illumina®

Infinium® Global Screening Array-24 v1.0

A powerful, high-quality, economical array for population-scale genetic studies.

Highlights

- **Global Content**
  Includes a multiethnic genome-wide backbone, expertly designed clinical research variants, quality control (QC) markers, and the option to add content
- **Broad Clinical Research Applications**
  Enables genotyping for a broad range of applications, including complex disease studies, pharmacogenomics research, lifestyle and wellness characterization, and more
- **High-Throughput Workflow**
  Supports high-throughput processing of thousands of samples per week for population-scale studies
- **Robust, High-Quality Assay**
  Maintains the same data quality of Illumina genotyping arrays with call rates > 99% and reproducibility > 99.9%

Introduction

The Infinium Global Screening Array-24 v1.0 (GSA) BeadChip is an advanced genotyping array that provides an economical solution for population-scale genetic studies, variant screening, and precision medicine research. Using the proven iScan® System, integrated analysis software, and Infinium high-throughput screening (HTS) Assay, this high-density, 24-sample BeadChip (Figure 1) provides optimized content for a broad range of applications, delivered with the same high-quality, reproducible data that Illumina genotyping arrays have provided for over a decade (Table 1). The GSA Kit includes convenient packaging containing BeadChips and reagents for amplifying, fragmenting, hybridizing, labeling, and detecting genetic variants using the high-throughput, streamlined Infinium workflow.

Widespread Adoption

The BeadChip builds on the success of the consortium version of the Infinium Global Screening Array that was developed by a community of human disease researchers, health care networks, consumer genomics companies, and genomic service providers. The consortium version has been widely adopted with over 5.5 million BeadChips ordered by a global community that provides a network of users that can help power discovery through collaboration and data sharing.

Optimized Global Content

The BeadChip combines highly optimized multiethnic genome-wide content, curated clinical research variants, and QC markers for a broad range of clinical research and variant screening applications. These applications include disease association and risk profiling studies, pharmacogenomics research, disease characterization, lifestyle and wellness characterization, and marker discovery in complex disease research (Figure 2).

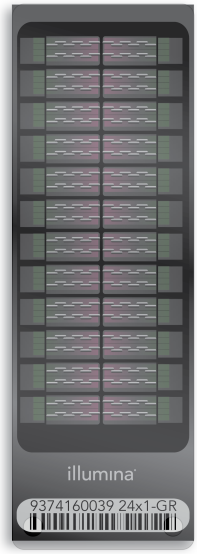

Figure 1: The BeadChip—The BeadChip is built on the trusted 24-sample Infinium HTS platform.

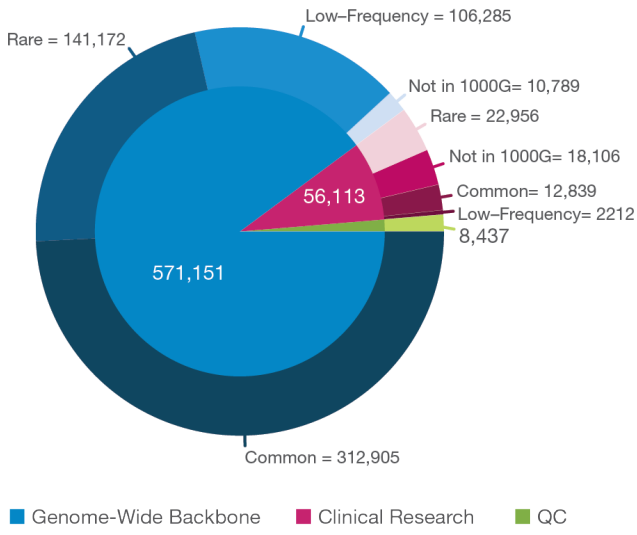

Figure 2: Summary of Content on the BeadChip–Genome-wide content enables a broad range of clinical research and genetic variant screening applications. Plotted in the inner pie is the proportion of the array that was selected for genome-wide coverage (blue), clinical research (pink), and quality control (green). The outer ring summarizes the weighted reference global allele frequency for variants present in the 1000 Genomes Project (1000G).[1] Variants not in 1000G are labeled.

Table 1: Product Information

Feature	Description	
Species	Human	
Total Number of Markers	642,824	
Capacity for Custom Bead Types	50,000	
Number of Samples per BeadChip	24	
DNA Input Requirement	200 ng	
Assay Chemistry	Infinium HTS	
Instrument Support	iScan or HiScan® System	
Sample Throughput[a]	~ 2304 samples/week	
Scan Time per Sample	iScan System 2.5 min	HiScan System 2.0 min

a. Estimate assumes 1 iScan System, 1 AutoLoader, 2 Tecan robots, and a 5-day work week.

Table 2: High-Value Content

Content	No. of Markers	Research Application/Note
ADME Core and Extended Genes	5816	Drug metabolism and excretion
ADME Core and Extended Genes +/- 10 kb	7246	Drug metabolism and excretion (plus regulatory regions)
APOE	17	Cardiovascular disease, Alzheimer's disease, immunoregulation, and cognition
Blood Phenotype Genes	1984	Blood phenotypes
COSMIC Genes	276,149	Somatic mutations in cancer
GO CVS Genes	82,984	Cardiovascular conditions
Database of Genomic Variants	494,268	Genomic structural variation
eQTLs	2680	Genomic loci regulating mRNA expression levels
Fingerprint SNPs	385	Human identification
HLA Genes	439	Disease defense, transplant rejection, and autoimmune disorders
Extended MHC[a]	8608	Disease defense, transplant rejection, and autoimmune disorders
KIR Genes	27	Autoimmune disorders and disease defense
Neanderthal SNPs	765	Neanderthal ancestry and human population migration
NHGRI GWAS Catalog	6988	Markers from published genome-wide association studies
RefSeq 3' UTRs	10,808	3' untranslated regions of known genes
RefSeq 5' UTRs	5268	5' untranslated regions of known genes
RefSeq All UTRs	15,614	All untranslated regions of known genes
RefSeq	310,926	All known genes
RefSeq +/- 10 kb	367,210	All known genes plus regulatory regions
RefSeq Promoters	13,567	2 kb upstream of all known genes to include promoter regions
RefSeq Splice Regions	1714	Variants at splice sites in all known genes

a. Extended MHC is a ~ 8 Mb region.

Abbreviations: ADME: absorption, distribution, metabolism, and excretion; APOE: apolipoprotein E; COSMIC: catalog of somatic mutations in cancer; GO CVS: gene ontology annotation of the cardiovascular system; eQTL: expression quantitative trait loci; HLA: human leukocyte antigen; KIR: killer cell immunoglobulin-like receptor; MHC: major histocompatibility complex; NHGRI: national human genome research institute; GWAS: genome-wide association study; UTR: untranslated region; RefSeq: reference sequence.

Broad Clinical Research Applications

The clinical research content of the BeadChip was designed through collaboration with medical genomics experts using multiple annotation databases[2-5] to create an informative, economical panel for clinical research applications (Tables 2 and 3).

Expertly Selected Content

Variants included on the array consist of known disease association markers based on ClinVar,[2] the Pharmacogenomics Knowledgebase (PharmGKB),[3] and the National Human Genome Research Institute (NHGRI) database[4] (Figure 3). In addition to disease associated markers, the GSA contains imputation-based tagSNPs for HLA alleles and putative functional content from the Exome Aggregation Consortium (ExAC) database.[5]

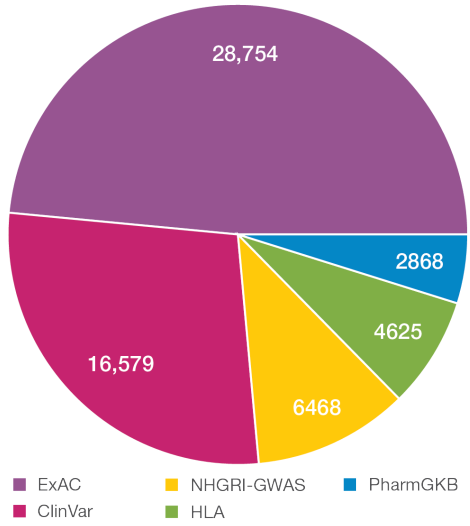

Figure 3: Clinical Research Content on the BeadChip–Clinical research content was expertly selected from scientifically recognized databases to create a highly informative array for clinical research applications.

Table 3: Marker Information

Marker Categories			No. of Markers
Exonic Markers[a]			66,199
Intronic Markers[a]			256,673
Nonsense Markers[b]			3232
Missense Markers[b]			43,342
Synonymous Markers[b]			5109
Mitochondrial Markers[c]			137
Indels[c]			3836
Sex Chromosomes[c]	X 16,927	Y 1456	PAR/Homologous 576

a. RefSeq - NCBI Reference Sequence Database[6]
b. Compared against the University of California, Santa Cruz (UCSC) Genome Browser[7]
c. NCBI Genome Reference Consortium, Version GRCh37[8]

Abbreviations: PAR: pseudoautosomal region

Broad Spectrum of Pharmacogenomics Markers and Exonic Content

The BeadChip features pharmacogenomics variants associated with absorption, distribution, metabolism, and excretion (ADME) phenotypes based on PharmGKB[3] and Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines[9] (Figure 4). It also features diverse exonic content from the ExAC database,[5] including both cross population and population specific markers (Figure 5) with either functionality or strong evidence for association.

Figure 4: Broad Spectrum of Pharmacogenomics Markers–Clinical research content features an extensive list of pharmacogenomics markers selected based on CPIC guidelines and the PharmGKB database.[10] Markers are arranged according to level of evidence as defined by the PharmGKB database. VIP: very important pharmacogene.

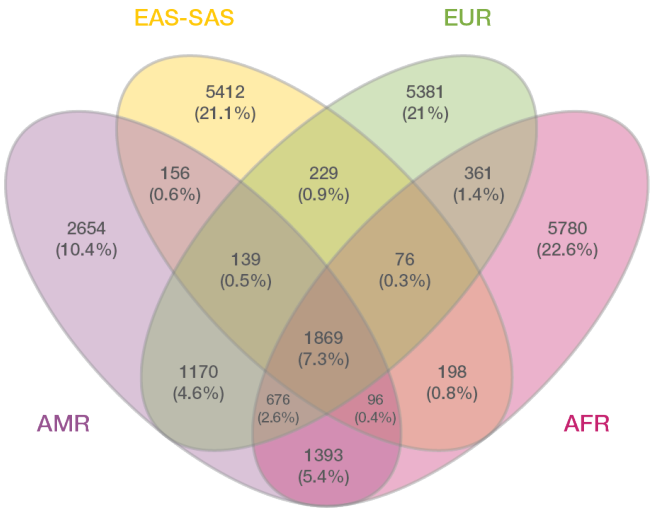

Figure 5: Global Exonic Content is Cross-Population and Population-Specific–Exonic content included on the BeadChip contains content that is present across several populations as well as population-specific content. The Venn diagram displays proportion of total content that either overlaps or is specific to certain populations. Abbreviations: EAS: East Asian; SAS: South Asian; AMR: Ad Mixed American; AFR: African; EUR: European.

Extensive Range of Disease Categories Covered

Including over 18,000 variants with established clinical associations based on the ClinVar database,[2] clinical research content on the BeadChip enables validation of disease associations, risk profiling, preemptive screening research, and pharmacogenomics studies. Variant selection includes a range of pathology classifications based on the ClinVar American College of Medical Genetics and Genomics (ACMG) annotations (Figure 6A).[11] There are over 7000 disease and trait associations from the ClinVar database (Figure 6B) and over 7000 variants selected from the NHGRI-GWAS catalog[4] (Figure 7), representing a broad range of phenotypes and disease classifications.

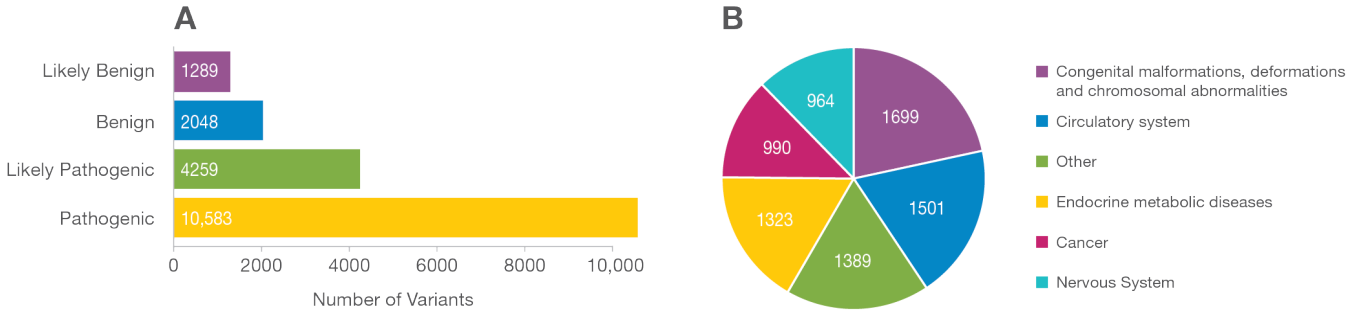

Figure 6: Broad Coverage of Disease Categories– (A) Variants sorted by range of pathology classifications according to ClinVar American College of Medical Genetics (ACMG) annotations. (B) BeadChip clinical research content features over 7000 markers based on the ClinVar database.

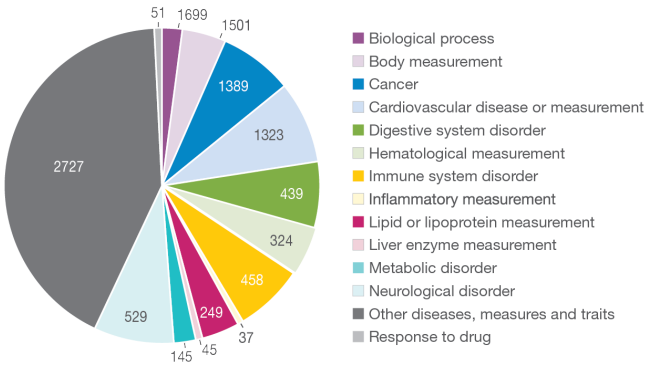

Figure 7: NHGRI Disease Categories– BeadChip clinical research content features over 7000 markers across 20 disease categories based on the NHGRI database.

QC Markers for Sample Identification, Tracking, and Stratification

The BeadChip includes QC and high-value markers for large-scale studies, enabling sample identification, tracking, ancestry determination, and stratification (Figure 8).

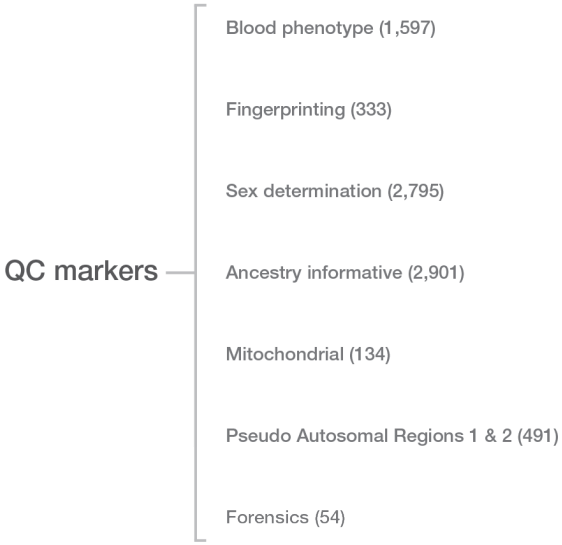

Figure 8: QC Markers–QC variants on the BeadChip enable a variety of capabilities for sample tracking such as sex determination, continental ancestry, and forensics.

Flexible Content Options

The BeadChip can be customized to incorporate up to 50,000 custom bead types or predesigned content panels (Table 4).

Table 4: Flexible Content Options

Optional Compatible Content	No. of Markers	Description
Custom Content	≤ 50,000 Bead Types	Custom design virtually any target (eg SNP, CNV, indel) using the DesignStudio™ Microarray Assay Designer[a]
Infinium Global Screening Array Multi-Ethnic Disease Drop-In Panel	~ 50,00 Markers	Fine-mapping content derived from exome sequencing and meta analysis of phenotype-specific consortia focused on the following traits: psychiatric, neurological, cancer, cardiometabolic, autoimmune, anthropometric
Infinium PsychArray-24 v1.0 Focused Content Panel	~ 30,00 Markers	Markers from the Infinium PsychArray-24 v1.1 BeadChip[b] associated with common psychiatric disorders including, schizophrenia, bipolar disorder, autism spectrum disorders, attention deficit hyperactivity disorder, major depressive disorders, obsessive compulsive disorder, anorexia, Tourette's syndrome

a. www.illumina.com/designstudio.html.
b. www.illumina.com/products/by-type/microarray-kits/infinium-psycharray.html.

Abbreviations: SNP: single nucleotide polymorphism; CNV: copy number variation; indel: insertion/deletion.

High-Throughput Workflow

The BeadChip uses the highly scalable 24-sample Infinium HTS format for high-throughput processing of thousands of samples per week for large, population-scale research and variant screening. The Infinium HTS format also provides a rapid 3-day workflow that allows genotyping service providers and clinical researchers to gather data and advance studies quickly (Figure 9).

Optional integration of the Illumina Laboratory Information Management System (LIMS) into the workflow provides high laboratory efficiency with automation functionality, process tracking, and QC data tracking. The Illumina ArrayLab Consulting Service offers customized solutions to high-throughput genotyping labs that desire increased efficiency and overall operational excellence.

Robust, High-Quality Assay

The BeadChip uses proven Infinium assay chemistry to deliver the same high-quality, reproducible data (Table 5) that Illumina genotyping arrays have provided for over a decade. The Infinium product line provides high call rates and high reproducibility for numerous sample types including, saliva, blood, solid tumors, fresh frozen, and buccal swabs. It is compatible with the Infinium FFPE QC and DNA Restoration Kits,[12] enabling genotyping of formalin-fixed, paraffin-embedded (FFPE) samples. In addition, the high signal-to-noise ratio of the individual genotyping calls from the Infinium assay provides researchers with access to genome-wide copy number variant (CNV) calling with a mean probe spacing of ~ 4.59 kb.

Table 5: Data Performance and Spacing

Data Performance	Value[a]	Product Specification	
Call Rate	99.9%	> 99% avg	
Reproducibility	99.9%	> 99.9%	
Log R Deviation	0.10	< 0.30[b]	
Spacing			
Spacing (kb)	Mean	Median	90th%[b]
	4.59	2.53	10.84

a. Values are derived from genotyping 308 HapMap reference samples.
b. Value expected for typical projects using standard Illumina protocols. Tumor samples and samples prepared by methods other than standard Illumina protocols are excluded.

Figure 9: The Infinium HTS Workflow–The Infinium HTS format provides rapid 3-day workflow with minimal hands-on time.

High Imputation Accuracy for Global Populations

Leveraging available whole-genome reference data from over 26 global populations in Phase 3 of the 1000 Genomes Project,[1] the genome-wide content on the BeadChiphas been selected to generate high imputation accuracy for low-frequency and common variants (minor allele frequencies (MAF) of > 1%) (Tables 6–10). High imputation accuracy provides increased power to support population-scale disease research and population-specific causal variant detection.

Table 6: Imputation Accuracy from 1000G[a] at Various MAF Thresholds

Population[b]	Imputation Accuracy		
	MAF ≥ 5%	MAF ≥ 1%	MAF 1–5%
AFR	0.91	0.86	0.79
AMR	0.95	0.92	0.85
EAS	0.94	0.89	0.77
EUR	0.95	0.93	0.87
SAS	0.94	0.89	0.78

a. Compared against Phase 3, version 5 of the 1000 Genomes Project (1000G). www.1000genomes.org. Accessed July 2016.
b. www.1000genomes.org/category/frequently-asked-questions/population

Abbreviations: MAF: minor allele frequency

Table 7: Number of Markers Imputed at r^2 ≥ 0.80 from 1000G[a]

Population[b]	Number of Markers Imputed at r^2 ≥ 0.80 (% of Total Markers)		
	MAF ≥ 5%	MAF ≥ 1%	MAF 1–5%
AFR	6.5M (76%)	11.1M (70%)	4.6M (63%)
AMR	5.6M (89%)	12.0M (90%)	6.4M (91%)
EAS	4.8M (86%)	8.6M (86%)	3.8M (85%)
EUR	5.5M (90%)	9.7M (89%)	4.2M (87%)
SAS	5.4M (87%)	9.6M (85%)	4.3M (82%)

a. Compared against Phase 3, version 5 of the 1000 Genomes Project (1000G). www.1000genomes.org. Accessed July 2016.
b. www.1000genomes.org/category/frequently-asked-questions/population

Table 8: No. of Markers LD r^2 ≥ 0.80 from 1000G[a] at Various MAF Thresholds

1000G Population[b]	LD Coverage (r^2 ≥ 0.80)		
	MAF ≥ 5%	MAF ≥ 1%	MAF 1–5%
AFR	1.8M (22%)	2.2M (14%)	279K (4%)
AMR	2.9M (47%)	3.7M (38%)	750K (21%)
EAS	3.2M (59%)	4.0M (53%)	818K (39%)
EUR	3.1M (52%)	4.3M (50%)	1.3M (47%)
SAS	3.1M (51%)	3.8M (43%)	660K (24%)

a. Compared against Phase 3, version 5 of the 1000 Genomes Project (1000G). www.1000genomes.org. Accessed July 2016.
b. www.1000genomes.org/category/frequently-asked-questions/population

Abbreviations: LD: linkage disequilibrium

Table 9: LD Mean r^2 from 1000G[a] at Various MAF Thresholds

Population[b]	LD Coverage (Mean r^2)		
	MAF ≥ 5%	MAF ≥ 1%	MAF 1–5%
AFR	0.44	0.30	0.11
AMR	0.69	0.58	0.35
EAS	0.75	0.68	0.49
EUR	0.71	0.68	0.58
SAS	0.71	0.61	0.35

a. Compared against Phase 3, version 5 of the 1000 Genomes Project (1000G). www.1000genomes.org. Accessed July 2016.
b. See www.1000genomes.org/category/frequently-asked-questions/population

Summary

The BeadChip provides an economical solution for population-scale genetic studies, variant screening, and precision medicine research. The BeadChip builds on the success of the consortium version of the Infinium Global Screening Array, which has been widely adopted with over 5.5 million BeadChips ordered worldwide. Using the proven iScan System, Infinium HTS Assay, and integrated analysis software, this high-density, 24-sample BeadChip provides optimized content for a broad range of clinical research applications.

Ordering Information

Infinium Global Screening Array-24 v1.0 Kit	Catalog No.
48 Samples	20005132
288 Samples	20005133
1152 Samples	20005134
Infinium Global Screening Array-24 v1.0 Kit*	**Catalog No.**
48 Samples	20005135
288 Samples	20005136
1152 Samples	20005137

*Enabled for additional custom content.

Learn More

To learn more about the BeadChip and other Illumina genotyping products and services, visit www.illumina.com/genotyping.html

References

1. The 1000 Genomes Project. www.1000genomes.org. Accessed July 16, 2016.
2. ClinVar Database. www.ncbi.nlm.nih.gov/clinvar. Accessed October 2016.
3. PharmGKB, The Pharmacogenomics Knowledgebase. www.pharmgkb.org. Accessed January 2017.
4. National Human Genome Research Institute. www.genome.gov/. Accessed January 2017.
5. Exome Aggregation Consortium (ExAC) Browser. exac.broadinstitute.org. Accessed October 2016.
6. RefSeq - NCBI Reference Sequence Database. www.ncbi.nlm.nih.gov/refseq. Accessed September 2016.
7. University of California, Santa Cruz (UCSC) Genome Browser. genome.ucsc.edu. Accessed July 2016.

8. NCBI Genome Reference Consortium. Version GRCh37. www.ncbi.nlm.nih.gov/grc/human. Accessed July 2016.
9. Clinical Pharmacogenetics Implementation Consortium (CPIC). cpicpgx.org. Accessed October 2016
10. PharmGKB, Clinical Annotation Levels of Evidence. www.pharmgkb.org/page/clinAnnLevels. Accessed January 2017.
11. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. www.ncbi.nlm.nih.gov/clinvar/docs/acmg/. Accessed January 2017.
12. Infinium FFPE QC and DNA Restoration Kit. www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_FFPE_DNA_restoration.pdf.

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

For Research Use Only. Not for use in diagnostic procedures.

illumina®