AFFYMETRIX®

Data Sheet

Affymetrix® Genome-Wide Human SNP Array 5.0

The new single-chip Affymetrix Genome-Wide Human SNP Array 5.0 features single nucleotide polymorphisms (SNPs) from the original two-chip Mapping 500K Array Set, as well as additional non-polymorphic probes that can measure other genetic differences, such as copy number variation. The SNP 5.0 Array gives researchers a significant increase in information above the original 500K Array Set, while reducing the array processing time.

Introduction

The new single-chip Affymetrix Genome-Wide Human SNP Array 5.0 contains all 500,568 single nucleotide polymorphisms (SNPs) from the two-array Mapping 500K Array Set as well as an additional 420,000 non-polymorphic probes that can measure other genetic differences, such as copy number variation. SNPs on the array are present on 200 to 1,100 base pair (bp) Nsp I or Sty I digested fragments in the human genome, and are amplified using the fifth generation of the Whole-genome Sampling Assay (WGSA). This assay now combines the Nsp and Sty fractions previously assayed on two separate arrays. Using the current version of the Affymetrix® Genotyping Console, a set of 440,794 SNPs on the array exhibit the performance capabilities detailed in this data sheet.

One hundred thousand non-polymorphic probes were chosen to cover 2,000 germline copy number variants (CNV) identified in the UCSC Genome Browser database with 50 probes each. The other 320,000 were chosen to give even spacing across the genome, concentrating on areas that were not already represented by SNPs. Current Affymetrix software does not support analysis of the non-polymorphic probes.

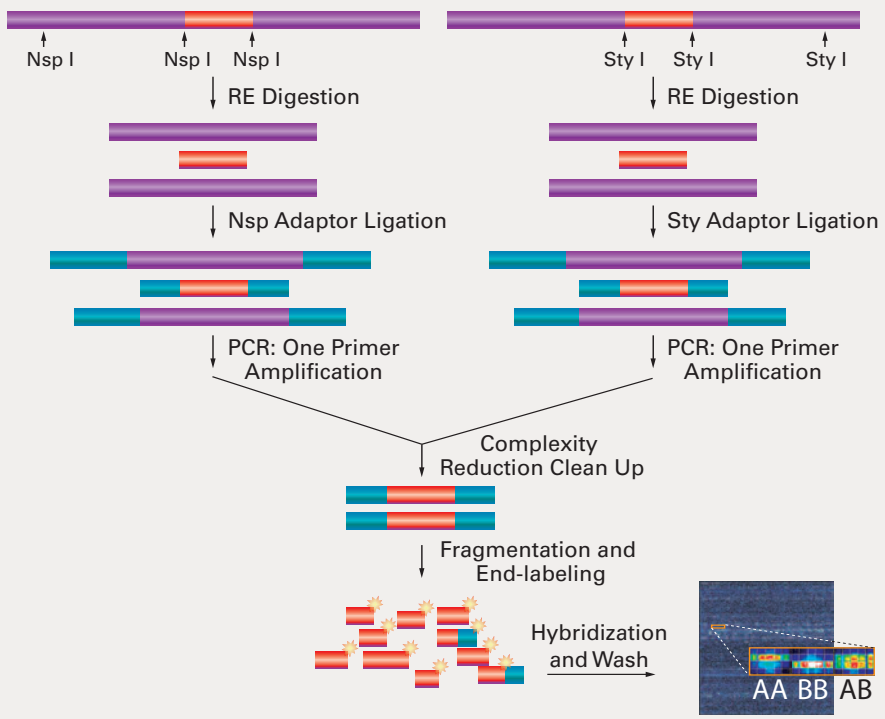
Please check the Affymetrix web site for additional updates to the genotyping calling algorithm and updates to support analysis of non-polymorphic probes.

The Whole-genome Sampling Assay

The Affymetrix Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0 (P/N 901152, 901015) was developed and validated for use in conjunction with the Genome-Wide Human SNP Array 5.0 (P/N 901167, 901071). Briefly, 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 Genome-Wide Human SNP Array 5.0.

The Affymetrix Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0 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 several control reagents. Kits are available for either 50 or 100 reactions (refer to Ordering Information).

Figure 1: The fifth-generation Whole-genome Sampling Assay.



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 5.0/6.0.

Genotype Calls Using Affymetrix® Genotyping Console

The Genome-Wide Human SNP Array 5.0 is used in conjunction with the Affymetrix® Genotyping Console. Genotyping Console implements two algorithms; the Dynamic Model algorithm (DM) generates a quality control (QC) call rate for each array by testing 3,022 SNPs specifically chosen for evaluating data quality, and the BRLMM-P algorithm uses data from multiple arrays to make genotype calls[1, 2]. BRLMM-P is an updated version of the previous BRLMM (Bayesian Robust Linear Model with Mahalanobis distance classifier) algorithm that only analyzes perfect match probes.

The QC call rate for an individual array (at a DM confidence threshold of 0.33) should be used to determine whether a sample should be repeated or used for downstream analysis. If the sample passes the QC call rate, it is expected to have a minimum BRLMM-P call rate of 97 percent. The user may adjust the default confidence score of BRLMM-P to allow genotyping with either greater accuracy or higher call rates, depending on what is needed for the application of interest[3].

Using the current version of the algorithm BRLMM-P, a set of 440,794 SNPs exhibit the performance capabilities detailed below. Future versions of the algorithm have the potential to add more SNPs to this default set of 440,794 SNPs. In addition, advanced users can analyze the full set of 500,568 SNPs using an alternative library file (CDF) that reveals all of the SNP content on the array. This advanced workflow cannot be conducted using Genotyping Console; it can only be performed with the command-line "snp5-probeset-genotype" software.

The details of the Genotyping Console data analysis workflow using BRLMM-P, including step-by-step instructions to install and run the Genotyping Console are described in the *Affymetrix® Genotyping Console Manual*. Further details can be found in the snp5_probeset_genotypemanual (www.affymetrix.com/products/software/specific/genotyping_console_software.affx).

Performance Data

To test the performance of the SNP Array 5.0, Affymetrix and the Broad Institute jointly ran the 270 samples from the International HapMap Project. In addition, two external sites and one internal validation group ran 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 BRLMM-P algorithm at the default setting of 0.05. 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 99.5 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.

Data analyzed with BRLMM-P (0.05).

	270 HapMap	Site 1	Site 2	Internal
Call Rate	99.71	99.55	99.37	99.63
HapMap Concordance	99.69	99.67	99.56	99.69
Mendelian Consistency	99.96	99.95	99.94	99.96
Reproducibility	NA	99.9	99.9	99.9

Refer to the *Affymetrix® Genome-Wide Human SNP Nsp/Sty Assay 5.0 Manual* (P/N 702419) 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
GenomeWideSNP5v1_450

LIBRARY FILES REQUIRED
GenomeWideSNP_5

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

These files are unique for each probe array type. Library files are available from the Affymetrix web site at: www.affymetrix.com/support/technical/libraryfilesmain.affx.

Two alternate SNP List (CDF) files are available. These files identify which SNPs are available for downstream analysis.

GenomeWideSNP_5.cdf is the default set of 440,794 SNPs that are accessible via the Genotyping Console (BRLMM-P algorithm).

GenomeWideSNP_5.Full.cdf is for advanced users who wish to look at all SNPs from the previous-generation Mapping 500K Array Set. This CDF file can only be accessed from the command-line tool, snp5-probeset-genotype, which is part of the Affymetrix Power Tools distribution. The full CDF file includes SNPs that may have lower per-SNP accuracy or call rates. It is expected that the performance of some of these SNPs will improve with different or future algorithms.

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

1. Rabbee N., *et al.* A genotype calling algorithm for Affymetrix SNP arrays. *Bioinformatics* **22**:7-12 (2006).
2. Affymetrix White Paper, BRLMM: An Improved Genotype Calling Method for the Mapping 500K Array Set.
3. Matsuzaki H., *et al.* Genotyping over 100,000 SNPs on a Pair of Oligonucleotide Arrays. *Nature Methods* **1**:109-111 (2004).
4. Papassotiropoulos A., *et al.* Common Kibra Alleles Are Associated with Human Memory Performance. *Science* **314**(5798):475-8 (2006).

Product Information

Number of SNPs on the array	500,568
Number of SNPs accessible using BAT 2.0	440,794
Number of Arrays	1
DNA Required	500ng
Expected minimum BRLMM-P Call Rate (0.05)	≥97 percent
Average Minor Allele Frequency (MAF)	0.22
Average Heterozygosity	0.31
PCR Primers	1 per sample
Instrumentation	GeneChip® Scanner 3000 7G with AutoLoader
Throughput	>21 million genotypes per day per scanner with three GeneChip® Fluidics Station 450s

Affymetrix® Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0 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

Ordering Information

Affymetrix® Genome-Wide Human SNP Array 5.0

901167 *Contains 50 arrays*
901071 *Contains 100 arrays*

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

901152 *Sufficient for 50 reactions*
901015 *Sufficient for 100 reactions*

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
UK and Others Tel: +44 (0) 1628 552550
France Tel: 0800919505
Germany Tel: 01803001334
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 web site for international distributor contact information.

For research use only. Not for use in diagnostic procedures.

Part No. 702087 Rev. 5