DATA SHEET

Infinium™ Chinese Genotyping Array-24 v1.0 BeadChip

A powerful, high-quality, cost-effective array for population-scale genetic studies in the Chinese population

- Updated GWAS content focusing on the Chinese population
- Genome-wide scaffold to detect common and low-frequency variants
- High call rates and reproducibility from diverse sample types, including, saliva, blood, solid tumors, and buccal swabs

illumina®

Introduction

The Infinium Chinese Genotyping Array-24 v1.0 BeadChip provides a scalable, and cost-effective solution for variant screening and precision medicine research in the Chinese population (Table 1, Figure 1). Content includes the powerful genome-wide backbones from the Infinium Global Screening Array-24 v3.0 and the Infinium Asian Screening Array-24 v1.0, and was designed in collaboration with key scientific leaders in China to ensure optimal coverage and imputation in the Chinese population (Figure 2). The BeadChip includes up-to-date clinical research variants for a broad range of applications, including complex disease studies, pharmacogenomics (PGx) research, and more.

Table 1: Infinium Chinese Genotyping Array-24 v1.0 at a glance

Feature	Description
Species	Human
Total number of markers[a]	684,023
Capacity for custom beadtypes	50,000
Number of samples per BeadChip	24
DNA input requirement	200 ng
Assay chemistry	Infinium HTS
Instrument support	iScan™ System
Sample throughput[b]	~2304 samples/week
Scan time per sample	2.5 min

a. Total number of markers calculated from commercial manifest
b. Estimated sample processing assumes 1 iScan System, 1 AutoLoader, 2 Tecan robots, and a 5-day work week

Abbreviations: HTS, high-throughput screening

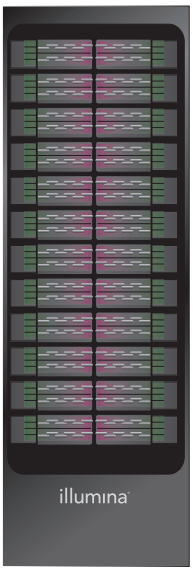

Figure 1: The Infinium Chinese Genotyping Array-24 v1.0 BeadChip—The BeadChip is built on the trusted 24-sample Infinium HTS platform and provides a cost-effective tool for accurately analyzing 684,023 SNPs in the Chinese population.

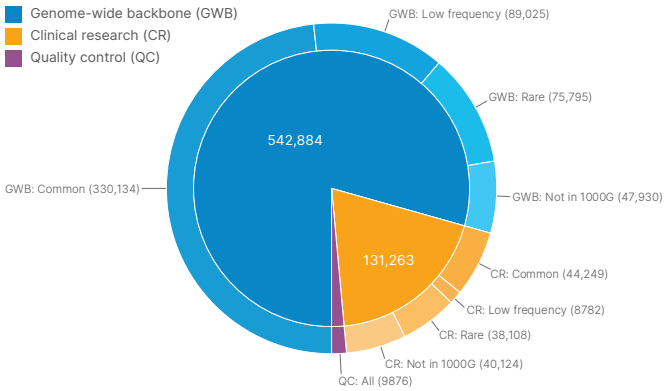

Figure 2: Summary of content—Genome-wide content enables a broad range of clinical research and genetic variant screening applications. Plotted in the inner pie is the proportion of the array that was selected for genome-wide coverage (blue), clinical research (yellow), and quality control (purple).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

Optimized Chinese content from whole-genome samples

The Infinium Chinese Genotyping Array-24 v1.0 BeadChip contains highly informative tag single nucleotide polymorphisms (SNP) in Chinese populations. It includes over 540,000 genome-wide backbone markers carefully selected from the Infinium Global Screening Array v3.0 and the Infinium Asian Screening Array-24 v1.0 for optimized genome-wide association studies in the Chinese population. There are also quality control (QC) markers enabling sample identification, tracking, ancestry determination, and stratification (Figure 3).

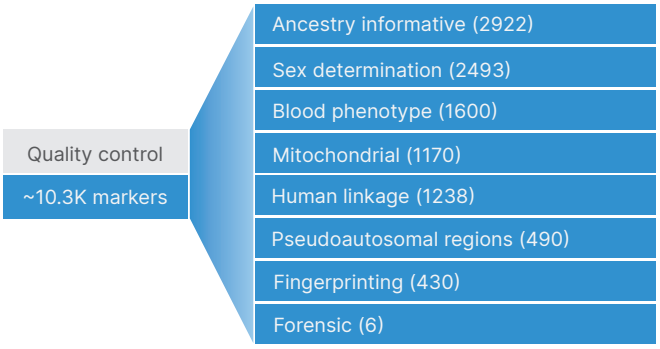

Figure 3: Infinium Chinese Genotyping Array-24 v1.0 QC markers—Approximately ~10.3K QC markers on the BeadChip enable various capabilities for sample tracking, such as sex determination, continental ancestry, and human identification.

In addition to previously identified variants, the Infinium Chinese Genotyping Array-24 v1.0 contains ~170K SNP markers contributed by collaborator-owned, Chinese whole-genome sequencing data. The resulting BeadChip will power growing biobank and translational research studies on the Chinese population. It supports broad applications, including disease marker association, risk profiling, PGx, lifestyle, wellness, and other marker discovery research (Table 2, Table 3).[1–5]

Table 2: Infinium Chinese Genotyping Array-24 v1.0 marker summary

Marker categories			No. of markers[a]
Exonic markers[4]			87,172
Nonsense markers[6]			5510
Missense markers[6]			48,198
Synonymous markers[6]			8972
Mitochondrial markers[4]			1170
Indels[6]			13,615
	X	Y	PAR/ homologous
Sex chromosomes	28,567	6661	880

a. Markers are calculated from commercial manifest

Abbreviations: indel, insertion/deletion; PAR, pseudoautosomal region

Chinese-specific content empowers clinical research

The clinical research content of the Infinium Chinese Genotyping Array-24 v1.0 BeadChip was designed through collaboration with Chinese medical genomics experts using multiple annotation databases and collaborator-owned, Chinese whole-genome sequencing data. The 130,000 clinical markers create a highly informative, cost-effective panel for clinical research in Chinese population (Table 3, Figure 4).

Variants included on the array consist of markers with known disease association based on ClinVar,[7] the Pharmacogenomics Knowledge Base (PharmGKB),[8] and the National Human Genome Research Institute (NHGRI)-EBI database. In addition to disease-associated markers, the Infinium Chinese Genotyping Array-24 v1.0 BeadChip contains imputation-based tag SNPs for HLA alleles, extended MHC region, the KIR gene, and exonic content from the gnomAD database.[9]

Table 3: Infinium Chinese Genotyping Array-24 v1.0 high-value content

Content[a]	No. of markers	Research application/note	Content[a]	No. of markers	Research application/note
ACMG[2] 59 2016 gene list	19,241	Variants with known clinical significance identified from clinical WGS and WES samples	GO[14] CVS genes	105,715	Cardiovascular conditions
ACMG 59 all annotations	16,026		Database of Genomic Variants[15]	531,844	Genomic structural variation
ACMG 59 benign	607		eQTLs[16]	3699	Genomic loci regulating mRNA expression levels
ACMG 59 likely benign	645		Fingerprint SNPs[17]	454	Human identification
ACMG 59 pathogenic	9104		gnomAD exome	78,610	Exome and whole-genome sequences from unrelated individuals sequenced as part of various studies
ACMG 59 likely pathogenic	2998		HLA genes	519	Disease defense, transplant rejection, and autoimmune disorders
ACMG 59 VUS	1506		Extended MHC[18d]	9276	Disease defense, transplant rejection, and autoimmune disorders
AIMs[c]	2595	Ancestry-informative markers	KIR genes[6]	29	Autoimmune disorders and disease defense
APOE[13]	16	Cardiovascular disease, Alzheimer's disease, immunoregulation, and cognition	Neanderthal SNPs[19]	805	Neanderthal ancestry and human population migration
Blood phenotype genes[13]	1902	Blood phenotypes	Newborn/carrier screening gene coverage[20]	26,007	Genes associated with severe, recessive childhood diseases included in the TruSight™ Inherited Disease Sequencing Panel
ClinVar[7] variants	51,852	Relationships among variation, phenotypes, and human health	NHGRI-EBI GWAS catalog[10]	33,335	Markers from published genome-wide association studies
ClinVar pathogenic	18,845		NHGRI diseases	27,251	Markers related to various diseases from published studies
ClinVar likely pathogenic	7158		PharmGKB[8]	4182	Human genetic variation associated with drug responses
ClinVar benign	9946		RefSeq[12] 3' UTRs	13,947	3' untranslated regions of known genes
ClinVar likely benign	5041		RefSeq 5' UTRs	6549	5' untranslated regions of known genes
COSMIC[13] genes	308,677	Somatic mutations in cancer	RefSeq All UTRs	19,893	All untranslated regions of known genes
CPIC[21] ALL	398	Variants with potential guidelines to optimize drug therapy	RefSeq	344,363	All known genes
CPIC-A	257		RefSeq +/– 10 kb	403,181	All known genes plus regulatory regions
CPIC-A/B	1		RefSeq Promoters	11,516	2 kb upstream of all known genes to include promoter regions
CPIC-B	16		RefSeq Splice Regions	2622	Variants at splice sites in all known genes
CPIC-C	37				
CPIC-C/D	1				
CPIC-D	63				

a. Imputation content is derived from commercial manifest
b. Variant counts are current as of May 2021—the number of markers for each content category is subject to change
c. Based on internal calculations
d. 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 Knowledge Base; RefSeq, NCBI Reference Sequence Database; UTR, untranslated region

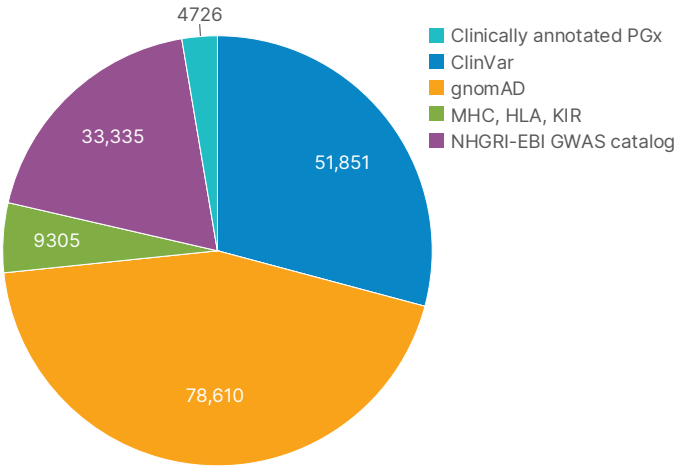

Figure 4: Clinical research content—Content was expertly selected from scientific databases to create a highly informative array for clinical research applications. Variant counts may be subject to change.

Broad spectrum of PGx markers and disease categories

The Infinium Chinese Genotyping Array-24 v1.0 BeadChip provides coverage of PGx variants associated with absorption, distribution, metabolism, and excretion (ADME) phenotypes based on PharmGKB (reference) and Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines (Figure 5).[21]

Clinical research content on the Infinium Chinese Genotyping Array-24 v1.0 BeadChip enables validation of disease associations, risk profiling, preemptive screening research, and PGx studies. Variant selection includes a range of pathology classifications based on ClinVar and American College of Medical Genetics (ACMG) annotations (Figure 6).[2] The BeadChip contains extensive coverage of phenotypes and disease classifications based on ClinVar and the NHGRI-EBI GWAS Catalog (Figure 7).

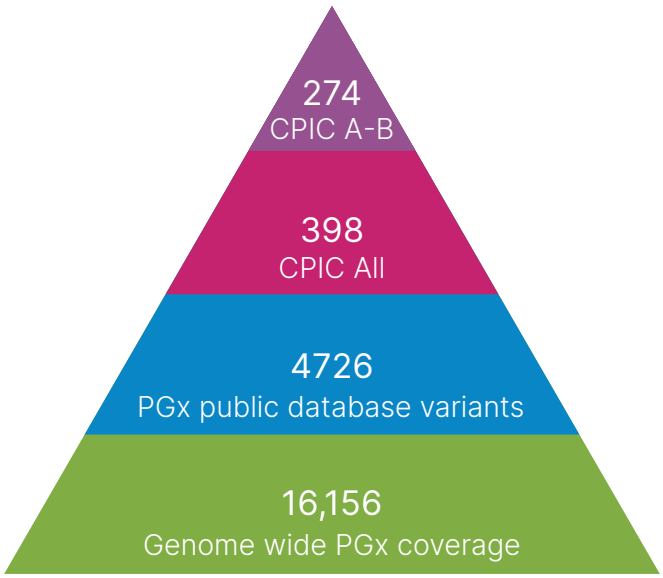

Figure 5: Clinical research content—The BeadChip features an extensive list of pharmacogenomics markers selected based on CPIC guidelines and the PharmGKB database.[8,21] From top to bottom: CPIC A-B, CPIC levels A, B, A/B, all have guidelines and function (decreased/increased/deficient); CPIC all, all CPIC variants regardless of function and level; PGx public database variants, all variants annotated in PharmGKB (with manual curation), PharmVar (with manual curation and manuscript lookups), manual haplotype tables; genome-wide PGx coverage, includes markers located in an extended ADME genes or CPIC level A genes, including targeted imputation tag SNPs and CPIC level A CNV tags.

Flexible content options

The Infinium Chinese Genotyping Array-24 v1.0 BeadChip can be customized to incorporate up to 50,000 custom SNP targets (Table 4). Visit the Human Consortia page (illumina.com/science/consortia/human-consortia.html) for more information about other collaboratively pre-designed panels.

Table 4: Infinium Chinese Genotyping Array-24 v1.0 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[22]
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

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

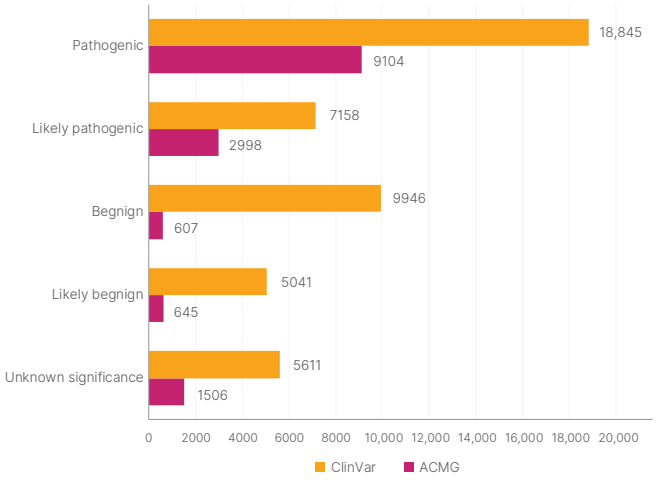

Figure 6: Broad coverage of disease categories—Variants sorted by range of pathology classifications according to ClinVar and American College of Medical Genetics (ACMG) annotations. Variant counts may be subject to change.

Robust, powerful genotyping

The Infinium Chinese Genotyping Array-24 v1.0 BeadChip uses trusted Infinium assay chemistry to deliver the same high-quality, reproducible data achieved with all Illumina genotyping microarrays (Table 5). 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 formalin-fixed, paraffin-embedded (FFPE) QC and DNA Restoration Kits[23], enabling genotyping of FFPE samples. In addition, the high signal-to-noise ratio of the individual genotyping calls from the assay provides access to genome-wide copy number variant (CNV) calling with a mean probe spacing of ~4.4 kb.

Figure 7: Disease research content covering diverse populations—The BeadChip includes extensive coverage of numerous phenotypes and disease classifications based on (A) ClinVar and (B) NHGRI-GWAS categories.

High coverage and imputation accuracy for Chinese populations

Infinium Chinese Genotyping Array-24 v1.0 BeadChip content was selected based on available whole-genome reference data. The resulting assay produces high imputation accuracy for Chinese populations that have limited representation in the 1000 Genomes Project (Table 6).

High-throughput workflow

The Infinium Chinese Genotyping Array-24 v1.0 BeadChip uses the highly scalable Infinium 24-sample Infinium HTS format that enables laboratories to efficiently scale as needed. For flexible throughput processing, the Infinium HTS assay provides the capability to run hundreds, and potentially thousands, of samples per week. The Infinium HTS assay provides a rapid, three-day workflow that allows users to gather and report data quickly (Figure 8). For labs interested in quickly scaling or increasing efficiency, the Illumina Consulting Service offers customized solutions.

Table 5: Data performance and spacing

Data performance	Value[a]	Specification[b]	
Call rate	99.37%	> 99.0% avg	
Reproducibility	99.99%	> 99.90%	
Log R deviation	0.11	< 0.30 avg[c]	
Spacing	**Mean**	**Median**	**90th. percentile[d]**
Spacing (kb)	0.42	0.22	1.0

a. Values are derived from genotyping 683 HapMap reference samples—excludes Y chromosome markers for female 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

Table 6: Imputation accuracy at various MAF thresholds for select populations[a]

Population	MAF ≥ 5%	MAF 1-5%	MAF 0.05-1%
Chinese	0.90	0.81	0.73

a. Underrepresented populations are defined by comparisons of population-specific samples included in the 1000 Genomes Project

Abbreviations: MAF: minor allele frequency

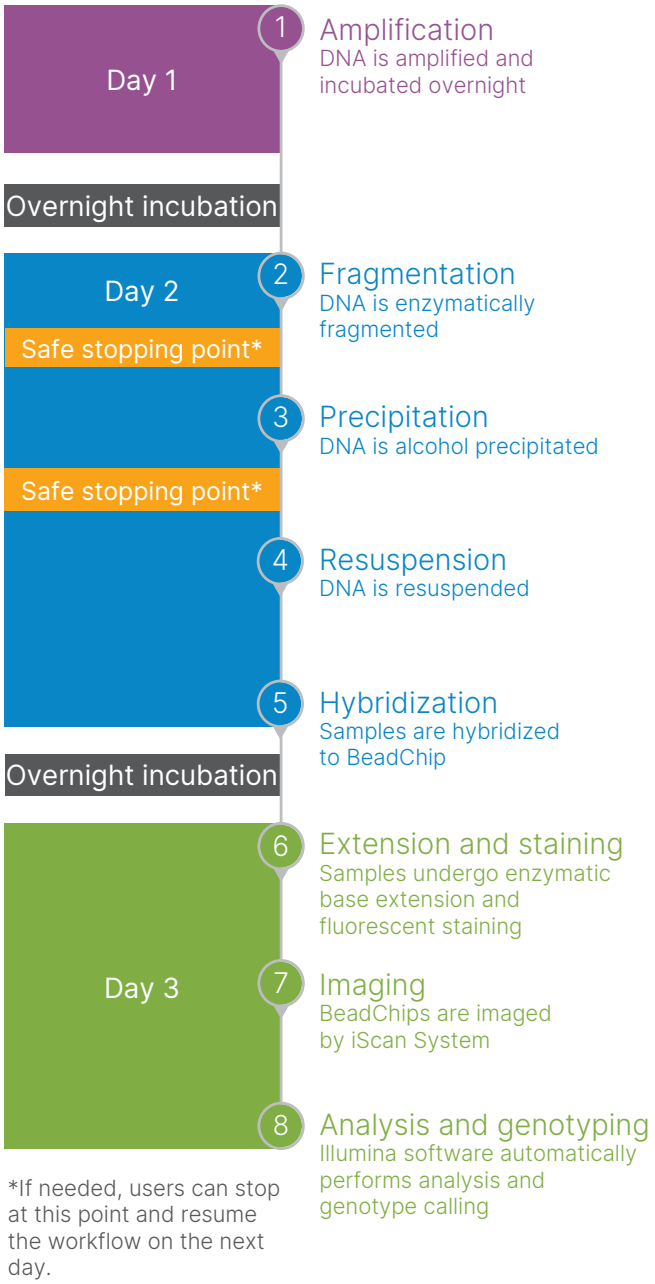

Figure 8: Fast workflow—The Infinium HTS format provides a rapid three-day workflow with minimal hands-on time.

Learn more

Infinium Chinese Genotyping Array-24 v1.0 BeadChip and other Illumina genotyping products, illumina.com/genotyping

Ordering information

For labs interested in higher throughput processing with the Infinium Chinese Genotyping Array, contact your local sales representative for information about Infinium high throughput (HTS) kit configurations.

Infinium Chinese Genotyping Array-24 v1.0 BeadChip	Catalog no.
48 samples	20039075
288 samples	20039076
1152 samples	20039077
4608 samples	20039078
23,040 samples	20039079
Infinium Chinese Genotyping Array-24+ v1.0 BeadChip[a]	**Catalog no.**
48 samples	20039014
288 samples	20039015
1152 samples	20039016
4608 samples	20039017
23,040 samples	20039018

a. Enabled for custom content

References

1. European Bioinformatics Institute. The 1000 Genomes Project website. 1000genomes.org. Accessed May 15, 2021.
2. American College of Medical Genetics and Genomics. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. ncbi.nlm.nih.gov/clinvar/docs/acmg. Accessed May 15, 2021.
3. PharmaADME. PharmaADME Gene List website. http://pharmaadme.org/joomla/index.php?option=com_content&task=view&id=12&Itemid=27. Accessed May 15, 2021.
4. University of California, Santa Cruz Genomics Institute. UCSC Genome Browser. genome.ucsc.edu. Accessed May 15, 2021.
5. National Center for Biotechnology Information. NCBI Reference Sequence Blood Group Antigen Gene Mutation Database. ncbi.nlm.nih.gov/projects/gv/rbc/xslcgi.fcgi?cmd=bgmut/systems. Accessed May 15, 2021.
6. Genome Reference Consortium. Genome Reference Consortium Version GRCh37. ncbi.nlm.nih.gov/grc/human. Accessed May 15, 2021.
7. National Center for Biotechnology Information. ClinVar Database. ncbi.nlm.nih.gov/clinvar. Accessed May 15, 2021.
8. Whirl-Carrillo M, McDonagh EM, Hebert JM, et al. Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther.* 2012;92(4):414-417.
9. gnomAD, Genome Aggregation Database website. gnomad.broadinstitute.org. Accessed May 15, 2021.
10. National Human Genome Research Institute. genome.gov. Accessed May 15, 2021.
11. Broad Institute. Exome Aggregation Consortium (ExAC) Browser. exac.broadinstitute.org. Accessed May 15, 2021.
12. National Center for Biotechnology Information. RefSeq: NCBI Reference Sequence Database. ncbi.nlm.nih.gov/refseq. Accessed May 15, 2021.
13. Sanger Institute. COSMIC, the Catalogue Of Somatic Mutations. cancer.sanger.ac.uk/cosmic. Accessed May 15, 2021.
14. Gene Ontology Consortium. The gene ontology resource. geneontology.org. Accessed May 15, 2021.
15. 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.
16. National Center for Biotechnology Information. NCBI eQTL Database. ncbi.nlm.nih.gov/projects/gap/eqtl/index.cgi. Accessed May 15, 2021.
17. The Allele Frequency Database. alfred.med.yale.edu/alfred/snpSets.asp. Accessed May 15, 2021.
18. 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.
19. Max Planck Institute for Evolutionary Anthropology. Neanderthal Genome Browser. neandertal.ensemblgenomes.org/index.html. Accessed May 15, 2021.
20. Illumina. TruSight Inherited Disease Sequencing Panel Data Sheet. illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_trusight_inherited_disease.pdf. Accessed May 15, 2021.
21. Clinical Pharmacogenetics Implementation Consortium (CPIC). cpicpgx.org. Accessed May 15, 2021.
22. Illumina. Microarray Assay Designer. Illumina.com. illumina.com/products/by-type/informatics-products/designstudio.html. Accessed May 15, 2021.
23. Illumina. Infinium FFPE QC and DNA Restoration Kit. Illumina.com. https://www.illumina.com/products/by-type/molecular-biology-reagents/infinium-ffpe-qc-dna-restoration.html. Accessed May 15, 2021.

illumina®

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

M-US-00004 v1.0