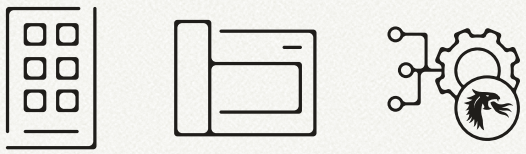
illumina®

Infinium™ H3Africa Consortium Array v2

Powerful array for genetic studies focused on African populations

- Comprehensive content includes > 2.2M markers with ~60K variants relevant to African populations
- Genome-wide scaffold detects common and low-frequency variants across a range of phenotypes
- Trusted Infinium chemistry delivers high-quality, accurate, and reproducible genotyping data



Overview

The Infinium H3Africa Consortium Array v2 (Figure 1, Table 1) is a powerful genotyping array for identifying genetic associations with common and rare traits among African populations. Specifically designed by the H3Africa Consortium, the array harnesses content from Illumina Omni2.5-8 and Omni5-4 BeadChips, in addition to custom content selected by the consortium based on whole-genome sequencing (WGS) data. The custom content was selected by including specific single nucleotide polymorphisms (SNPs) requested by H3Africa Consortium projects, SNPs within the MHC region, X-chromosome and mitochondrion SNPs, and SNPs of clinical or pharmacogenomic interest. Remaining custom SNPs were selected to improve coverage, imputation accuracy, and enrichment in novel but common variants in African populations based on sequencing data.

Consortium-selected content

The H3Africa Consortium selected approximately 10,000 variants relevant to specific diseases of interest, including variants known to be associated with kidney disease, diabetes, sickle cell disease, cardiometabolic diseases, and susceptibility to infectious diseases. Additional variants from PharmGKB,[1,2] NHGRI-EBI genome-wide association study (GWAS) catalog,[3] ClinVar,[4] and the COSMIC[5] database were also identified and selected by the consortium. For PharmGKB and the GWAS catalog, 4000 and 24,000 variants, respectively, that occur with a minor allele frequency (MAF) ≥ 0.01 in at least one of the African populations were selected for inclusion. For ClinVar, the consortium restricted their selection to 27,000 SNPs with MAF ≥ 0.05 in at least one of the African populations. For variants in the COSMIC database, the consortium retained 20,000 variants that were substitutions, characterized as pathogenic in the database and showed MAF ≥ 0.05 in at least one of the African populations. A part of this list was already present in the fixed content, and the remaining 60,000 SNPs were added as a component of the custom content. To optimize ancestry inference, the consortium identified a panel consisting of ~2000 variants of carefully chosen mitochondrial, Y-chromosome, and African-centric ancestry-informative markers (AIMs) (Table 2, Table 3).

Figure 1: Infinium H3Africa Consortium Array v2

The 8-sample BeadChip includes common SNP variation and ~10K variants relevant to diseases of interest in African populations.

Table 1: Product specifications

Feature	Description
Species	Human
Total number of markers	2,271,503
No. of samples per BeadChip	8
DNA input requirement	200 ng
Assay chemistry	Infinium LCG
Instrument support	iScan™ System
Sample throughput	~1728 samples per week
Scan time per BeadChip	35 min

Table 2: Marker category information

Marker categories[a]	No. of markers	
Exonic markers[b]	88,785	
Intronic markers[b]	1,074,881	
Nonsense markers[c]	313	
Missense markers[c]	23,548	
Synonymous markers[c]	21,654	
Mitochondrial markers[c]	234	
Insertion/deletions[d]	318	
Sex chromosomes[c]	X	Y
	36,347	2528

a. Number of markers calculated from the consortium manifest.
b. RefSeq: NCBI Reference Sequence Database.[6]
c. Compared against the UCSC Genome Browser.[7]
d. NCBI Genome Reference Consortium, version GRCh37.[8]

Infinium H3Africa Array reference samples

WGS data obtained from ~3480 individuals from 17 African countries were used in the array design. The sequencing coverage for these samples varied from 4× to 30×. Included in this cohort, the H3Africa Consortium contributed ~350 samples for sequencing at the Baylor College of Medicine to generate high coverage data and fill some of the gaps in missing populations or countries. The TrypanoGen project from H3Africa had additional medium coverage sequence data for 118 samples that were contributed for the design.

Table 3: High-value content

Content	No. of markers[a]	Application/notes
ACMG[9] 59 2016 gene coverage	5201	Variants with known clinical significance identified from WGS and WES samples
ADME[2] CPIC genes	3258	Drug metabolism and excretion (including regulatory regions)
ADME[2] core and extended + CPIC genes ± 10 kb	24,168	
AIMs	2563	Ancestry-informative markers
ClinVar[4] variants	8590	Relationships among variation, phenotypes, and human health
ClinVar[4] pathogenic	85	
ClinVar[4] likely pathogenic	24	
ClinVar[4] benign	4679	
ClinVar[4] likely benign	4097	
COSMIC[5] genes	81,841	Somatic mutations in cancer
eQTLs[10]	8219	Genomic loci regulating mRNA expression
gnomAD exome	65,982	WES and WGS results from unrelated individuals
HLA genes[11]	1164	Disease defense, transplant rejection, autoimmune disorders
Extended MHC[11, b]	24,411	
NHGRI-EBI GWAS catalog[3]	44,344	Published GWAS markers
PharmGKB[1] phenotype annotation	1846	Human genetic variation associated with drug responses, variants affecting a phenotype, with or without drug information
PharmGKB[1] drug annotation	1729	Variants affecting drug response, dose, metabolism, etc
PharmGKB[1] functional analysis annotation	149	*In vitro* and functional analysis-type associations

a. Number of markers for each category subject to change.
b. Extended MHC is an 8 Mb region.

ACMG, American College of Medical Genetics; ADME, absorption, distribution, metabolism, excretion; COSMIC, catalog of somatic mutations in cancer; eQTL, expression quantitative trait loci; gnomAD, Genome Aggregation Database; HLA, human leukocyte antigen; MHC, major histocompatibility complex; NHGRI, National Human Genome Research Institute; PharmGKB, Pharmacogenomics Knowledgebase.

Learn more →

Contact your local sales representative to learn more about the Infinium H3Africa Consortium Array v2:

North America: 800.809.4566

Europe, Middle East, Africa: +44.1799.534000

Other regions: www.illumina.com/company/contact-us.html

H3Africa Consortium: https://h3africa.org/

Ordering information	
Product	Catalog no.
Infinium H3Africa Consortium Array v2 (48 samples)	15056943
Infinium H3Africa Consortium Array v2 (96 samples)	15056944
Infinium H3Africa Consortium Array v2 (384 samples)	15056945

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.23502.
2. Relling M, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Ther.* 2011;89(3):464-467. doi:10.1038/clpt.2010.279.
3. Sollis E, Mosaku A, Abid A, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* 2023;51(D1):D977-D985. doi:10.1093/nar/gkac1010.
4. NCBI. ClinVar Database website. ncbi.nlm.nih.gov/clinvar. Accessed February 3, 2025.
5. Tate JG, Bamford S, Jubb HC, et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* 2019;47(D1):D941-D947. doi:10.1093/nar/gky1015.
6. NCBI. Reference Sequence Database website. ncbi.nlm. nih.gov/refseq. Accessed February 3, 2025.
7. UCSC Genomics Institute. UCSC Genome Browser. genome.ucsc.edu. Accessed February 3, 2025.
8. Genome Reference Consortium. Human Genome Overview Version GRCh37 website. ncbi.nlm.nih.gov/grc/human. Accessed February 3, 2025.
9. Green RC, erg J, 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.
10. Wong KM, Langlais K, Tobias G, 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.
11. 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.

illumina®

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

M-GL-03384 v1.0