Human genotyping

Axiom PangenomiX Array

Ethnic diversity at your fingertips—the largest population coverage on a high-throughput array

The Applied Biosystems™ Axiom™ PangenomiX Array is a human genotyping research array, designed for whole-genome imputation with diverse and global population coverage. It is an essential research tool in human genomics, including applications such as genome-wide association studies (GWAS), population health initiatives, polygenic risk score development and implementation, and clinical trials in drug discovery. The Axiom PangenomiX Array can scan the whole genome from as little as 100 ng genomic DNA. It can enable identification of single-nucleotide polymorphism (SNP) targets, copy number variation (CNV) analysis, human leukocyte antigen (HLA) typing, and more, in a single, cost-effective assay with ready-to-use data analysis.

More than 800,000 markers were selected for high genomic coverage from the 1000 Genomes Project phase 3, yielding coverage for European, African, admixed American, East Asian, and South Asian populations. This means variants prevalent in different populations are accurately represented and accounted for, leading to more inclusive research outcomes.

In addition to markers selected specifically to maximize imputation power and ethnic diversity, additional markers were chosen from broadly referenced public databases, including ACMG 73, ClinVar, the NHGRI-EBI GWAS catalog, CPIC, PharmGKB, and PharmaADME, and can be directly genotyped for each sample.

The Axiom PangenomiX Array also offers CNV analysis for fixed genomic regions and *de novo* copy number discovery to detect copy number changes across the whole genome. The Axiom PangenomiX Plus Array enables genotyping of markers of important genes in difficult-to-genotype regions, such as the pharmacogene *CYP2D6*, and comes with Pharmacogenomic Translation Reports with general metabolizer status.

Coverage highlights

- Genome-wide association study (GWAS) imputation module with ~800,000 markers across all ancestral populations (Table 1)
- Evidence-based markers for relevant variants, including ClinVar and ACMG 73 for comprehensive gene coverage
- Pharmacogenomics (PGx) variants cited in Clinical Pharmacogenomics Implementation Consortium (CPIC) guidelines and Pharmacogenomics Knowledge Base (PharmGKB)
- Disease-related variants (Alzheimer's disease, cancer, cardiovascular and cardiometabolic diseases, diabetes, and neurological disorders)
- Blood phenotyping variants for blood typing of common and rare blood groups, bleeding disorders, and blood conditions such as sickle cell anemia
- HLA typing of 11 major histocompatibility complex (MHC) Class I and Class II loci
- COVID-19, host response, and immune-related markers

Table 1. Content of the Axiom PangenomiX Array compared with other catalog Axiom arrays.

Category	Description	Axiom PangenomiX Array	Axiom PMD* Research Array	Axiom PMR* Array	Axiom UK Biobank Array
GWAS module	Markers to maximize coverage in ancestral populations, especially in the 1–5% minor allele frequency (MAF) range, enabling cross-platform and cross-cohort metadata analysis	>800,000	>800,000	>800,000	>600,000
NHGRI-EBI GWAS catalog	Includes content covering the complete NHGRI catalog of published GWAS as of May 2020	>25,000	>16,000	>15,000	>8,000
ACMG 73	Markers from the list of 73 genes published by ACMG covering highly penetrant genetic disorders	>16,000	>15,000	>11,000	>9,000
ClinVar variants	Covers pathogenic or likely pathogenic associations from ClinVar archives (accessed December 2020)	>23,000	>24,000	>23,000	>7,500
High-value markers associated with inherited disorders	Markers in high-value genes such as *APOE* (Alzheimer's disease), *BRCA1/2* (breast cancer), *DMD* (Duchenne muscular dystrophy), and *CFTR* (cystic fibrosis) [1,2]	>5,300	>5,500	>2,000	>75
PGx (absorption, distribution, metabolism, excretion (ADME))	• Number of ADME genes covered	>1,100	>1,100	>660	>900
	• Markers from PharmGKB with known relevance to drug metabolism	>5,000	>5,000	>1,950	>2,400
	• Total number of markers in ADME genes	>92,000	>92,000	>49,000	>67,000
Blood phenotypes and disorders	Covers variants that are used to identify the common and rare blood group types	>2,700	>1,100	>550	>650
COVID-19 and host response	Variants in immune response genes and related pathways including published GWAS hits, expression quantitative trait loci (eQTLs), and nonsynonymous exome variants, as well as markers for HLA type imputation	>53,000	–	–	–
Immunity, inflammation, and HLA	Markers from HLA genes, killer cell immunoglobulin-like receptor (KIR) genes, and variants in research on specific autoimmune and inflammatory disorders such as ulcerative colitis, Crohn's disease, Graves' disease, Hashimoto's thyroiditis, and celiac disease	>13,000	>14,000	>10,400	>8,200
Cancer research risk variants	Cancer risk variants from the NHGRI-EBI GWAS catalog, various publications, and the OMIM® database; include variants associated with risks for colorectal [3], prostate [4], ovarian [5], lung [6], and gastric cancers, as well as typical blood cancers such as myeloma and lymphoma	>6,000	>10,000	>10,000	>6,500
Loss of function (LOF)	Markers to detect genetic changes that are predicted to completely disrupt the function of protein-coding genes, including rare and likely deleterious LOF alleles, predicted relevant variants, and common LOF variants in nonessential genes	>3,600**	>3,100**	>33,000	>30,000
eQTLs	eQTLs with MAF >0.01% to support mapping of functional noncoding variations to identify associations with gene transcription variability and differential gene expression	>17,000	>3,000	>16,000	>17,000
Neuropsychiatric conditions and lung function; CNV regions for developmental delay	Includes markers associated with increased risk for neurological conditions such as Alzheimer's disease and Parkinson's disease, schizophrenia, and autism [7,8]	>6,000	>230	>180	>2,300
Genetic testing	Markers associated with lifestyle health conditions, most notably obesity, alcohol and smoking addiction, skin conditions, and asthma and allergies	>1,400	>1,500	>1,200	>250
Y chromosome markers	Markers on the Y chromosome that are suitable for applications covering deep ancestry	>1,000	>440	>5	>800
Mitochondrial markers	Common mitochondrial DNA variants	>1,100	>700	>115	>350
Fingerprinting and sample tracking	Includes fingerprint SNPs used by the University of Washington and the Broad Institute of MIT and Harvard; these markers are shared among several major genotyping platforms to facilitate sample tracking	>300	>300	>300	>300
Total markers		>900,000	>900,000	>900,000	>800,000

* PMD: Precision Medicine Diversity, PMR: Precision Medicine Research.
** Using curated LOF variants from gnomAD [9,10].

GWAS grid

The Axiom PangenomiX Array includes over 800,000 markers in the GWAS module. Common variants are intelligently selected via a proprietary imputation-based marker selection strategy for genome-wide coverage in the five major ancestral populations (Table 2). This process allows access to a vast number of low-frequency markers (minor allele frequency (MAF) >1%) and common markers (MAF >5%) for any given population, through imputation. The intelligent, imputation-aware design helps ensure that the selection of markers offers the highest imputation accuracy across all ancestral populations. The autosomal markers in the GWAS grid are valuable in ascertaining the ethnic breakdown of individuals genotyped with the Axiom PangenomiX Array. Combined with the mitochondrial and Y chromosome markers, the Axiom PangenomiX Array is a powerful array for determining ancestry and migration patterns in genetic testing [11].

Table 2. Number of imputed markers with r^2 ≥0.8 and MAF >1%.

	Number of imputed variants	Imputation accuracy	
Population	**MAF >1%**	**MAF >1%**	**MAF >5%**
African (AFR)	15.2 M	0.90	0.92
Admixed American (AMR)	10.3 M	0.92	0.94
East Asian (EAS)	7.5 M	0.88	0.93
European (EUR)	8.8 M	0.91	0.95
South Asian (SAS)	9.0 M	0.90	0.94

ACMG 73 and ClinVar

The Axiom PangenomiX Array includes a set of markers covering 73 genes from guidelines published by the ACMG. The relevant variants in the ACMG 73 genes identify and manage risk for selected highly penetrant genetic disorders through established interventions aimed at preventing or significantly reducing morbidity and mortality. More than 16,000 variants are selected on the Axiom PangenomiX Array to interrogate the ACMG 73 genes. These markers are known to be of high importance, such as those in the *BRCA1/2*, *CFTR*, *DMD*, *RYR1/2*, *LDLR,* and *APOB* genes. Advanced probe set designs for variants with very high GC content (>78%) in flanking sequences allow for accurate genotyping of such complex variants.

The Axiom PangenomiX Array also includes variants from ClinVar archives curated for pathogenic or likely pathogenic significance. The module includes updated and well-annotated content from ClinVar retrieved on December 2020. Pathogenic variants often carry information about the penetrability associated with a disease. A list of some of the disease research categories and number of associated variants is shown in Table 3.

Table 3. Markers of the Axiom PangenomiX Array, classified into disease categories and important subcategories according to NHGRI, OMIM, and ClinVar databases using a search with Medical Subject Headings (MeSH)-controlled vocabulary.

Categories and subcategories	Number of markers
Cancer risk variants	**>13,000**
Myeloma	>70
Lung cancer	>400
Breast cancer	>1,800
Ovarian cancer	>1,500
Gastric cancer	>900
Leukemia	>3,000
Lymphoma	>700
Colorectal cancer	>2,200
Mental, behavioral, neurological, and neurodevelopmental risk variants	**>4,300**
Alzheimer's disease	>300
Parkinson's disease	>300
Schizophrenia	>700
Autism	>200
Inherited eye disease risk variants	**>3,700**
Macular degeneration	>500
Glaucoma	>150
Retinal dystrophy	>100
Retinitis pigmentosa	>400
Optic atrophy	>10
Autoimmune and inflammatory disease risk variants	**>1,150**
Celiac disease	>90
Crohn's disease	>400
Graves' disease	>35
Loss-of-function variants, autosomal inheritance	**>3,600**
Autosomal recessive	>300
• Fanconi anemia	>60
• Cystic fibrosis	>3
• Thalassemia	>3
Autosomal dominant	>340
• Familial hypercholesterolemia	>20
• Mitochondrial diseases	>10
Cardiovascular disease risk variants	**>8,500**
Respiration disorder risk variants	**>500**
Diabetes risk variants	**>1,500**
Musculoskeletal disease risk variants	**>5,900**

Pharmacogenomics research

The Axiom PangenomiX Array includes over 5,000 variants in 1,100 genes of known PGx relevance [12]. This evidence-based content allows researchers to gain valuable insight into an individual's ability to process drugs based upon high, moderate, and preliminary scientific evidence.

The pharmacogenomics content module includes:

- 2,000 markers in Very Important Pharmacogenes as identified by PharmGKB
- 300 markers associated with PharmGKB level 1A–2B annotations
- >550 reportable alleles mentioned in Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines

The Axiom PangenomiX Array provides genotype calling for variants in the PGx module. In addition, the Applied Biosystems™ Axiom™ PangenomiX Plus Array, used in conjunction with the Applied Biosystems™ Axiom™ 2.0 Plus Assay or Axiom™ Propel Plus Assay, unlocks over 100 additional markers associated with important haplotypes in genes including *CYP2D6, CYP2A6, CYP2B6, CYP2C19, CYP1A2, CYP2C8*, and *SULT1A1*. This unique assay opens up the ability to genotype these important PGx markers that are in highly homologous regions of the genome. Based on gene-specific amplification, the Axiom 2.0 Plus Assay overcomes limitations observed in other hybridization-based microarray technologies, making it the array of choice for PGx research.

Furthermore, the Axiom PangenomiX Plus Array comes with Pharmacogenomic Translation Reports that include star allele calling and phenotype prediction. Star allele calling is informed by copy number measurement for genes like *CYP2D6* (Table 4).

The Pharmacogenomics Translation Reports include:

- >110 genes
- >75 genes reporting haplotypes
- >1,000 reportable haplotypes
- >30 genes with phenotype predictions (e.g., intermediate metabolizer)

Table 4. The *CYP2D6* gene requires both determination of copy number state and SNP genotype calling for accurate star allele reporting. This table shows the workflow in Applied Biosystems™ Axiom™ Analysis Suite Software to get to the final diplotype/metabolizer status for a sample.

	Step	Output	Interpretation
1	CNV state detection	*CYP2D6*	There is one copy number state present for the *CYP2D6* gene in this sample
2	SNP genotype calling	*rs16947 'A'* *rs1135840 'G'*	Variant haploid genotypes
3	Star allele calling	Diplotype: *CYP2D6 *2/*5* Phenotype: intermediate metabolizer	A sample with one normal-function allele (*2) and one no-function allele (*5) is considered an "intermediate metabolizer" for *CYP2D6*

Blood phenotypes and disorders

The Axiom PangenomiX Array includes over 2,500 markers that can be used to type common (ABO, Rh, Kell) and rare blood groups to perform research in immunohematology, alloimmunization, maternal–fetal incompatibility, and hemoglobinopathies. The Axiom PangenomiX Array leverages past efforts in blood typing studies [13], and knowledge gained on blood typing from the UK Biobank study [14] and the Applied Biosystems™ Axiom™ UK Biobank v2 (UKBBv2) Array designed in collaboration with the Blood Transfusion Genomics Consortium [15]. The blood typing content on the array includes genetic research markers for anemia, bleeding disorders, and thalassemia. The large number of variants on the array helps enable investigation of rare blood types. The array supports detailed copy number calling in the *RHD* gene, and genotyping of variants in the highly homologous *RHD* and *RHCE* genes in the Rh blood group system.

High-resolution HLA typing

The HLA complex is the human version of the MHC. This complex includes genes responsible for immune function. Genetic variations in these genes affect immune responses, including transplant rejection and disease susceptibility.

The highly polymorphic nature of this region and the prevalence of pseudogenes create challenges in traditional genotyping methods. The Axiom PangenomiX Array combines the use of direct genotyping with advanced imputation methods over the extended MHC region, allowing accurate HLA typing from SNP genotype data.

Coverage of major HLA genes helps enable:

- Determination of the HLA types of 11 major MHC Class I and Class II loci (Figure 1) with two-digit and four-digit resolution
- Integration of HLA typing with genotyping data for insight into immune system variation associations with GWAS

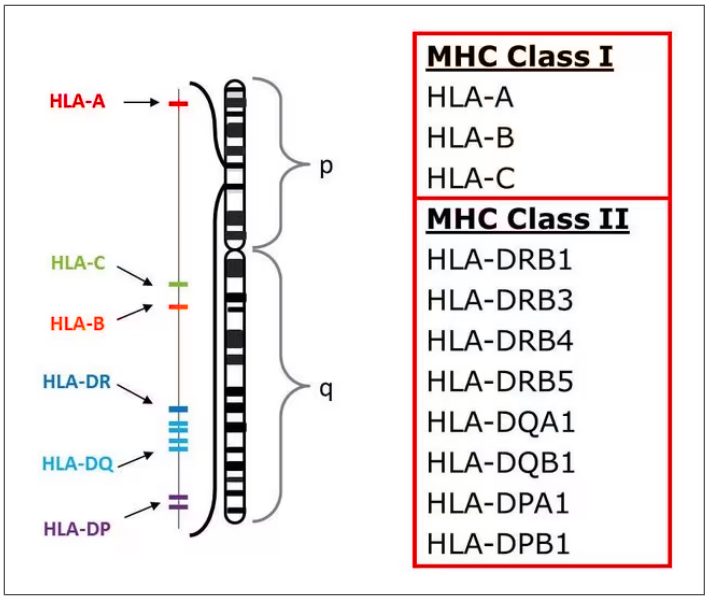

Figure 1. Example of HLA typing of 11 major MHC Class I and Class II loci.

Copy number analysis—fixed regions and discovery

CNVs are structural changes of DNA that include gains or losses, that account for significant variation among human genomes [16]. They have been associated with increased risk for various human diseases, abnormalities, conditions, and developmental disorders [17]. Therefore, in addition to genotyping SNPs and insertions or deletions (indels), the Axiom PangenomiX Array is designed to perform CNV analysis with integrated workflows using Axiom Analysis Suite Software 5.3.0 or higher. Two CNV analysis methods are enabled: (1) fixed-region analysis within genes of importance in PGx (*CYP2D6* (Figure 2), *CYP2A6, GSTM1, GSTT1, UGT2B17,* and *SULT1A1*) and blood typing (*RHD, RHCE, GYPA, GYPB, GYPC,* and *GYPE*), and (2) *de novo* discovery analysis to detect copy number changes across the whole genome. For more information, see assets.thermofisher.com/TFS-Assets/GSD/Technical-Notes/axiom-copy-number-analysis-tech-note.pdf

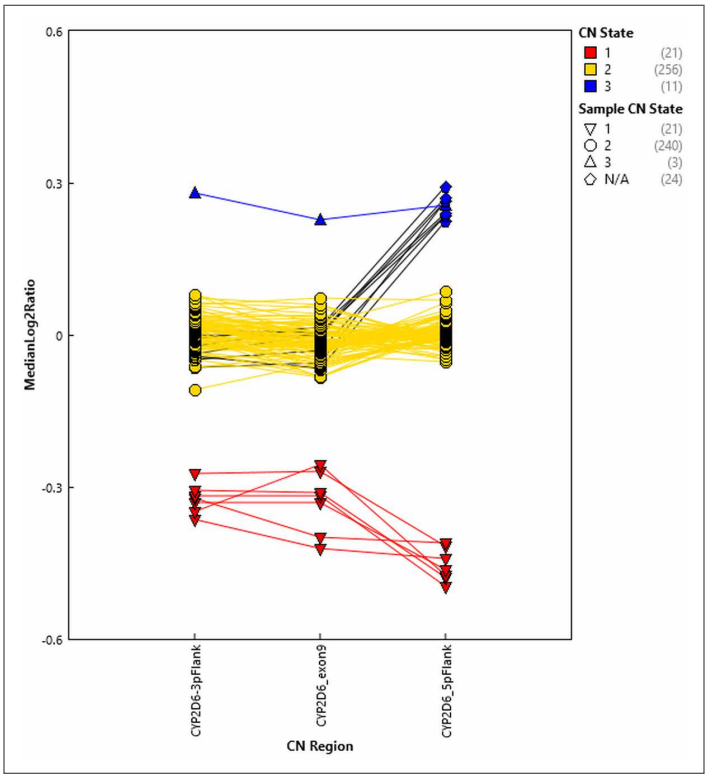

Figure 2. Copy number states (1, 2, or 3) detected for *CYP2D6* gene.

Applications in genetic testing

In recent years, genetic testing has evolved from using genomics to identify ancestry and lifestyle preferences to understanding risks for acquiring an inherited disease or condition. The successful completion of the UK Biobank study demonstrated that polygenic risk scores are relevant when considering disease risk.

The Axiom PangenomiX Array includes markers for identification of ancestry for diverse populations, PGx research, and rare blood typing. It includes additional genetic variants that have been identified in lifestyle phenotypes and traits such as diet, weight, metabolism, tastes, and more. The variants in the Axiom PangenomiX Array genetic testing module are shown in Table 5.

Table 5. Markers for genetic testing applications.

Lifestyle phenotypes and traits	Number of markers
Alcohol dependence and sensitivity	>500
Asthma	>700
Allergies	>1,000
Smoking and addiction	>1,000
Vitamin absorption	>100
Weight and obesity	>900

Assay and workflow

The Axiom PangenomiX Array utilizes the Axiom 2.0 Assay and workflow (Figure 3). The Axiom 2.0 workflow is a standard 3- to 4-day workflow, inclusive of whole-genome amplification through array hybridization, staining, and scanning on the Applied Biosystems™ GeneTitan™ MC Fast Scan Instrument.

The Axiom PangenomiX Plus Array utilizes the Axiom 2.0 Plus Assay and workflow with an extra step introduced for gene-specific amplification for PGx markers that are in highly homologous regions of the genome.

For higher-throughput customers that require more than 96 samples to be processed at a time, the Axiom Propel Assay with Fast Wash and Axiom Propel Plus Assay with Fast Wash (Figure 4) are available, allowing 4 x 96 samples or 8 x 96 samples at a time to be prepared.

The Axiom Analysis Suite Software automates data analysis and includes allele-calling algorithms and user-friendly visualization tools. The analysis workflow is described in the Axiom Genotyping Solution Data Analysis Guide (Pub. No. 702961) and enables high flexibility to help in finding the most informative content for each study.

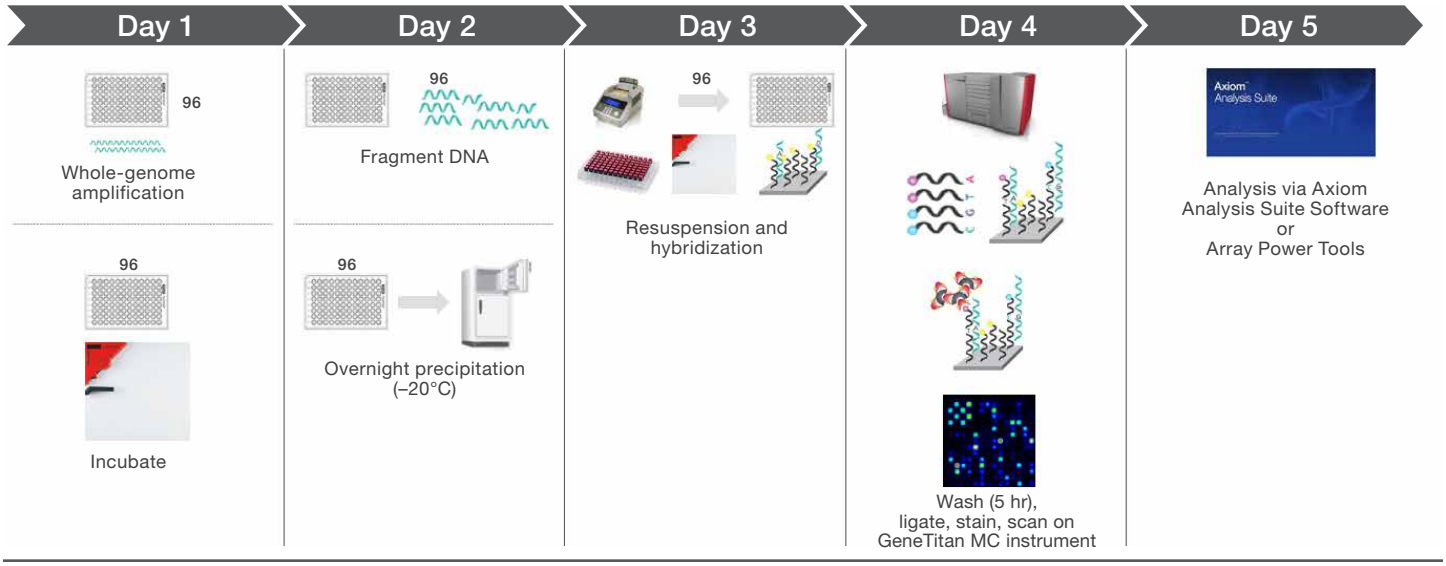

Figure 3. Axiom PangenomiX Array analysis with Axiom 2.0 Assay and workflow.

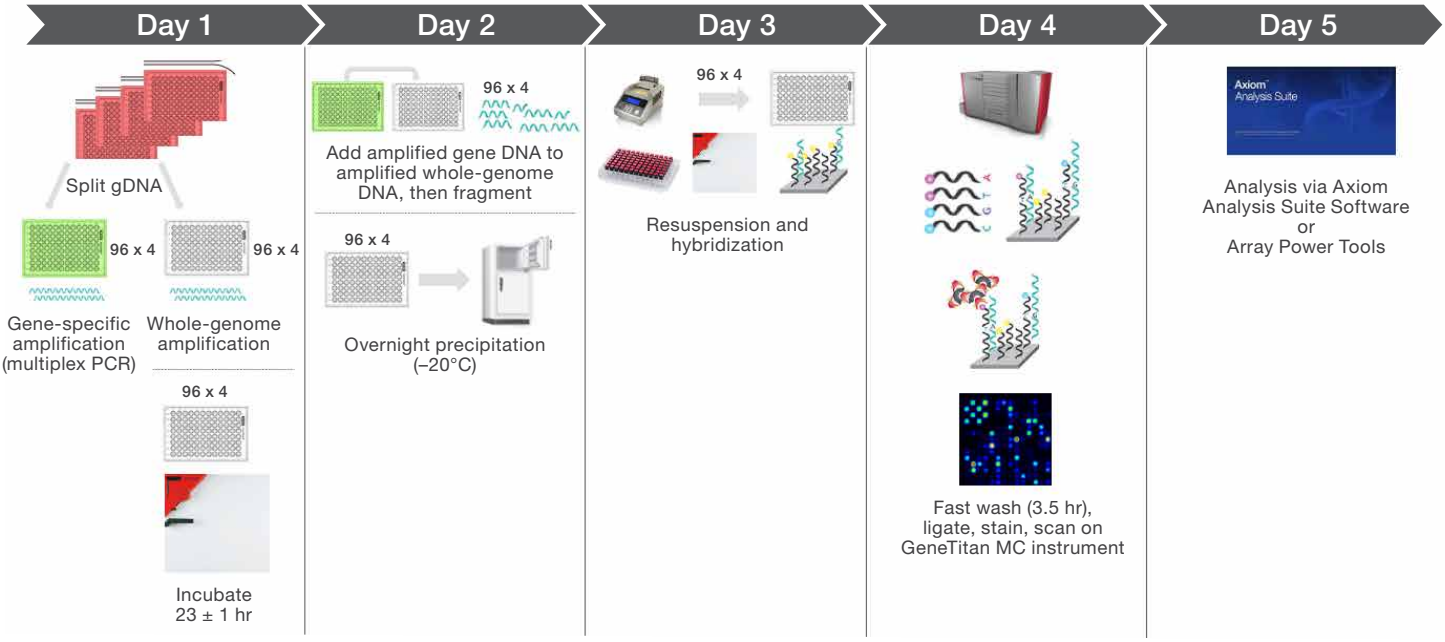

Figure 4. Axiom PangenomiX Plus Array analysis with Axiom 2.0 Plus Assay and workflow incorporating gene-specific amplification and a faster washing step.

Specifications

The Axiom PangenomiX Array's genotyping performance has been evaluated on 192 samples from the International HapMap Project using stringent quality control metrics that cover average sample call rate, sample concordance, and reproducibility. The samples were processed using the Axiom 2.0 Plus Assay to help ensure performance in regions of high sequence homology (e.g., markers in genes such as *CYP2D6*). Concordance and reproducibility were also evaluated on these markers (Table 6).

Table 6. Performance of the Axiom PangenomiX Array across 192 samples used with the Axiom 2.0 Plus Assay.

Metric	Specification	Performance
Number of samples	–	192
Sample pass rate	>95%	100.0%
Average call rate	≥99.5%	99.7%
Reproducibility	≥99.8%	99.9%
Average HapMap concordance	≥99.8%	99.9%
Average call rate of markers that require gene-specific amplification*	≥99.5%	99.9%
Concordance of markers that require gene-specific amplification*	≥99.5%	99.8%

* Two plates in the 96-array format.

The Axiom PangenomiX Array brings ethnic diversity to researchers' fingertips, accounting for global population coverage without compromising on directly genotyped variants, copy number analysis, and other important content such as HLA and blood types. This array will help enable researchers to identify potential population-specific associations for better understanding of complex diseases, leading to diverse genomic datasets and inclusive outcomes for the genomics community and predictive genomics applications.

References

1. Lambert JC et al. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 45:1452–1458.
2. Joshi PK et al. (2016) Variants near *CHRNA3/5* and *APOE* have age- and sex-related effects on human lifespan. *Nat Commun* 7:11174.
3. Al-Tassan NA et al. (2015) A new GWAS and meta-analysis with 1000 Genomes imputation identifies novel risk variants for colorectal cancer. *Sci Rep* 5:10442.
4. Hoffmann TJ et al. (2015) Imputation of the rare *HOXB13* G84E mutation and cancer risk in a large population-based cohort. *PLoS Genet* 11:e1004930.
5. Permuth JB et al. (2016) Exome genotyping arrays to identify rare and low frequency variants associated with epithelial ovarian cancer risk. *Hum Mol Genet* 25:3600–3612.
6. Kachuri L et al. (2016) Fine mapping of chromosome 5p15.33 based on a targeted deep sequencing and high density genotyping identifies novel lung cancer susceptibility loci. *Carcinogenesis* 37:96–105.
7. Gale CR et al. (2016) Pleiotropy between neuroticism and physical and mental health: findings from 108,038 men and women in UK Biobank. *Transl Psychiatry* 6:e791.
8. Smith DJ et al. (2016) Genome-wide analysis of over 106,000 individuals identifies 9 neuroticism-associated loci. *Mol Psychiatry* 21:749–757.
9. Singer-Berk M et al. (2020) Loss-of-function curations in gnomAD. **gnomad.broadinstitute.org/news/2020-10-loss-of-function-curations-in-gnomad/**.
10. Karczewski KJ et al. (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581:434–443.
11. Baughn LB et al. (2018) Differences in genomic abnormalities among African individuals with monoclonal gammopathies using calculated ancestry. *Blood Cancer J* 8:96.
12. Lemieux Perreault LP et al. (2018) Pharmacogenetic content of commercial genomewide genotyping arrays. *Future Med* 19:1159–1167.
13. Guo Y et al. (2018) Development and evaluation of a transfusion medicine genome wide genotyping array. *Transfusion* 59:101–111.
14. Bycroft C et al. (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562:203–209.
15. Gleadall NS et al. (2020) Development and validation of a universal blood donor genotyping platform: a multinational prospective study. *Blood Advances* 4(15):3495–3506.
16. Conrad D et al. (2010) Origins and functional impact of copy number variation in the human genome. *Nature* 464:704–712.
17. Kendall KM et al. (2017) Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK Biobank subjects. *Biol Psychiatry* 82:103–110.

Ordering information

Product	Quantity	Cat. No.
Axiom PangenomiX Array—combo kits (array, reagents, and consumables)		
Axiom PangenomiX Array Kit with Axiom 2.0 Assay	96 samples	952519
Axiom PangenomiX Array Kit with Axiom Propel 4X Assay and Fast Wash	4 x 96 samples	952528
Axiom PangenomiX Array Kit with Axiom Propel 8X Assay and Fast Wash	8 x 96 samples	952529
Axiom PangenomiX Plus Array—combo kits (array, reagents, mPCR reagents, and consumables)		
Axiom PangenomiX Plus Array Kit with Axiom 2.0 Plus Assay	96 samples	952521
Axiom PangenomiX Plus Array Kit with Axiom Propel Plus 4X Assay and Fast Wash	4 x 96 samples	952530
Axiom PangenomiX Plus Array Kit with Axiom Propel Plus 8X Assay and Fast Wash	8 x 96 samples	952531
Axiom PangenomiX Array—training kits (array, reagents, consumables, and DNA sample plate)		
Axiom PangenomiX Array Training Kit with Axiom 2.0 Assay	96 samples	952522
Axiom PangenomiX Array Training Kit with Axiom Propel 4X Assay and Fast Wash	4 x 96 samples	952416
Axiom PangenomiX Array Training Kit with Axiom Propel 8X Assay and Fast Wash	8 x 96 samples	952417
Axiom PangenomiX Plus Array—training kits (array, reagents, mPCR reagents, consumables, and DNA sample plate)		
Axiom PangenomiX Plus Array Training Kit with Axiom 2.0 Plus Assay	96 samples	952523
Axiom PangenomiX Plus Array Training Kit with Axiom Propel Plus 4X Assay and Fast Wash	4 x 96 samples	952524
Axiom PangenomiX Plus Array Training Kit with Axiom Propel Plus 8X Assay and Fast Wash	8 x 96 samples	952525

Learn more at **thermofisher.com/pangenomix**