DATA SHEET

Infinium Asian Screening Array-24 v1.0 BeadChip

Enable population-scale genetic studies in East and Southeast Asia

- Includes optimized, population-specific content designed in collaboration with a consortium of human disease researchers and genomic service providers
- Enables broad clinical research applications including complex disease studies, pharmacogenomics research, and more
- Maintains the same data quality of Illumina genotyping arrays with call rates > 99% and reproducibility > 99.9%

illumina®

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

Introduction

The Infinium Asian Screening Array-24 v1.0 BeadChip is an advanced genotyping array that provides a high-value, scalable, and cost-effective solution for population-scale genetic studies, variant screening, and precision medicine research (Table 1). Using the 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. The Infinium Asian Screening Array-24 v1.0 Kit includes convenient packaging containing BeadChips and reagents for amplifying, fragmenting, hybridizing, labeling, and detecting genetic variants using the high-throughput, streamlined Infinium workflow.

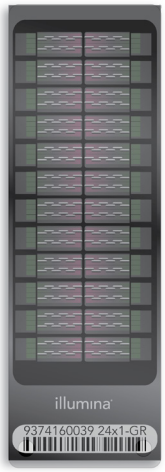

Figure 1: The Infinium Asian Screening Array-24 v1.0 BeadChip— Built on the trusted 24-sample Infinium HTS platform

Widspread adoption

The Infinium Asian Screening Array-24 v1.0 BeadChip builds on the success of the commercial version of the Infinium Global Screening Array v2.0 that has been widely adopted by a community of human disease researchers, health care networks, consumer genomics companies, and genomic service providers. Over 20 million samples of Infinium screening arrays have been ordered by a global community of users powering discovery through collaboration and data sharing.

Table 1: Product information

Read length	Description
Species	Human
Total number of markers	659,184
Capacity for custom bead types	50,000
Number of samples per BeadChip	24
DNA input requirement	200 ng
Assay chemistry	Infinum HTS
Instrument support	iScan System
Sample throughput	~ 2304 samples/week
Scan time per sample[a]	2.5 min

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

Optimized East and Southeast Asian conent from whole-genome samples

The Asian Screening Array consortium contributed whole-genome sequencing (WGS) data for variant selection to improve coverage among underrepresented samples relative to current reference panels. WGS data used to build the the genome-wide backbone, exome content, and sub-continental ancestry informative markers (AIMS) within East and Southeast Asian populations, including Korean, Mongolian and Malaysian samples currently not represented in publicly available reference populations (Table 2).

The Infinium Asian Screening Array-24 v1.0 BeadChip combines highly optimized multiethnic genome-wide content, curated clinical research variants, an East and Southeast Asian discovery panel, and quality control (QC) markers. The array enables a broad range of applications, including disease association and risk profiling studies, pharmacogenomics research, disease characterization, lifestyle and wellness characterization, and marker discovery in complex disease research.

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

Table 2: Whole-genome sequences used in Asian Screening Array development

Population	Cohort	Sequence depth
Chinese	Genomics Institute of Singapore	30×
Southern Han Chinese	1KGP[a] Southern Han Chinese	2-4×
Chinese Dai	1KGP Chinese Dai in Xishangbanna	2-4×
Han Chinese	1KGP Bejing	2-4×
Japanese	Riken W13	15×
Japanese	Riken W1-6, W18	30×
Japanese	Kyoto	3-15×
Japanese	1KGP Tokyo	2-4×
Korean	Macrogen	15-30×
Korean	KPGP[b]	35×
Malaysian	Singapore Population Health Study	30×
Mongolian	Macrogen	30×
Vietnamese	1KGP Kinh Ho Chin Minh	2-4×

a. 1000 Genomes Project (1KGP). www.1000genomes.org
b. The Personal Genome Project (KPGP)

The Infinium Asian Screening Array-24 v1.0 BeadChip achieves exceptional genomic coverage and imputation performance in all five major sub-populations within Asia. The array addresses gaps in the publicly available reference populations within Asia from South Korea, China, Japan, Mongolia, and Malaysia. Polymorphic content on the array enables discovery applications.

Broad clinical research applications

Clinical research content on the Infinium Asian Screening Array-24 v1.0 BeadChip was designed through collaboration with medical genomics experts using multiple annotation databases1–4 to create an informative, cost-effective panel for clinical research applications (Figure 2 and Table 3).

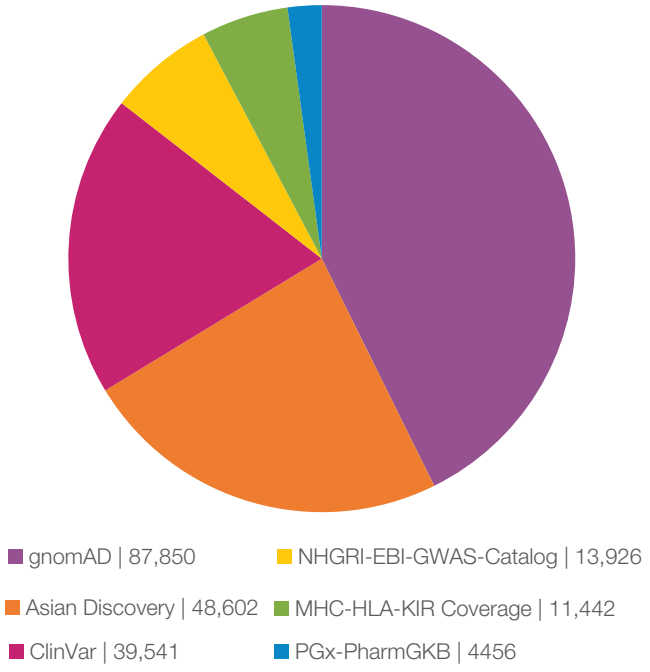

Figure 2: Clinical research content—Content was expertly selected from scientifically recognized databases to create a highly informative array for clinilcal research applications. Variant counts may be subject to change.

Expertly selected content

Variants included on the array consist of markers with known disease association based on ClinVar,[1] the Pharmacogenomics Knowledgebase (PharmGKB),[2] and the National Human Genome Research Institute (NHGRI)-EBI database[3] (Table 3 and Table 4). In addition to disease-associated markers, the array contains imputation-based tagSNPs for HLA alleles, extended MHC region, the KIR gene, and exonic content from the gnomAD database.[4] Exclusive to the Infinium Asian Screening Array-24 v1.0 BeadChip, the clinical research content also contains ~50K unpublished markers from various consortia members in Asia.

Table 3: High-value content

Content	No. of markers	Research application/note	Content	No. of markers	Research application/ note
ACMG[5] 2016 gene list	16,363	Genes from genome and exome sequencing by ACMG	gnomAD exome	87,850	Consists of exome and whole-genome sequences from unrelated individuals
ADME core and extended + CPIC genes[6]	15,342	Drug metabolism and excretion	HLA genes[7]	803	Disease defense, transplant rejection, and autoimmune disorders
ADME core and extended + CPIC genes[6] ± 10 kb	18,291	Drug metabolism and excretion (plus regulatory regions)	Extended MHC[14c]	10,579	Disease defense, transplant rejection, and autoimmune disorders
AIMs[b]	2824	Ancestry-informative markers	KIR genes[7]	60	Autoimmune disorders and disease defense
APOE[7]	14	Cardiovascular disease, Alzheimer's disease, immunoregulation, and cognition	Neanderthal SNPs[16]	1515	Neanderthal ancestry and human population migration
Blood phenotype genes[8]	1859	Blood phenotypes	Newborn/carrier screening gene coverage[15]	25,291	Genes included in the TruSight™ Inherited Disease Sequencing Panel
ClinVar[1] variants	39,541	Relationships among variation, phenotypes, and human health	NHGRI-EBI GWAS catalog[3]	13,926	Markers from published GWAS
ClinVar[1] pathogenic	13,005	Cause Mendelian disorders	PharmGKB[2]	4456	Human genetic variation associated with drug responses
ClinVar[1] likely pathogenic	5317	Likely cause Mendelian disorders	RefSeq[17] 3′ UTRs	14,915	3′ untranslated regions of known genes
ClinVar[1] benign	3747	Do not cause Mendelian disorders	RefSeq[17] 5′ UTRs	7644	5′ untranslated regions of known genes
ClinVar[1] likely benign	4308	Likely do not cause Mendelian disorders	RefSeq[17] All UTRs	21,926	All untranslated regions of known genes
COSMIC[9] genes	317,039	Somatic mutations in cancer	RefSeq[17]	351,194	All known genes
GO[10] CVS genes	104,838	Cardiovascular conditions	RefSeq[17] ± 10 kb	406,671	All known genes plus regulatory regions
Database of Genomic Variants[11]	516,905	Genomic structural variation	RefSeq[17] Promoters	15,748	2 kb upstream of all known genes to include promoter regions
eQTLs[12]	2383	Genomic loci regulating mRNA expression levels	RefSeq[17] Splice Regions	3006	Variants at splice sites in all known genes
Fingerprint SNPs[13]	428	Human identification			

a. The number of markers for each content category may be subject to change
b. Based on internal calculations
c. Extended MHC is a ~ 8 Mb region

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.

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

Table 4: Marker information

Markers categories			No. of markers
Exonic markers[a]			110,588
Nonsense markers[b]			6814
Missense markers[b]			74,490
Synonymous markers[b]			9145
Mitochondrial markers[c]			1042
Indels[c]			9012
Sex chromosomes[c]	X 26,622	Y 4771	PAR/homologous 676

a. RefSeq - NCBI Reference Sequence Database.[17] Accessed November 20, 2020.
b. Compared against the University of California, Santa Cruz (UCSC) Genome Browser.[7] Accessed November 20, 2020.
c. NCBI Genome Reference Consortium, Version GRCh37.[18] Accessed November 20, 2020.
Abbreviations: indel: insertion/deletion; PAR: pseudoautosomal region

Broad spectrum of pharmacogenomics markers and exonic content

The Infinium Asian Screening Array-24 v1.0 BeadChip features pharmacogenomics variants associated with absorption, distribution, metabolism, and excretion (ADME) phenotypes based on PharmGKB[2] and Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines[19] (Figure 3). It also features diverse exonic content from the ExAC database,[4] including both cross population and population specific markers (Figure 4) with either functionality or strong evidence for association.

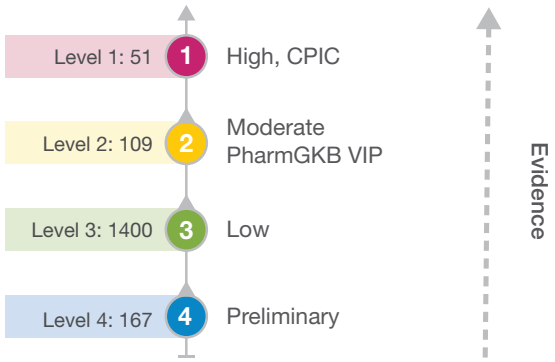

Figure 3: Broad spectrum of pharmacogenomics markers—Selected based on CPIC guidelines and the PharmGKB database. Markers are arranged according to level of evidence as defined by the PharmGKB database. VIP: very important pharmacogene. Variant counts may be subject to change.

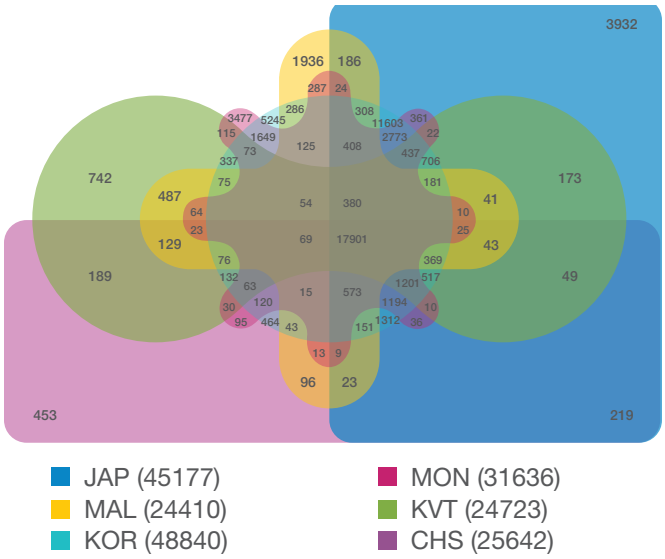

Figure 4: Exonic content across WGS samples—Exonic content includes markers that are common across several populations as well as population-specific content. The Venn diagram displays the proportion of total content that either overlaps or is specific to certain populations. Abbreviations: CHS= Han Chinese, JAP=Japanese, KOR=Korean, KVT=Vietnamese, MAL=Malay, MON=Mongolian. Variant counts may be subject to change.

Extensive range of disease categories covered

Including over 18K variants with established clinical associations based on the ClinVar database,[1] clinical research content on the Infinium Asian Screening Array-24 v1.0 BeadChip enables validation of disease associations, risk profiling, preemptive screening research, and pharmacogenomics studies. Variant selection includes a range of pathology classifications based on ClinVar American College of Medical Genetics and Genomics (ACMG) annotations (Figure 5A).[5] There are over 7K disease and trait associations from the ClinVar database (Figure 5B) and over 12K variants selected from the NHGRI-GWAS catalog[3] (Figure 6), representing a broad range of phenotypes and disease classifications.

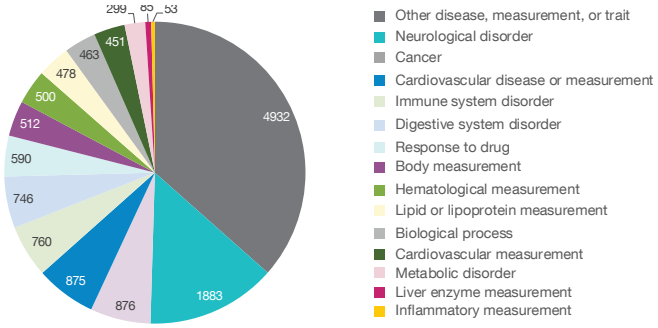

Figure 6: NHGRIEBI disease categories—Features over 12K markers across 15 disease categories based on the NHGRI-EBI database. Variant counts may be subject to change.

QC markers

The Infinium Asian Screening Array-24 v1.0 BeadChip includes QC and high-value markers for large-scale studies, enabling sample identification, tracking, ancestry determination, and stratification (Figure 7).

Figure 7: QC markers—QC variants on the array enable a variety of capabilities for sample tracking such as sex determination, continental ancestry, and human identifcation.

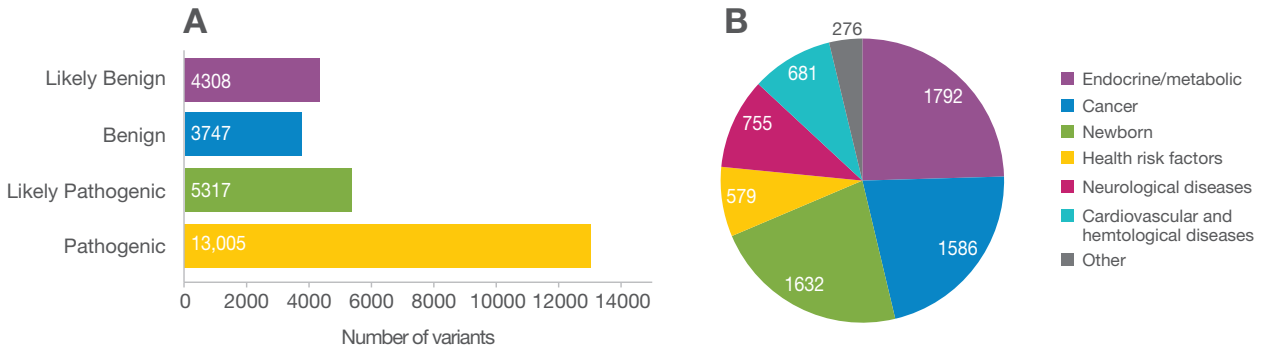

Figure 5: Broad coverage of disease categories—(A) Variants sorted by range of pathology classifications according to ClinVar ACMG annotations. (B) Clinical research content features over 7K disease traits and associations based on the ClinVar database. Variant counts may be subject to change.

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

Powering genomics in East and Southeast Asia

Asian discovery panel

More than 45K markers have been designed to develop the Asian discovery panel, which provides an opportunity for discovery in two ways. First, Asian-specific content seen in exomes from GnomAD (Figure 4) are captured in combination with a hypothesis-free approach using bioinformatics to determine evidence for functionality in Asian Screening Array WGS samples. Second, variants are selected with the Illumina imputation selection pipeline for coverage from previous expert-selected arrays, which includes the Infinium DrugDev Consotrium Array, the Infinium PsychArray, the Infinium ImmunoArray, and the Infinium NeuroX Consortium Array (Figure 8).

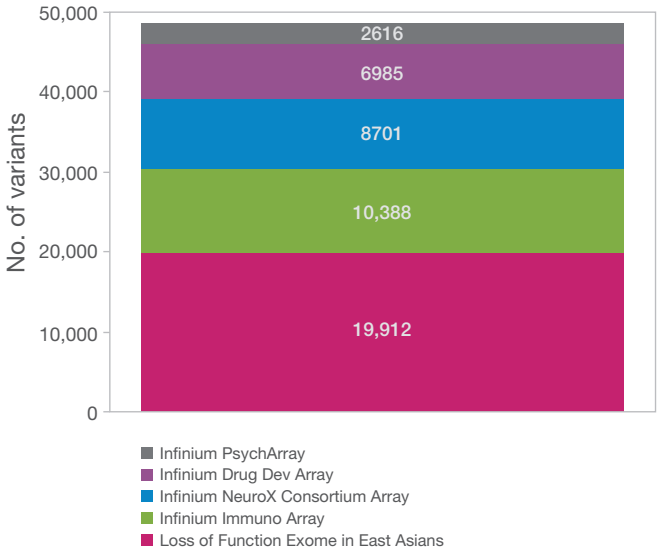

Figure 7: Asian discovery panel—Includes over 45K variants selected by tag SNPs for imputation coverage of targeted arrays (top 4 bars) and putative loss of function (LoF) variants from East and Southeast Asian exomes.

Subcontinental ancestry and HLA coverage

The Infinium Asian Screening Array-24 v1.0 BeadChip uses WGS samples to develop a unique subset of AIMs that can differentiate sub-population structure between Korean, Chinese, Japanese, Malaysian, and Mongolian, including additional AIMs that can capture HAN Chinese minority groups, differentiating between North and South China.

The Infinium Asian Screening Array-24 v1.0 BeadChip contains optimized HLA content aimed to support immune system research. To do this, the array captures > 11K MHC imputation tags based on an iterative imputation process in 1K Japanese and 10K Han Chinese HLA reference samples. These tag SNPs yield > 80% imputation accuracy for 4-digit and 2-digit HLA alleles.

Flexible content options

The Infinium Asian Screening Array-24 v1.0 BeadChip can be customized to incorporate up to 50K custom bead types or a predesigned content panel (Table 5).

Table 5: Flexible content options

Optional compatible content	No. of markers	Description
Custom content	≤ 50,000	Custom design virtually any target (eg, SNP, CNV, indel) using the DesignStudio™ Microarray Assay Designer[a]
Multi-disease drop-in panel	~ 50,000	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

a. www.illumina.com/designstudio.html
Abbreviations: SNP: single nucleotide polymorphism; CNV: copy number variation; indel: insertion/deletion

High-throughput workflow

The Infinium Asian Screening Array-24 v1.0 BeadChip uses the highly scalable 24-sample Infinium HTS format, which enables laboratories to efficiently increase throughput to support population-scale research and variant screening applications. For flexible throughput processing, the Infinium HTS assay provides the capability to run hundreds

to thousands of samples per week. The Infinium HTS assay provides a rapid, three-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 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.

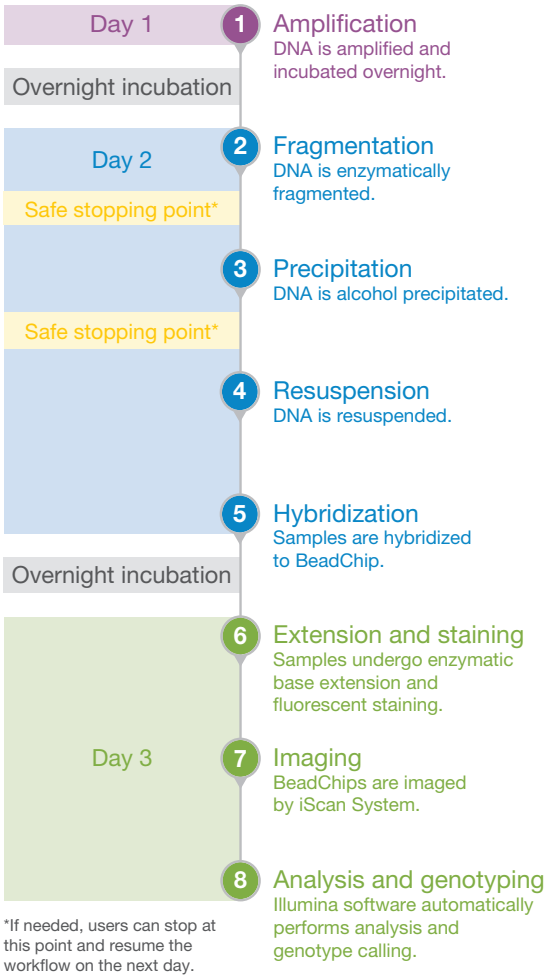

Figure 9: Infinium HTS workflow—Provides a rapid three-day workflow with minimal hands-on time.

Robust, high-quality assay

The Infinium Asian Screening Array-24 v1.0 BeadChip uses trusted Infinium assay chemistry to deliver the same high-quality, reproducible data (Table 6) 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,[21] 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 ~ 4.4 kb.

Table 6: Data performance and spacing

Data performance	Value[a]	Product specification[b]	
Call rate	99.5%	> 99.0% avg	
Reproducibility	99.99%	> 99.90%	
Log R deviation	0.11[c]	< 0.30 avg[d]	
Spacing			
Spacing (kb)	Mean 4.4	Median 2.0	90th%[c] 11.1

a. Values are derived from genotyping 1248 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. Tumor samples and samples prepared by methods other than standard Illumina protocols are excluded.

High coverage and imputation accuracy for East and Southeast Asian populations

Leveraging available whole-genome reference data, the genome-wide content on the Infinium Asian Screening Array-24 v1.0 BeadChip has been selected to generate high imputation accuracy for populations with limited representation in the 1000 Genomes Project (1KGP) (Table 7).

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

Thousands of samples were derived from various regions of Asia, including South Korea, China, Taiwan, Singapore, Japan, Mongolia, Malaysia, and Vietnam, several of which are population underrepresented or missing from the 1KGP.

High imputation accuracy provides increased power to support population-scale disease research and population-specific causal variant detection. Variant selection was largely focused on those of low minor allele frequencies (MAF = 1-5%), maximizing representation and imputation accuracy of recent population sub-structures by leveraging population-specific targets. High coverage of low MAF variants and high levels of polymorphic markers are key features of the Infinium Asian Screening Array-24 v1.0 BeadChip and a differentiator compared to previously developed arrays focused on common variants that lack coverage of Asian populations (Table 8).

Table 7: Imputation accuracy[a]

Population	Imputation accuracy		
	MAF ≥ 5%	MAF 1-5%	MAF 0.5-1%
Japanese	0.84	0.80	0.76
Korean	0.79	0.86	0.88
Malaysian	0.80	0.81	0.92

a. At various MAF thresholds for select populations underrepresented or missing from the 1KGP. Underrepresented populations are defined by comparison of population-specific samples included in the 1KGP

Abbreviations: MAF, minor allele frequency

Table 8: Variant counts[a]

Population	Imputation accuracy		
	MAF ≥ 5%	MAF 1-5%	MAF 0.5-1%
Chinese	548,557	243,348	120,263
Japanese	536,638	235,034	122,888
Korean	557,837	258,361	124,097
Malaysian	544,462	222,845	135,838

a. At various MAF thresholds for select populations

Abbreviations: MAF, minor allele frequency

Summary

The Infinium Asian Screening Array-24 v1.0 BeadChip provides a cost-effective solution for population-scale genetic studies, variant screening, and precision medicine research. The Infinium ASA-24 v1.0 BeadChip builds on the success of the widely adopted commercial version of the Infinium Global Screening Array. Using the 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.

Learn more

Learn more about the Infinium Asian Screening Array-24 v1.0 BeadChip and other Illumina genotyping products and services at www.illumina.com/genotyping.html.

For labs interested in higher throughput processing with the Infinium Asian Screening Array, contact your local sales representative for more information about Infinium HTS Extra high throughput kit configurations.

Ordering information

Infinium Asian Screening Array-24 v1.0 BeadChip Kit	Catalog no.
48 samples	20016317
288 samples	20016318
1152 samples	20016319
Infinium Asian Screening Array-24+ v1.0 BeadChip Kit[a]	**Catalog no.**
48 samples	20016320
288 samples	20016321
1152 samples	20016322

b. Enabled for custom content

References

1. ClinVar Database. www.ncbi.nlm.nih.gov/clinvar. Accessed November 20, 2020.
2. PharmGKB, The Pharmacogenomics Knowledgebase. www.pharmgkb.org. Accessed April 24, 2023.
3. National Human Genome Research Institute. www.genome.gov/. Accessed April 24, 2023.
4. Exome Aggregation Consortium (ExAC) Browser. exac.broadinstitute.org. Accessed April 24, 2023.
5. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. www.ncbi.nlm.nih.gov/clinvar/docs/acmg/. Accessed April 24, 2023.
6. PharmaADME Gene List. www.pharmaadme.org. Accessed April 24, 2023.
7. University of California, Santa Cruz (UCSC) Genome Browser. genome.ucsc.edu. Accessed April 24, 2023.
8. NCBI Reference Sequence Blood Group Antigen Gene Mutation Database. www.ncbi.nlm.nih.gov/projects/gv/rbc/xslcgi.fcgi?c-md=bgmut/systems. Accessed April 24, 2023.
9. Catalog of somatic mutations in cancer. cancer.sanger.uk/cosmic. Accessed April 24, 2023.
10. Gene Ontology Consortium. www.geneontology.org. Accessed April 24, 2023.
11. Database of Genomic Variants. dgv.tcag.ca/dgv/app/home. Accessed April 24, 2023.
12. NCBI eQTL Database. www.ncbi.nlm.nih.gov/projects/gap,eqtl/index.cgi. Accessed April 24, 2023.
13. The Allele Frequency Database. alfred.med.yale.edu/alfred/snpSets.asp. Accessed April 24, 2023.
14. de Bakker PIW, McVean G, Sabeti PC, et al. A high-resolution HLA and SNP haplotype map for disease association studies
15. Illumina (2017). TruSight Inherited Disease Sequencing Panel Data Sheet. Accessed April 24, 2023.
16. Neanderthal Genome Browser. neandertal.ensemblgenomes.org/index.html. Accessed April 24, 2023.
17. RefSeq - NCBI Reference Sequence Database. www.ncbi.nlm.nih.gov/refseq. Accessed April 24, 2023.
18. NCBI Genome Reference Consortium. Version GRCh37. www.ncbi.nlm.nih.gov/grc/human. Accessed April 24, 2023.
19. Clinical Pharmacogenetics Implementation Consortium (CPIC). cpicpgx.org. Accessed April 24, 2023.
20. PharmGKB, Clinical Annotation Levels of Evidence. www.pharmgkb.org/page/clinAnnLevels. Accessed April 24, 2023.
21. Infinium FFPE QC and DNA Restoration Kit.www.illumina.com/content/dam/illumina-marketing/documents/products/data-sheets/datasheet_FFPE_DNA_ restoration.pdf. Accessed April 24, 2023.

illumina®

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

© 2023 Illumina, Inc. All rights reserved. All trademarks are the property of Illumina, Inc. or their respective owners. For specific trademark information, see www.illumina.com/company/legal.html.
Pub. no. M-GL-01548 v1.0.