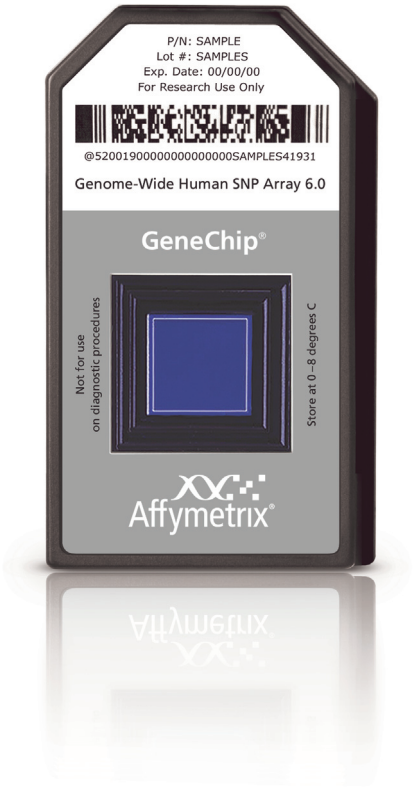
Data Sheet

Genome-Wide Human SNP Array 6.0



The Genome-Wide Human SNP Array 6.0 features more than 1.8 million markers of genetic variation, including single nucleotide polymorphisms (SNPs) as well as probes for the detection of copy number variation. The SNP Array 6.0 allows researchers to perform association studies with large sample sizes in both initial scan and replication phases, thereby significantly increasing the overall genetic power of their studies.

The SNP Array 6.0 demonstrates industry-leading performance and represents more genetic variation on a single array than any other product, providing maximum panel power and the highest physical coverage of the genome (median marker spacing of 680 bases).

Introduction

The Genome-Wide Human SNP Array 6.0 contains more than 906,600 single nucleotide polymorphisms (SNPs) and more than 946,000 probes for the detection of copy number variation. SNPs on the array are present on 200 to 1,100 base pairs (bp) Nsp I or Sty I digested fragments in the human genome, and are amplified using the Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0. This assay, which is also compatible with the SNP Array 5.0, now combines the Nsp and Sty fractions previously assayed on two separate arrays.

SNPs on the SNP Array 6.0 were screened in more than 500 distinct samples, including 270 HapMap samples and separate diversity samples. Approximately 482,000 SNPs are derived from the previous-generation Mapping 500K and SNP 5.0 Arrays. The remaining 424,000 SNPs include tag SNP markers derived from the International HapMap Project. These novel markers have better representation of SNPs on chromosomes X and Y, mitochondrial SNPs, SNPs in recombination hotspots, and new SNPs added to the dbSNP database after completion of the GeneChip® Human Mapping 500K Array Set.

This array contains a total of 946,000 non-polymorphic copy number probes. These probes—744,000 originally selected for their spacing and 202,000 selected based on known copy number changes reported in the Toronto Database of Genomic Variants (DGV)—enable you to detect de novo copy number changes and perform association studies by genotyping both SNP and known copy number polymorphism (CNP) loci (as reported by McCarroll, *et al*.).

The median inter-marker distance over all 1.8 million SNP and copy number markers combined is less than 700 bases.

The whole-genome sampling assay

The Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0 (P/N 901152, 901015, and 901192) is validated for use in conjunction with the SNP Array 6.0. Total genomic DNA (500 ng) is digested with Nsp I and Sty I restriction enzymes and ligated to adaptors that recognize the cohesive 4 bp overhangs. All fragments resulting from restriction enzyme digestion, regardless of size, are substrates for adaptor ligation. A generic primer that recognizes the adaptor sequence is used to amplify adaptor-ligated DNA fragments. PCR conditions have been optimized to preferentially amplify fragments in the 200 to 1,100 bp size range. PCR amplification products for each restriction enzyme digest are combined and purified using polystyrene beads. The amplified DNA is then fragmented, labeled, and hybridized to a SNP Array 6.0 (see Figure 1).

The Genome-Wide Human SNP Nsp/Sty Assay Kit contains validated and qualified reagents for the most critical steps in the assay. This includes the PCR primer and adaptors, reagents to fragment and label the PCR products, and control reagents. Manual processing kits are available for either 50 or 100 reactions. An automated assay kit (for processing 96 reactions) is also available (see ordering information). Whole-genome-amplified material prepared by the Qiagen REPLI-g® kits may also be used as the starting material for the Genome-Wide Human SNP Assay Kit.

Performance data

To validate the performance of the SNP Array 6.0, Affymetrix tested 270 samples from the International HapMap Project. In addition, two external sites tested a plate of 44 HapMap DNAs, which includes 30 unique samples, 10 trios, and five samples with multiple replicates.

The arrays that passed the QC call rate threshold were analyzed using the Birdseed algorithm at the default setting of 0.1. The average call rate for each set was greater than 99 percent, and the concordance with HapMap genotypes was observed to be greater than or equal to 99.7 percent. For the 10 trios, the Mendelian inheritance consistency was found to be greater than 99.9 percent. Reproducibility was measured at 99.9 percent (see results in Table 1).

Genotype calls using Genotyping Console™ Software

When the Genome-Wide Human SNP Array 6.0 is used in conjunction with Genotyping Console™ Software, the following applications are enabled:

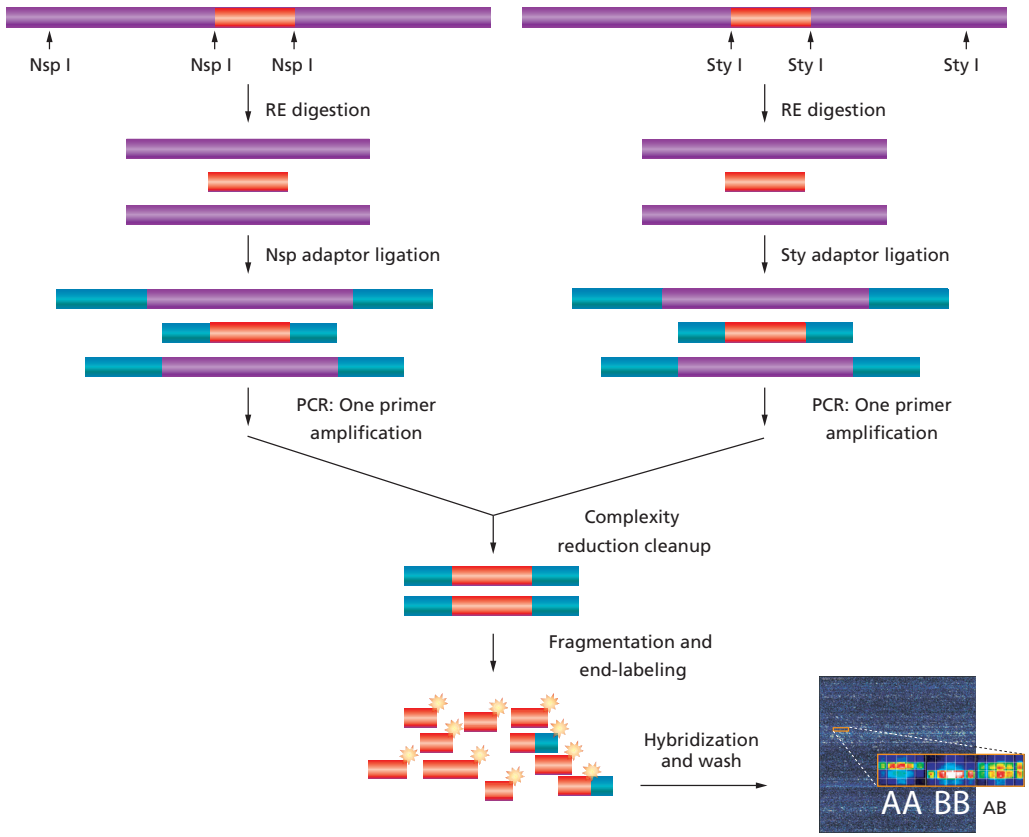
- SNP genotyping for association studies
- CNP genotyping (McCarroll, *et al.*) for association studies
- De novo CNV and LOH detection for association studies and cytogenetics research

Table 1: Data analyzed with Birdseed (0.1).

	270 HapMap	Site 1	Site 2
Call rate	99.8%	99.7%	99.7%
HapMap concordance	99.8%	99.7%	99.8%
Mendelian consistency	99.97%	99.95%	99.96%
Reproducibility	NA	99.9%	99.9%
SNP completeness*	99.9%	99.7%	99.8%

*SNP completeness is defined as the proportion of SNPs with per-SNP call rate greater than 85 percent.

Figure 1: Overview of the Genome-Wide Human SNP Assay 5.0/6.0.



Genotyping Console™ Software utilizes strong quality control metrics to streamline genotyping. QC metrics include the Contrast Quality Control (CQC) metric to predict genotyping performance and the Median Absolute Pairwise Difference (MAPD) value for copy number measurements. Flexible visualization and filtering tools allow you to generate customized PDF reports for downstream statistical analysis and bridge CNP and SNP genotyping for a truly genome-wide association study.

Refer to the *Genome-Wide Human SNP Nsp/Sty Assay 5.0 or 6.0 User Guide* (P/N 702419-2, P/N 702504-3) for details on the QC call rate thresholds, as well as procedures on DNA target preparation, target hybridization, fluidics setup, array scanning, and data analysis.

Fluidics protocol required
GenomeWideSNP6_450

Library files required
GenomeWideSNP_6

Library files contain information about probe array design layout, probe use and content, scanning and analysis parameters, and other characteristics.

These files are unique for each probe array type. Library files are available from the Affymetrix website at www.affymetrix.com/support/technical/libraryfilesmain.affx. Affymetrix products can be purchased directly from Affymetrix in the United States and many European countries. For all other territories, refer to our list of distribution partners located at www.affymetrix.com/site/contact/index.affx.

References

Rabbee, N., *et al.* A genotype calling algorithm for Affymetrix SNP arrays. *Bioinformatics* 22:7-12 (2006).

Affymetrix, Inc. BRLMM: An Improved Genotype Calling Method for the Mapping 500K Array Set.

Matsuzaki, H., *et al.* Genotyping over 100,000 SNPs on a Pair of Oligonucleotide Arrays. *Nature Methods* 1:109-111 (2004).

Frayling, T. M., *et al.* A Common Variant in the FTO Gene is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science* (ePub 2007).

McCarroll, S. A., *et al.* Integrated detection and population—genetic analysis of SNPs and copy number variation. *Nature Genetics* **40**(10):1166-1174 (2008).

Product information

Number of SNPs on the array	906,600
Number of non-polymorphic probes for copy number detection	946,000
Number of arrays	1
DNA required	500 ng
Average Birdseed call rate (0.1)	>99 percent*
Average minor allele frequency (MAF)	19.6% in HapMap Caucasians 18.2% in HapMap Asians 20.6% in HapMap Africans
Average heterozygosity	26.7% in HapMap Caucasians 24.6% in HapMap Asians 28.5% in HapMap Africans
PCR primers	1 per sample
Instrumentation	GeneChip® Scanner 3000 7G with AutoLoader
Throughput	>40 million genotypes per day, per scanner with three GeneChip® Fluidics Station 450s

* Average scan from four independent studies

Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0 components

Adaptor, Nsp I or Adaptor, Sty I	Two annealed oligonucleotides specific for ligation to the Nsp I or Sty I restriction site
PCR Primer 002	PCR primer to amplify ligated genomic DNA
Reference Genomic DNA, 103	Human genomic DNA control, with consensus genotypes
GeneChip® Fragmentation Reagent	DNase I enzyme, formulated to fragment purified PCR amplicons
10X Fragmentation Buffer	Buffer for fragmentation reaction
GeneChip® DNA Labeling Reagent (30mM)	Proprietary biotin-labeled reagent for end-labeling fragmented PCR amplicons
Terminal Deoxynucleotidyl Transferase	Enzyme used to end-label fragmented PCR amplicons with the GeneChip® DNA Labeling Reagent
5X Terminal Deoxynucleotidyl Transferase Buffer	Buffer for labeling reaction
Oligo Control Reagent, 0100	Mixture of five biotin-labeled oligonucleotides, which hybridize to control regions (gridding and array controls) on the SNP Array 5.0/6.0

Ordering information

Part Number	Description
Genome-Wide Human SNP Array 6.0	
901153	Contains 50 arrays
901150	Contains 100 arrays
Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0	
901152	Sufficient for 50 reactions
901015	Sufficient for 100 reactions
901192	Sufficient for 96 reactions, automated processing

Affymetrix, Inc.
3420 Central Expressway
Santa Clara, CA 95051 USA
Tel: 1-888-DNA-CHIP (1-888-362-2447)
Fax: 1-408-731-5441
sales@affymetrix.com
support@affymetrix.com

Affymetrix UK Ltd.
Voyager, Mercury Park,
Wycombe Lane, Wooburn Green,
High Wycombe HP10 0HH
United Kingdom
Tel: +44 (0) 1628 552550
Fax: +44 (0) 1628 552585
saleseurope@affymetrix.com
supporteurope@affymetrix.com

Affymetrix Japan K.K.
Mita NN Bldg., 16 F
4-1-23 Shiba, Minato-ku,
Tokyo 108-0014 Japan
Tel: +81-(0)3-5730-8200
Fax: +81-(0)3-5730-8201
salesjapan@affymetrix.com
supportjapan@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 702509 Rev. 2