

SNP genotyping using the Affymetrix® Axiom® Genome-Wide Pan-African (PanAFR) Array Set

Abstract

Genome-wide association studies (GWAS) in African populations pose unique challenges, given the extensive genetic heterogeneity that has been observed in these populations (Figures 1 and 2). Adding to the GWAS challenge is the increase in the number of known SNP and insertion/deletion (in/del) markers (driven by sequencing in novel populations). Population-optimized SNP arrays that include this novel sequencing content and overall high genomic coverage offer great power to detect disease-associated markers. To this end, the Axiom® Genome-Wide PanAFR Array Set was designed to maximize coverage of novel common and rare variants in populations of Yoruba (YRI) ancestry. We have also demonstrated the coverage of this array in the following HapMap populations: Africans from the Southwestern USA (ASW), Maasai (MKK), Luhya (LWK), European (CEU), and Asian (Chinese + Japanese, ASI) (Figure 3).

The Axiom® Genome-Wide PanAFR Array Set contains ~2.2 million markers (SNPs and insertion/deletion) obtained from HapMap, dbSNP, the 1000 Genomes Project, and the Southern African Genomes Project (Figure 5). In addition to high genome-wide coverage, these markers were also selected for coverage of the following categories: chromosomes X and Y, mitochondria, coding regions, recombination hotspots, ADME, miRNA, and disease-associated regions. The array set has been tested using manual and automated workflows with multiple sample types such as cell line, saliva, and blood-derived DNA. Assay performance was evaluated and shown to have an average sample call rate greater than 99.0 percent; trio concordance and average sample concordance to independent DNA genotype information (HapMap) were greater than 99.5 percent, and intra- and inter-run reproducibility were greater than 99.8 percent (Figure 6).

Figure 1: Marker selection strategy for the GWAS challenge

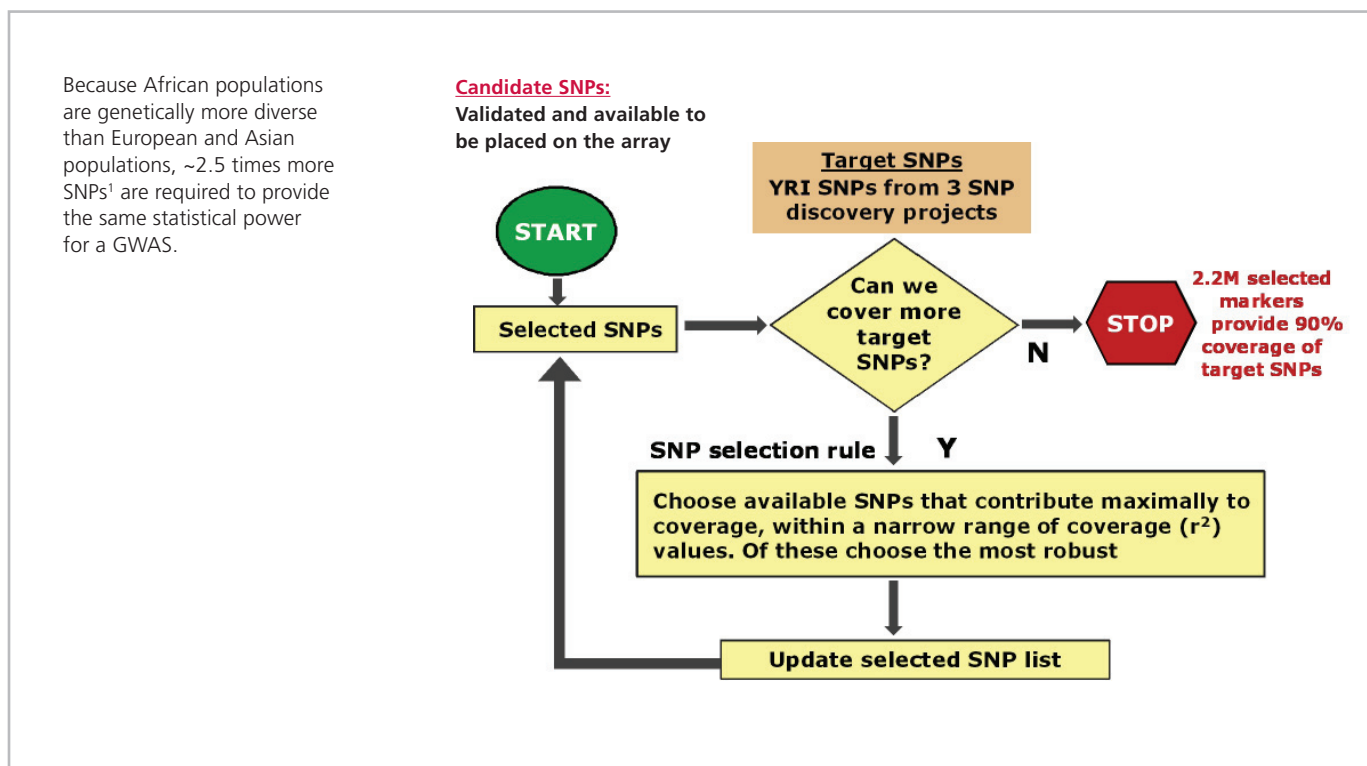
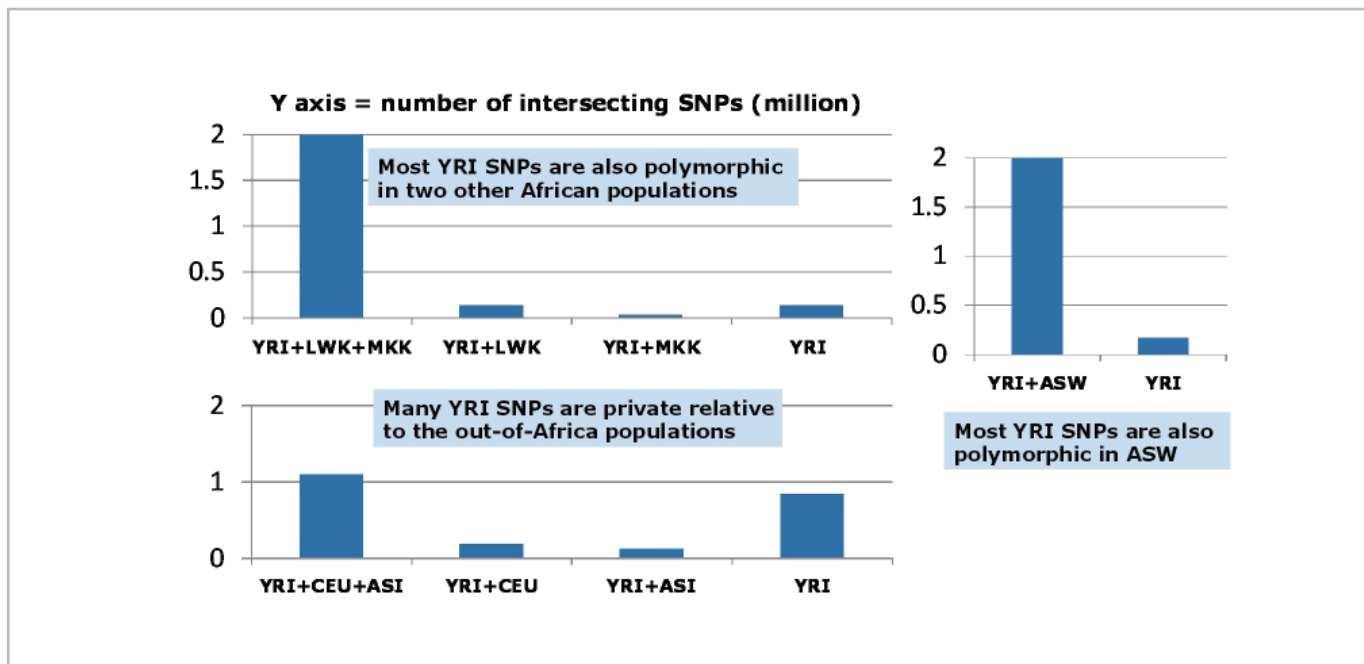
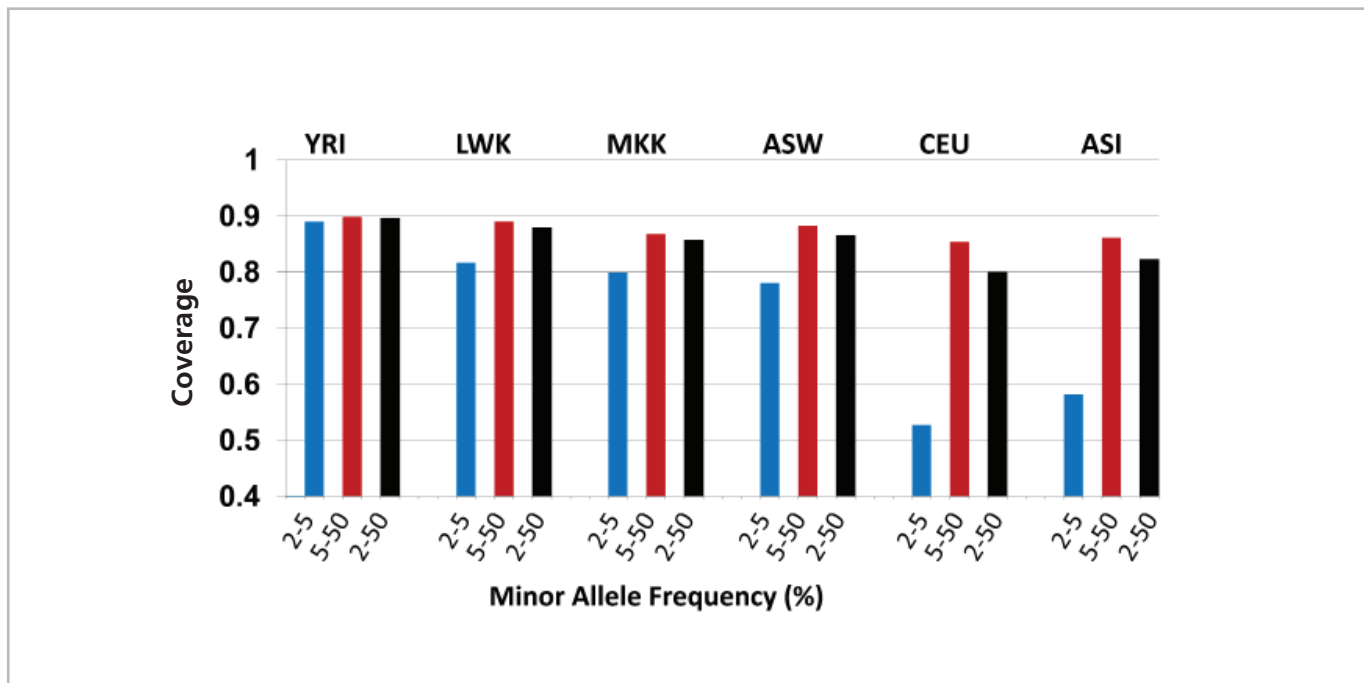


Figure 2: Intersection of YRI SNPs with various populations



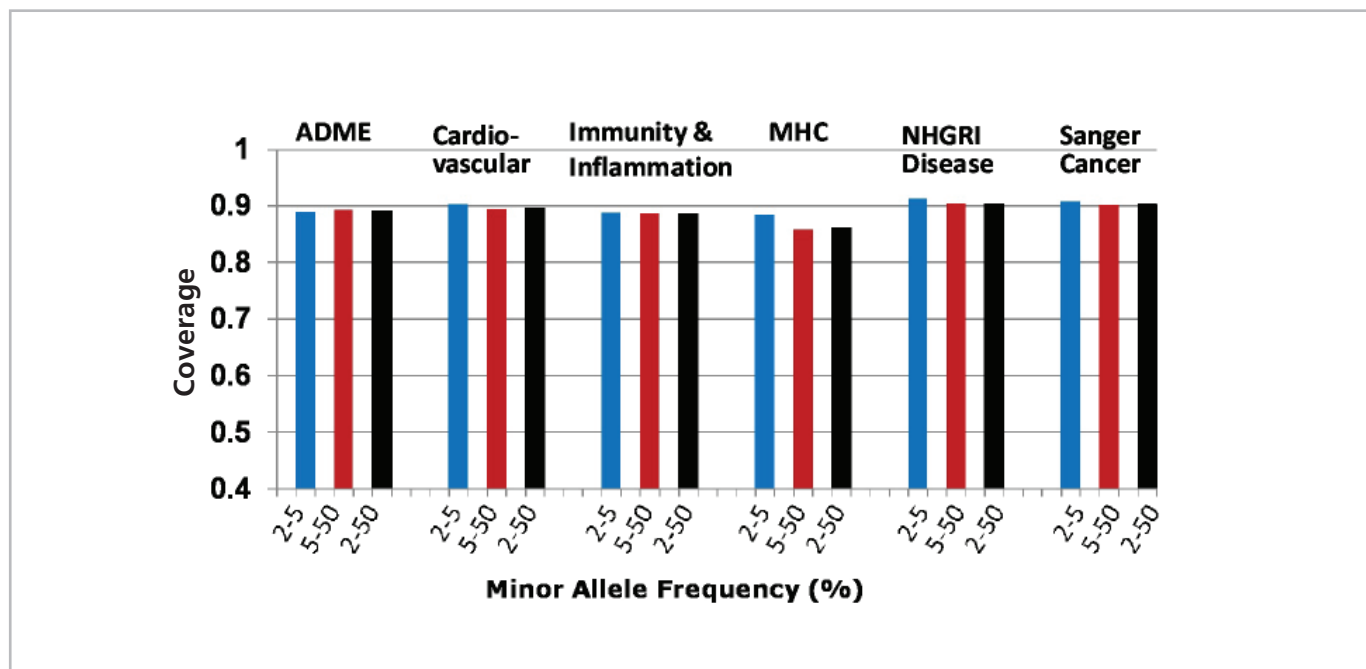
Intersection of arrayed markers polymorphic in various populations. Counts of polymorphic SNPs were determined by genotyping the ASI, ASW, CEU, LWK, MKK, and YRI HapMap samples with the Genome-Wide PanAFR Array Set.

Figure 3: Overall genetic coverage



Genetic coverage: Single-marker coverage is the proportion of SNPs and in/dels with $r^2 \geq 0.8$ to a marker on the Axiom Genome-Wide PanAFR Array.

Figure 4: Coverage of markers in biological categories



Coverage of biologically relevant SNPs: Single-marker coverage ($r^2 \geq 0.8$ for YRI) of SNPs and in/dels associated with selected genes and regions.

Figure 5: Markers for the Axiom® Genome-Wide PanAFR Array are from three discovery projects

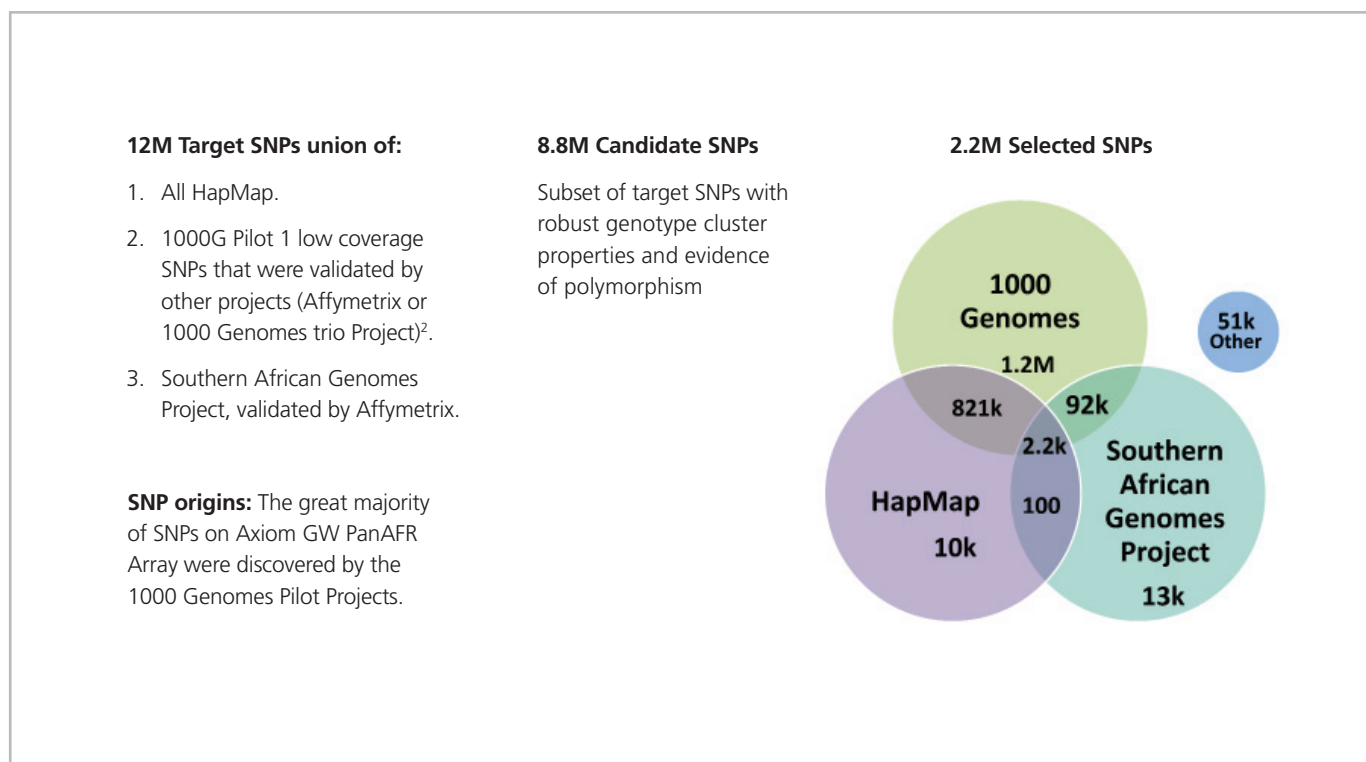
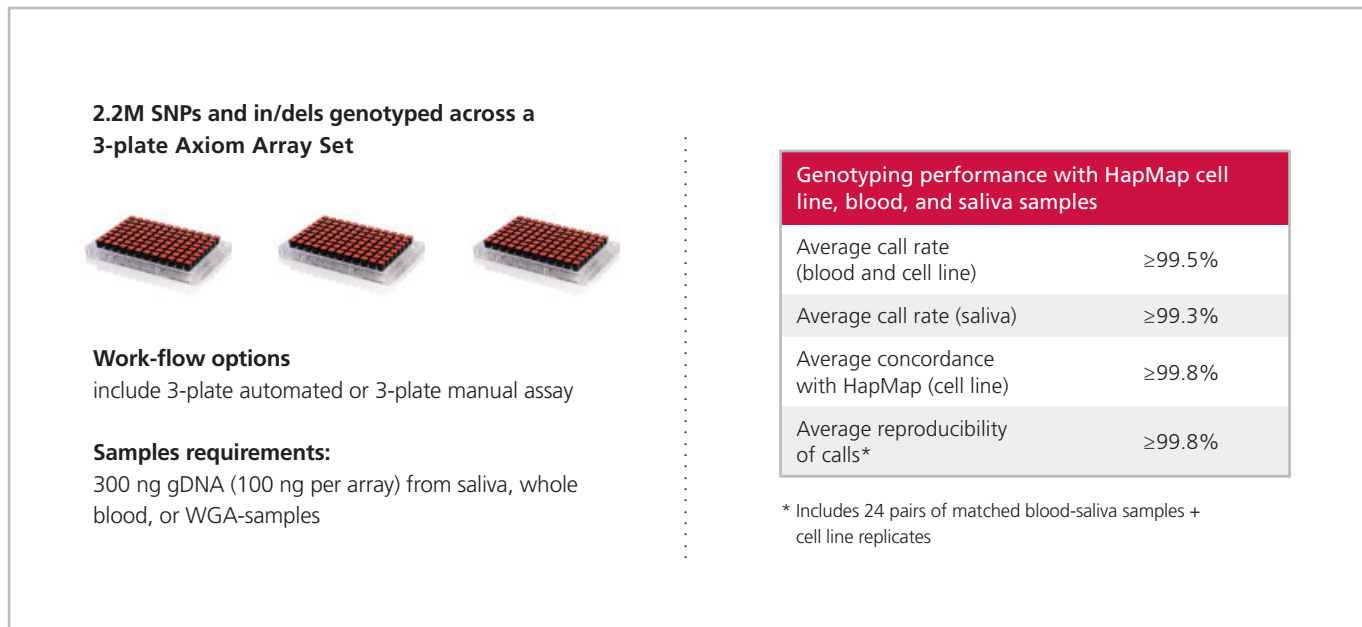


Figure 6: Axiom® Genome-Wide PanAFR Array Set specifications



References:

1. Teo, *et al.*, Methodological challenges of genome-wide association analysis in Africa. *Nature Reviews Genetics* **11** (2010).
2. The 1000 Genomes Project Consortium, A map of human genome variation from population-scale sequencing. *Nature* **467** (2010).

Affymetrix, Inc. Tel: +1-888-362-2447 ■ Affymetrix UK Ltd. Tel: +44-(0)-1628-552550 ■ Affymetrix Japan K.K. Tel: +81-(0)3-6430-4020
Panomics Solutions Tel: +1-877-726-6642 www.panomics.com ■ USB Products Tel: +1-800-321-9322 www.usb.affymetrix.com

www.affymetrix.com Please visit our website for international distributor contact information.

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

P/N DNA01074 Rev. 1

©Affymetrix, Inc. All rights reserved. Affymetrix®, Axiom®, Command Console®, CytoScan™, DMET™, GeneAtlas®, GeneChip®, GeneChip-compatible™, GeneTitan®, Genotyping Console™, myDesign™, NetAffx®, OncoScan™, Powered by Affymetrix™, PrimeView™, Procarta®, and QuantiGene® are trademarks or registered trademarks of Affymetrix, Inc. All other trademarks are the property of their respective owners.

Products may be covered by one or more of the following patents: U.S. Patent Nos. 5,445,934; 5,744,305; 5,945,334; 6,140,044; 6,399,365; 6,420,169; 6,551,817; 6,733,977; 7,629,164; 7,790,389 and D430,024 and other U.S. or foreign patents. Products are manufactured and sold under license from OGT under 5,700,637 and 6,054,270.