illumina®

Infinium® Multi-Ethnic Global BeadChip

A cost-effective array for understanding complex disease in diverse human populations.

Introduction

The Infinium Multi-Ethnic Global BeadChip harnesses content from Phase 3 of the 1000 Genomes Project (1kGP)[1], Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA), Population Architecture using Genomics and Epidemiology (PAGE), T2D-Genes Consortium, OMIM, ClinVar, ACMG, carrier screening panels, and other resources to create a multipurpose, multiethnic array. With > 1.7 million expertly selected markers, the Infinium Multi-Ethnic Global BeadChip enables identification of genetic associations with common and rare traits, providing insight across diverse populations to epidemiologists, health care researchers, population geneticists, and genomic researchers (Tables 1–5).

Maximized Imputation Accuracy

Consortium partners developed content for the Infinium Multi-Ethnic Global BeadChip using tagging strategies with the power to perform more effective association studies in diverse populations. The novel algorithm selects population-specific and transethnic tag SNPs that maximize imputation accuracy, as imputation has become a standard practice in the interpretation of genotyping data and allows for more accurate statistical inference of genotypes not directly genotyped.

Expert-Selected Content

The Infinium Multi-Ethnic Global BeadChip combines expertly selected markers and content from the most popular Illumina commercial arrays with the most current genomic information. Researchers can detect both common and rare variants across the most commonly studied 5 superpopulations and impute variants in a vast number of subpopulations.

The Infinium Multi-Ethnic Global BeadChip contains the following content:

- Infinium HumanCore-24 BeadChip content with highly informative genome-wide tag SNPs
- African Diaspora Consortium Power Chip content identified through sequencing of 692 individuals by CAAPA
- Genome-wide coverage for diverse populations selected by PAGE using a new cross-population tagging strategy
- Total exonic content of > 420,000 markers including the Infinium Exome-24 BeadChip content and Multiethnic exome content designed by PAGE
- Over 17,000 variants chosen to be relevant to clinical and pharmacogenetic studies and 23,000 hand-curated variants picked for functional, immunological, oncological, ancestry, and forensic applications

Table 1: Multi-Ethnic Global BeadChip Product Information

Feature	Description	
Total No. of Markers	1,779,819	
Capacity for Custom Bead Types	245,000	
No. Samples per BeadChip	8	
DNA Input Requirement	200 ng	
Assay Chemistry	Infinium LCG	
Instrument Support	iScan® or HiScan® System	
Sample Throughput[a]	~ 1067 samples/week	
Scan Time per Sample	iScan System 11.3 min	HiScan System 6.5 min
Data Performance	**Value[b]**	**Product Specification**
Call Rate	99.87%	> 99% avg.
Reproducibility	99.99%	> 99.9%
Log R Deviation	0.10	< 0.30
Spacing **Mean**	**Median**	**90th %[c]**
Spacing (kb) 1.68	0.78	4.22

a. Estimated sample throughput based on use of 1 HiScan System, 1 AutoLoader 2.x, 1 Tecan robot, and a 5-day work week.
b. Values are derived from genotyping 708 HapMap reference samples.
c. Values are expected for typical projects using standard Illumina protocols. Tumor samples and samples prepared by methods other than standard Illumina protocols are excluded.

Table 2: Imputation Accuracy for 5 Populations from 1kGP at Different MAF Thresholds

	Minor Allele Frequency (MAF) Threshold		
Population[a]	**0.5–1%**	**1–5%**	**≥ 5%**
AFR	78.1%	89.5%	95.8%
AMR	82.7%	90.2%	96.9%
EAS	57.4%	82.4%	96.1%
EUR	70.3%	87.9%	97.2%
SAS	61.6%	84.8%	96.4%

a. AFR: African; AMR: Ad-mixed American; EAS: East Asian; EUR: European; SAS: South Asian.[1]

Table 3: LD Mean r^2 for 5 Populations from 1kGP at Different MAF Thresholds

	Minor Allele Frequency (MAF) Threshold		
Population[a]	**0.5–1%**	**1–5%**	**≥ 5%**
AFR	0.171	0.412	0.671
AMR	0.449	0.629	0.841
EAS	0.300	0.636	0.873
EUR	0.307	0.616	0.870
SAS	0.302	0.609	0.854

a. AFR: African; AMR: Ad-mixed American; EAS: East Asian; EUR: European; SAS: South Asian.[1]

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

Table 4: Multi-Ethnic Global BeadChip Marker Information

Marker Categories		Number of Markers	
Exonic Markers		424,843	
Intronic Markers		626,336	
Nonsense Markers		20,671	
Missense Markers		332,835	
Synonymous Markers		22,558	
Mitochondrial Markers		802	
Indels		24,458	
Sex Chromosomes	X 50,970	Y 3602	PAR / Homologous 4314

Ordering Information

Infinium Multi-Ethnic Global-8 v1.0 Kit	Catalog No.
16 samples	WG-316-1001
48 samples	WG-316-1002
96 samples	WG-316-1003
384 samples	WG-316-1004
Infinium Multi-Ethnic Global-8+ v1.0 Kit[a]	**Catalog No.**
16 samples	WG-316-1011
48 samples	WG-316-1012
96 samples	WG-316-1013
384 samples	WG-316-1014

a. Enabled for additional custom content.

Learn More

To learn more about the Infinium Multi-Ethnic Global v1.0 BeadChip and other Illumina genotyping products and services, visit www.illumina.com/techniques/popular-applications/genotyping.html

References

1. 1000 Genomes (www.1000genomes.org). Accessed July 21, 2015.
2. COSMIC: Catalogue of somatic mutations in cancer (http://cancer.sanger.ac.uk/cosmic). Accessed July 21, 2015.
3. Gene Ontology Consortium (www.geneontology.org). Accessed July 21, 2015.

Table 5: Multi-Ethnic Global BeadChip High-Value Content

Content	No. of Markers	Research Application / Note
ADME Core and Extended Genes[a]	18,234	Drug metabolism and excretion
ADME Core and Extended Genes[a] +/- 10 kb	21,955	Drug metabolism and excretion (+ regulatory regions)
APOE	91	Cardiovascular disease, Alzheimer's disease, immunoregulation, and cognition
Blood Phenotype Genes	2766	Blood phenotypes
COSMIC Genes[a]	901,123	Somatic mutations in cancer
GO CVS Genes[a]	243,272	Cardiovascular conditions
Database of Genomic Variants	1,360,391	Genomic structural variation
eQTLs	6100	Genomic loci regulating mRNA expression levels
Fingerprint SNPs	475	Human identification
HLA Genes	849	Disease defense, transplant rejection, and autoimmune disorders
Extended MHC[a]	16,526	Disease defense, transplant rejection, and autoimmune disorders
KIR[a]	121	Autoimmune disorders and disease defense
Neanderthal SNPs	2012	Neanderthal ancestry and human population migration
NHGRI GWAS Catalog[a]	10,061	Markers from published genome-wide association studies
RefSeq 3' UTRs	40,175	3' untranslated regions of known genes
RefSeq 5' UTRs	26,988	5' untranslated regions of known genes
RefSeq All UTRs	65,229	All untranslated regions of known genes
RefSeq	992,773	All known genes
RefSeq +/-10 kb	1,119,262	All known genes +/- 10 kb to include regulatory regions
RefSeq Promoters	39,280	2 kb 5' of all known genes to include promoter regions
RefSeq Splice Regions	10,338	Variants at splice sites in all known genes

a. ADME: absorption, distribution, metabolism, and excretion; COSMIC: catalog of somatic mutations in cancer[2]; GO CVS: Gene Ontology annotation of the cardiovascular system[3]; eQTL: expression quantitative trait loci; HLA: human leukocyte antigen; KIR: killer cell Ig-like receptor; MHC: major histocompatibility complex; NHGRI: national human genome research institute.

Illumina • 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.

© 2016 Illumina, Inc. All rights reserved. Illumina, HiScan, Infinium, iScan, and the pumpkin orange color are trademarks of Illumina, Inc. and/or its affiliate(s) in the U.S. and/or other countries. Pub. No. 370-2016-001 Current as of 25 January 2016

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGACG
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA