DATA SHEET

Infinium™ Global Screening Array with Enhanced PGx-48 v4.0

High-throughput genotyping array for pharmacogenomic and precision medicine research

- Comprehensive backbone featuring ~646K annotated variants enabling a range of research applications
- Supplemental ~41K markers supporting advanced pharmacogenomic research
- Automated workflow for high-throughput population genomics studies

illumina®

Introduction

The Infinium Global Screening Array with Enhanced PGx-48 v4.0 is a high-throughput genotyping microarray developed for pharmacogenomic (PGx) research, polygenic risk score development, ancestry determination, and genetic disease research. The array uses Infinium EX chemistry and is capable of assessing 646,681 backbone markers and 41,767 enhanced PGx markers accurately for 48 human samples in a single assay (Figure 1, Table 1).

The Infinium Global Screening Array with Enhanced PGx-48 v4.0 includes markers for ~16,000 absorption, distribution, metabolism, and excretion (ADME) markers spanning more than 2000 genes and exceptional coverage of priority level A and B Clinical Pharmacogenetics Implementation Consortium (CPIC) variants (Figure 2, Table 2).[1,2] Infinium EX chemistry features an optimized workflow with a targeted gene amplification step that supports detection of high-impact PGx genes like *CYP2D6, CYP2B6, and TPMT* that have historically been challenging to assess.

Table 1: Product specifications

Feature	Description
Species	Human
Total number of markers	656,275
Number of samples per BeadChip	48
DNA input requirement	200 ng
Enhanced PGx markers	41,767
Assay chemistry	Infinium EX
Instrument support[a]	iScan™ System Infinium Amplification System Infinium Automated Pipetting System with ILASS
Maximum iScan System sample throughput[b]	11,520 samples/week[c]
Scan time per BeadChip[b]	~30 minutes

a. The iScan System and Infinium Automated Pipetting System with ILASS are required for a fully supported solution. The Infinium Amplification System is optional.
b. Approximate values, scan times, and maximum throughput will vary depending on laboratory and system configurations.
c. Achieved with integration of AutoLoader 2.x automated array loading and five Infinium Automated Pipetting Solution with ILASS liquid handlers.

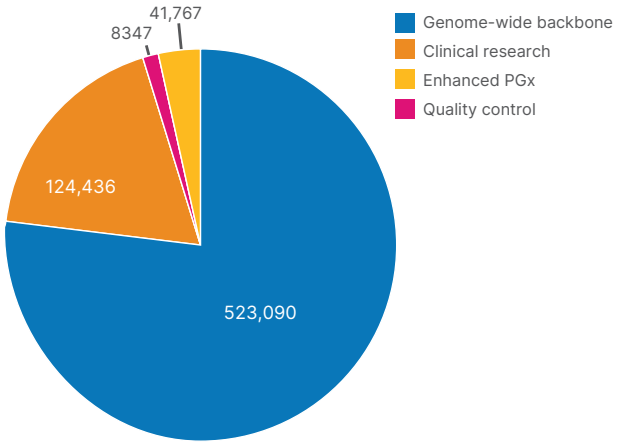

Figure 1: Annotated research database content summary—Distribution of markers for genome-wide coverage, clinical research, enhanced PGx, and quality control (QC).

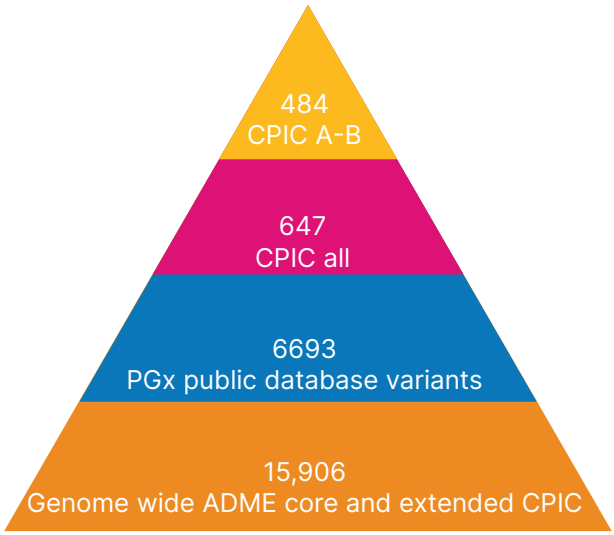

Figure 2: Broad spectrum of pharmacogenomics markers—Clinical research content developed from an extensive list of pharmacogenomics markers selected based on CPIC guidelines and the PharmGKB database.[1,2] Content includes PGx public database variants, variants annotated in PharmGKB, CPIC, genome-wide PGx coverage, extended ADME genes,[3] CPIC level A genes, including targeted imputation tag SNPs, and CPIC level A CNV tags.

Table 2: High-value content from key research databases

Content	No. of markers[a]	Research application/note	Content	No. of markers	Research application/note
ACMG[4] 59 2016 gene coverage	19,814	Variants with known clinical significance identified from clinical WGS and WES samples	GO[9] CVS genes	112,551	Cardiovascular conditions
ACMG 59 all annotations	17,851		Database of Genomic Variants[10]	511,192	Genomic structural variation
ACMG 59 pathogenic	9621		eQTLs[11]	2694	Genomic loci regulating mRNA expression levels
ACMG 59 likely pathogenic	3312		Fingerprint SNPs[12]	431	Human identification
ACMG 59 benign	955		gnomAD[13] exome	74,421	WES and WGS results from unrelated individuals from various studies
ACMG 59 likely benign	983		HLA genes[14]	1060	Disease defense, transplant rejection, and autoimmune disorders
ACMG 59 VUS	1788		Extended MHC[14c]	9655	Disease defense, transplant rejection, and autoimmune disorders
ADME[3] core and extended + CPIC genes	15,906	Drug absorption, distribution, metabolism, and excretion	KIR genes[5]	22	Autoimmune disorders and disease defense
ADME core and extended + CPIC genes +/- 10 kb	18,336	Includes regulatory regions	Neanderthal SNPs[15]	1537	Neanderthal ancestry and human population migration
AIMs	2796[b]	Ancestry-informative markers	Newborn/carrier screening gene coverage	26,823	Genes associated with childhood diseases included in the TruSight™ Inherited Disease Sequencing Panel[16]
APOE[5]	11	Cardiovascular disease, Alzheimer's disease, and cognition	NHGRI-EBI GWAS catalog[17]	7408	Markers from published GWAS
Blood phenotype genes[6]	2007	Blood phenotypes	PharmGKB[1] all	5296	Human genetic variation associated with drug responses
ClinVar[7] variants	56,780	Relationships among variation, phenotypes, and human health	PharmGKB level 1A	332	
ClinVar pathogenic	18,121		PharmGKB level 1B	7	
ClinVar likely pathogenic	7133		PharmGKB level 2A	58	
ClinVar benign	14,588		PharmGKB level 2B	51	
ClinVar likely benign	6376		PharmGKB level 3	1958	
ClinVar VUS	6799		PharmGKB level 4	487	
COSMIC[8] genes	307,966	Somatic mutations in cancer	RefSeq[18] 3' UTRs	14,506	3' untranslated regions[d]
CPIC[1] all	647	Variants with potential guidelines to optimize drug therapy	RefSeq 5' UTRs	6390	5' untranslated regions[d]
CPIC-A	454		RefSeq All UTRs	20,271	Untranslated regions[d]
CPIC-A/B	3		RefSeq	341,566	All known genes
CPIC-B	27		RefSeq +/- 10 kb	399,167	Regulatory regions[d]
CPIC-C	45		RefSeq Promoters	16,212	2 kb upstream to include promoter regions[d]
CPIC-C/D	2		RefSeq Splice Regions	3224	Variants at splice sites[d]
CPIC-D	61				

a. The number of markers for each category may be subject to change.
b. Based on internal calculations.
c. Extended MHC is an 8 Mb region.
d. Of all known genes.

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; VUS, variant of unknown significance; WES, whole-exome sequencing; WGS, whole-genome sequencing.

Diverse backbone with enhanced exonic coverage

The Infinium Global Screening Array with Enhanced PGx-48 v4.0 is built on a high-density SNP backbone that is optimized for cross-population imputation coverage. The genome-wide content includes enhanced tagging in exonic regions and enriched coverage of loci from genome-wide association studies (GWAS) with known disease or trait associations (Figure 2, Table 3).

More than 74,000 exome markers were selected from individuals representing diverse ethnic backgrounds, including African Americans, Hispanics, Pacific Islanders, East Asians, Europeans, and individuals of mixed ancestry. The Infinium Global Screening Array with Enhanced PGx-48 v4.0 also features exonic content from populations in the Exome Aggregation Consortium (ExAC) database, including cross-population and population-specific markers with functionality annotations or strong evidence for association (Table 4). The inclusive design allows for multiple applications, including polygenic risk scoring, ancestry examination, and genetic disease research based on reported variants.

Table 3: Marker information

Marker categories			No. of markers
Exonic markers[a]			84,571
Intronic markers[a]			272,922
Nonsense markers[b]			5347
Missense markers[b]			46,443
Synonymous markers[b]			8816
Mitochondrial markers[b]			1108
Indels[c]			10,728
Sex chromosomes[c]	X	Y	PAR/homologous
	28,636	3893	821

a. RefSeq-NCBI Reference Sequence Database.[18]
b. Compared against the UCSC Genome Browser.[5]
c. NCBI Genome Reference Consortium, Version GRCh38.[19]

Table 4: Exonic coverage across populations

Population(s)[a,b]	No. of markers
NFE	57,614
EAS	32,724
AMR	46,913
AFR	43,583
SAS	41,674
NFE/EAS/AMR/AFR/SAS	23,177

a. internationalgenome.org/category/population.
b. Based on gnomAD, gnomad.broadinstitute.org.

NFE, non-Finnish European; EAS, East Asian; AMR, Ad Mixed American; AFR, African; SAS, South Asian

Broad coverage of variants with known disease associations

Content on the Infinium Global Screening Array with Enhanced PGx-48 v4.0 is designed for high-value pharmacogenomic and clinical research applications. The array provides extensive coverage of variants selected from the National Human Genome Research Institute genome-wide association studies (NHGRI-GWAS) catalog[17] representing a range of phenotypes and disease classifications. The selection of markers provides extensive opportunities for researchers interested in studying diverse populations.

The clinical research content on the array enables validation of previously identified disease associations, risk profiling, predictive screening research, and pharmacogenomic studies. Variant selection includes a range of pathology classifications based on ClinVar and American College of Medical Genetics (ACMG) annotations.[3] Selected variants represent an extensive range of phenotypes and disease classifications based on ClinVar and the NHGRI-GWAS catalog (Figure 3). The content also covers ACMG and ClinVar database variants with a range of phenotypes pathogenic, likely pathogenic, and variants of unknown significance (VUS), as well as benign variants (Figure 4).

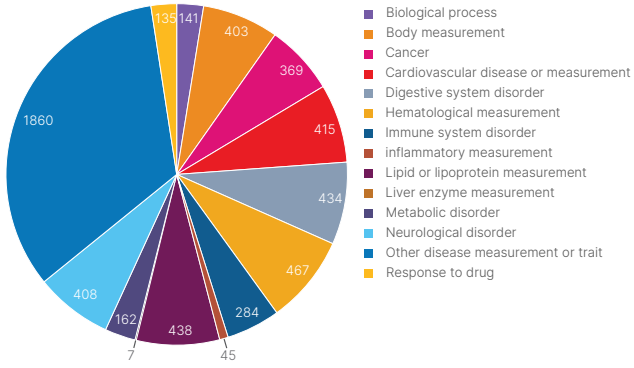

Figure 3: Disease research content covering diverse populations based on NHGRI GWAS database categories.

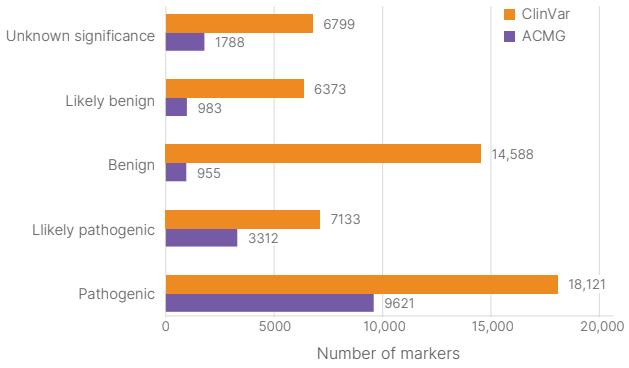

Figure 4: Distribution of variant pathology classifications according to ClinVar and ACMG annotations

Updated and research content

Databases, such as ClinVar, are constantly evolving with the addition of new variants and as variants change designation to "pathogenic" or "likely pathogenic" categories. The Infinium Global Screening Array with Enhanced PGx-48 v4.0 provides updated information for many high-value variants contained within these annotated databases. Variants included on the array consist of markers with known disease association selected from ClinVar, PharmGKB, and the NHGRI-EBI database.[8,18] The array also provides imputation-based tag SNPs for HLA alleles, extended MHC region, the KIR gene, and exonic content from the gnomAD database (Table 2, Figure 5).[13]

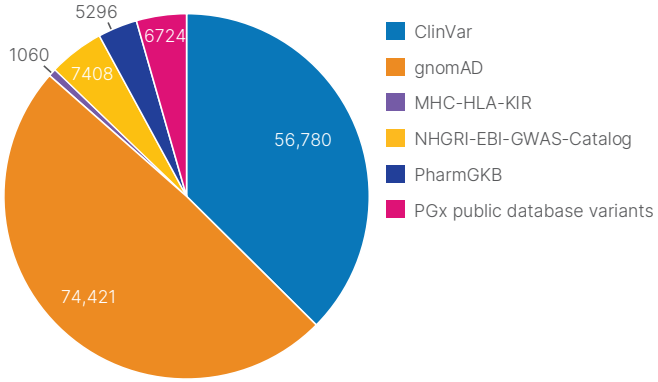

Figure 5: Expertly selected clinical research content from key databases supports a broad range of applications.

QC markers

The Infinium Global Screening Array with Enhanced PGx-48 v4.0 includes ~8.3K quality control (QC) markers. QC markers on the array are selected to facilitate high-throughput studies and enable sample tracking functions, including ancestry determination, and stratification (Figure 6).

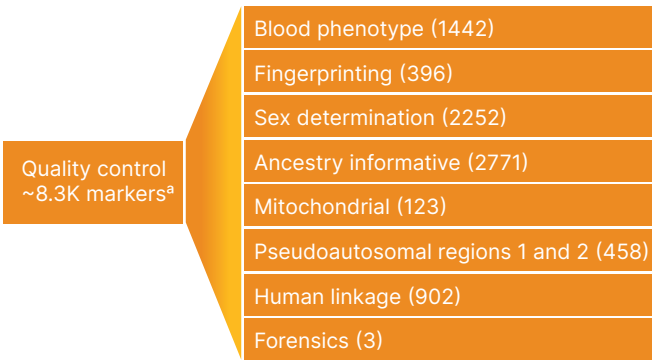

Figure 6: QC content by category—The Infinium Global Screening Array with Enhanced PGx-48 v4.0 contains ~8.3K QC markers enabling various sample tracking functions such as sex determination, continental ancestry, human linkage, and more.
a. Counts contain some markers that are represented in multiple QC categories.

High-performance assay

The Infinium Global Screening Array with Enhanced PGx-48 v4.0 uses trusted Infinium assay chemistry to deliver a high-performance, accurate genotyping solution (Table 5). In addition to workflow improvements, the targeted gene amplification step increases performance for the detection of difficult targets, such as *CYP2D6*, and pseudogene disambiguation. The high signal-to-noise ratio of individual genotyping calls from the assay also provides access to copy number variant calling and star allele annotation via DRAGEN™ Array secondary analysis.

Infinium EX chemistry

Infinium EX chemistry is optimized for automation and compatible with Infinium Amplification System and Infinium Automated Pipetting System with ILASS. The automated workflow substantially reduces hands-on time and the potential for human error when using the Infinium Global Screening Array with Enhanced PGx-48 v4.0. The proven Infinium EX workflow provides results in as little as two to three days (Figure 7).

Table 5: Data performance and spacing

	Observed[a]	Product specification[b]
Call rate	99.6%	> 99.0% avg
Reproducibility	99.99%	> 99.90%
Log R deviation	0.13[c]	< 0.30 avg[d]

	Mean	Median	90th percentile[c]
Probe spacing (kb)	4.5	2.2	10.8

a. Values are derived from genotyping 1446 HapMap reference samples.
b. Excludes Y chromosome markers for female samples.
c. Based on results from high-quality GenTrain sample set.
d. Value expected for typical projects using standard Illumina protocols.

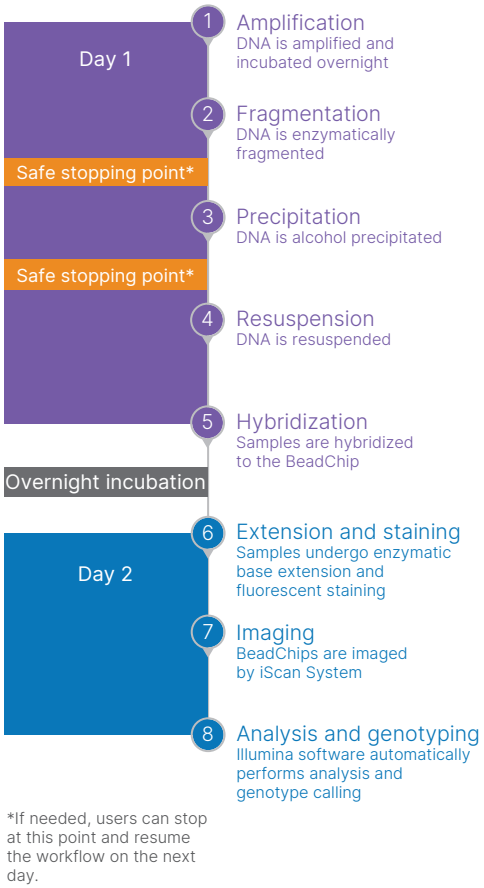

Figure 7: The Infinium EX 48-sample workflow provides a rapid workflow with minimal hands-on time.

Accurate and efficient secondary analysis

DRAGEN Array secondary analysis software is recommended for analysis of the Infinium Global Screening Array with Enhanced PGx-48 v4.0. DRAGEN Array is a powerful bioinformatics software that uses cutting-edge data analysis tools to maximize genomic insights. The software includes SNP analysis, PGx star allele and variant coverage across > 2400 targets for over 50 genes, hybrid and allele-specific copy number detection, PGx CNV coverage on seven target genes across nine target regions, and more.

DRAGEN Array secondary analysis is fast, generating accurate results in multiple output file formats for easy downstream analysis, including the capability to generate VCF files from Infinium array-based assays in as little as 10 seconds per sample and full PGx analysis results in approximately 20 seconds per sample.

DRAGEN Array secondary analysis has two deployment options. A local analysis option provides a command-line interface for granular control. No specialized DRAGEN server or FPGA hardware is required for the local installation solution. A cloud-based package with an intuitive graphical user interface is also available with the user-friendly BaseSpace™ Sequence Hub. The cloud-based option offers easy access and additional functionality, such as polygenic risk scoring for arrays.

Summary

The Infinium Global Screening Array with Enhanced PGx-48 v4.0 is a high-density genotyping assay that is ideal for high-throughput PGx and population genomics research. In addition, the ~646K backbone markers enable a range of population and clinical research studies. The Infinium EX chemistry delivers high-accuracy detection of challenging targets and fast turnaround times. When combined with the Infinium Automated Pipetting System with ILASS and the Infinium Amplification System, the Infinium Global Screening Array with Enhanced PGx-48 v4.0 offers a high-throughput option for labs looking to process large numbers of samples, with limited hands-on processing.

Learn more

Infinium Global Screening Array with Enhanced PGx-48 v4.0

Infinium Automation Option Packages

DRAGEN Array secondary analysis

Ordering information

Product	Catalog no.
Infinium Global Screening Array with Enhanced PGx-48 v4.0 kit (48 samples)	20065219
Infinium Global Screening Array with Enhanced PGx-48 v4.0 kit (96 samples)	20068346
Infinium Global Screening Array with Enhanced PGx-48 v4.0 kit (1152 samples)	20068347
DRAGEN Array Local - star allele annotation	20109885
DRAGEN Array Cloud - star allele annotation[a]	20109886

a. An Illumina Connected Analytics annual subscription is required for cloud analysis along with iCredits for data storage and analysis.

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.2350
2. Clinical Pharmacogenetics Implementation Consortium (CPIC). cpicpgx.org. Accessed September 13, 2024.
3. Gaedigk A, Casey ST, Whirl-Carrillo M, Miller NA, Klein TE. Pharmacogene Variation Consortium: A Global Resource and Repository for Pharmacogene Variation. *Clin Pharmacol Ther.* 2021;110(3):542-545. doi:10.1002/cpt.2321.
4. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing [published correction appears in Genet Med. 2017;19(5):606]. *Genet Med.* 2013;15(7):565-574.

5. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996-1006. doi:10.1101/gr.229102.
6. National Center for Biotechnology Information. Reference Sequence Blood Group Antigen Gene Mutation Database. https://ftp.ncbi.nlm.nih.gov/pub/mhc/rbc/Final%20Archive/. Accessed April 25, 2023.
7. National Center for Biotechnology Information. ClinVar Database. ncbi.nlm.nih.gov/clinvar. Accessed September 13, 2024.
8. Forbes SA, Bhamra G, Bamford S, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet*. 2008;Chapter 10:Unit-10.11. doi:10.1002/0471142905.hg1011s57
9. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25-29. doi:10.1038/75556.
10. 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. doi:10.1093/nar/gkt958
11. National Center for Biotechnology Information. eQTL Database. www.ncbi.nlm.nih.gov/projects/gap/eqtl/index.cgi. Accessed Seqptember 13, 2024.
12. Osier MV, Cheung KH, Kidd JR, Pakstis AJ, Miller PL, Kidd KK. ALFRED: an allele frequency database for diverse populations and DNA polymorphisms--an update. *Nucleic Acids Res*. 2001;29(1):317-319. doi:10.1093/nar/29.1.317.
13. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7
14. 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
15. Green RE, Krause J, Briggs AW, et al. A draft sequence of the Neandertal genome. *Science*. 2010;328(5979):710-722. doi:10.1126/science.1188021.
16. Illumina. TruSight Inherited Disease Sequencing Panel Data Sheet. support.illumina.com/content/dam/illumina-marketing/apac/china/documents/datasheet_trusight_inherited_disease.pdf. Published 2017. Accessed September 13, 2024.
17. National Human Genome Research Institute. genome.gov. Accessed August 23, 2024.
18. RefSeq - NCBI Reference Sequence Database. ncbi.nlm.nih.gov/refseq. Accessed August 23, 2024.
19. NCBI Genome Reference Consortium. Version GRCh38. ncbi.nlm.nih.gov/grc/human. Accessed August 23, 2024.

illumina®

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

M-GL-00711 v1.0