChip uses trusted Infinium assay chemistry to deliver a high-performance, accurate genotyping solution (Table 5).

Table 5: Data performance and spacing

Data performance	Observed	Product specification[b]
Call rate	99.6%	> 99.0% avg
Reproducibility	99.99%	> 99.90%
Log R deviation	0.11[c]	< 0.30 avg[d]
	Mean	Median	90th percentile[c]
Probe spacing	4.6 kb	2.4 kb	11.0

a. Excludes Y chromosome markers for female samples.
b. Based on results from GenTrain sample set.
c. Value expected for typical projects using standard Illumina protocols.

Infinium EX chemistry

Infinium EX chemistry is optimized for automation and compatible with Infinium Amplification and Infinium Automated Pipetting Systems. The automated workflow substantially reduces hands-on time and the potential for human error when using the Infinium Global Screening Array-48 v4.0 BeadChip. The Infinium EX workflow provides results in as little as two days (Figure 7).

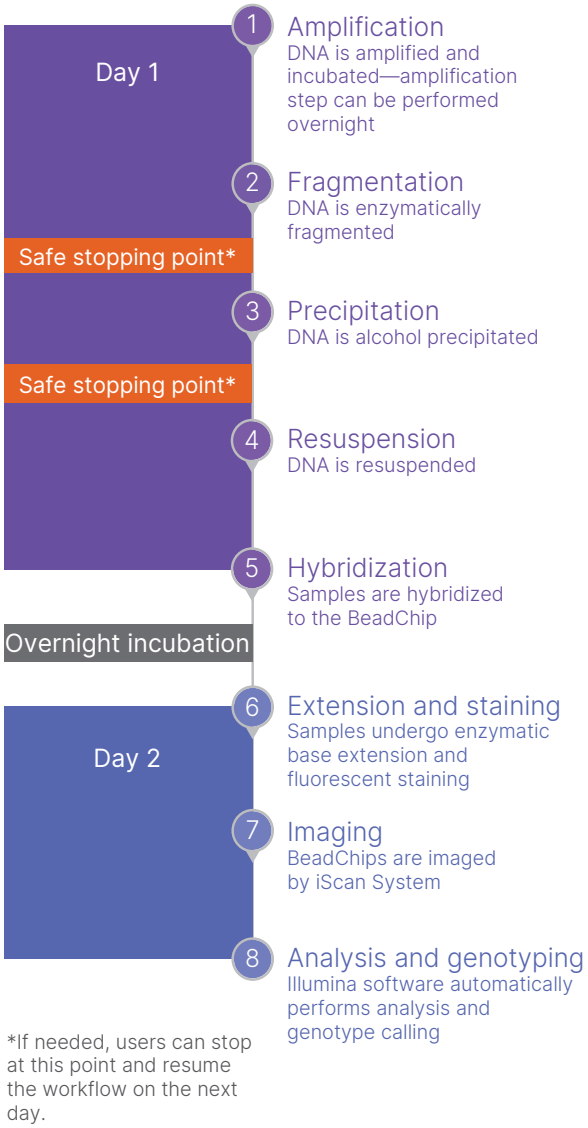

Figure 7: The Infinium EX 48-sample workflow—The Infinium EX 48-sample workflow provides an optional two or three-day workflow with minimal hands-on time.

Powerful, accurate secondary analysis software options

Illumina offers multiple solutions for array analysis to meet the needs of large and small labs. DRAGEN™ Array analysis software provides accurate, comprehensive, and efficient secondary analysis of Infinium microarray data. It is available in two environments. A local command-line package with an intuitive interface makes it easy for power users to have granular control and flexibility to support large-scale microarray genomic studies. DRAGEN Array is also available as a cloud offering via the user-friendly BaseSpace™ Sequence Hub interface. Both versions offer efficient genotyping performance and high data accuracy. Results can be outputted in VCF format for downstream analysis compatibility.

GenomeStudio™ Software provides an integrated platform for data analysis of all Illumina array-based genotyping assays. The graphical display of results generated from data analysis enables a thorough understanding of the large data sets generated by Illumina genotyping assays by providing multimodal examination from a broad, genome-wide view down to a fine-grained single-base view. GenomeStudio Software displays results at all scales to enable researchers to examine high-resolution genome-wide data effectively.

Summary

The Infinium Global Screening Array-48 v4.0 is a high-density genotyping assay that is ideal for medium- to high-throughput applications. The ~650K markers are carefully selected to support a range of applications, including population and clinical research studies. Infinium EX chemistry delivers high accuracy, detection of challenging targets, and fast turnaround times. When combined with the Infinium Automated Pipetting System with ILASS and the Infinium Amplification System the BeadChip offers a high-throughput option for labs looking to process large numbers of samples, with limited hands-on processing.

Learn more

The Infinium Global Screening Array-48 v4.0

DRAGEN Array software [Link pending]

GenomeStudio Software

Ordering information

Infinium Global Screening Array-48 v4.0 kit	Catalog no.
48 samples	20065217
96 samples	20068344
1152 samples	20068345
Infinium Global Screening Array-48+ v4.0 kit	**Catalog no.[a]**
48 samples	20068360
96 samples	20068361
1152 samples	20068362

a. Catalog no. 20068360, 20068361, and 20068362 support the addition of custom marker content.

References

1. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014;42(Database issue):D980-D985. doi:10.1093/nar/gkt1113
2. National Human Genome Research Institute. genome.gov. Accessed April 25, 2023.
3. 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
4. Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res*. 2017;45(D1):D840-D845. doi:10.1093/nar/gkw971
5. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing [published correction appears in Genet Med. 2017 May;19(5):606]. *Genet Med*. 2013;15(7):565-574. doi:10.1038/gim.2013.73
6. PharmaADME Gene List. pharmaadme.org. Accessed Decenber 15, 2020.
7. 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
8. Patnaik SK, Helmberg W, Blumenfeld OO. BGMUT Database of Allelic Variants of Genes Encoding Human Blood Group Antigens. *Transfus Med Hemother*. 2014;41(5):346-351. doi:10.1159/000366108.
9. 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
10. Clinical Pharmacogenetics Implementation Consortium (CPIC). cpicpgx.org. Accessed April 25, 2023.
11. 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
12. 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
13. Wong KM, Langlais K, Tobias GS, et al. The dbGaP data browser: a new tool for browsing dbGaP controlled-access genomic data. *Nucleic Acids Res*. 2017;45(D1):D819-D826. doi:10.1093/nar/gkw1139
14. Osier MV, Cheung KH, Kidd JR, Pakstis AJ, Miller PL, Kidd KK. ALFRED: an allele frequency database for diverse populations and DNA polymorphisms--an update. *Nucleic Acids Res*. 2001;29(1):317-319. doi:10.1093/nar/29.1.317
15. gnomAD, Genome Aggregation Database. gnomad.broadinstitute.org. Accessed April 25, 2023.
16. 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
17. 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
18. 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 April 25, 2023.
19. 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
20. NCBI. RefSeq - NCBI Reference Sequence Database. ncbi.nlm.nih.gov/refseq. Accessed April 25, 2023.
21. NCBI. Genome assembly GRCh37. ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.13/. Accessed April 25, 2023

illumina®

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

M-GL-00712 v1.0