Axiom® Soybean Genotyping Array
Whole-genome high-density (HD) genotyping for soybean

Axiom® Soybean Genotyping Array (also called the Axiom® SoyaSNP 96-Array Plate) includes 180,961 markers for genotyping cultivated and wild varieties of soybean. The array was designed through Affymetrix’ Expert Design Program in collaboration with the Korea Research Institute of Bioscience and Biotechnology (KIRIBB), the Rural Development Administration, and the National Institute of Crop Science. The SNP sequences were provided by Dr. Soon-Chun Jeong, Namshin Kim, and Dr. Jung-Kyung Moon. The array was described by Dr. Soon-Chun Jeong from KIRIBB at the Plant and Animal Genome Conference, Singapore, 2014.¹

Soybean is an important worldwide source of dietary protein and oil. Over a period of time, soybean breeding practices have resulted in narrow diversity in the cultivated soybean germplasm. This has been attributed to underuse of wild soybeans as a breeding resource. Axiom Soybean Genotyping Array has markers from wild and cultivated soybean, making it ideal for whole-genome association studies and high-density genetic mapping² to identify important adaptive genes in wild and cultivated soybean germplasms.

The high marker density on the array avoids the underrepresentation of the soybean genome. This underrepresentation is often observed when employing other technologies, which offer insufficient SNP density or employ markers that exhibit bias towards certain regions of the genome.

The distribution of SNPs mapped to the Williams 82 reference sequence (Wm82.a2.v1) across the 20 chromosomes is superior to any other commercially available soybean SNP array.

**Highlights**

**Content**
A total of 180,961 SNPs were selected across the 20 soybean chromosomes and represent cultivated (max) and wild (soja) soybean accessions. The number of SNPs is estimated to provide approximately 1 SNP for every 6.1 kb.

- 114,735 SNPs or 63.4% of total markers are located in 40,631 genes.
- 22,952 SNPs are in 13,259 regions within 5 kilobase pairs upstream or downstream of genes.
- 43,274 SNPs are located in intergenic regions.

**Diversity**
- SNP discovery and validation were completed by using a diverse set of 16 soybean accessions from Korean populations and 31 accessions from Chinese populations.
- The array was evaluated on a diverse set of 228 lines, which included high-depth resequenced lines, duplicated DNA samples from different origins, duplicated DNA samples, diverse cultivated and wild lines, and multiple F2 and recombinant inbred lines.

**Applications**

**Genome structure and complex trait research**
- Perform GWAS and fine mapping
- Identify traits of economic importance
- Construct high-resolution genetic maps and detect natural variations in genome structure

**Molecular breeding**
- Perform association mapping and genomic selection in soybean breeding programs

**Population and evolutionary genetics**
- Distinguish between soybean populations of different origins
- Distinguish between cultivated and wild populations

**SNP selection**
The marker discovery was completed by high-depth sequencing of Korean³ populations containing 10 cultivated and 6 wild accessions and low-depth sequencing of 31 Chinese⁴ populations containing 14 max and 17 soja soybean lines. For each SNP discovery process, the sequences were aligned to the Williams 82 reference sequence (Wm82.a2.v1). The SNPs were filtered to eliminate SNPs with minor allele frequency ≤0.021 and SNPs that had mutations within 40 bp of the high-quality SNP. Putative SNPs were submitted to Affymetrix to calculate in silico design scores. The 180,961 SNPs selected on the final array included 40,136 SNPs from the KR population, 71,930 from the Chinese population, and 68,895 SNPs that were common between the two populations.
References


Results

The performance of the array was evaluated by genotyping 228 soybean lines. The lines included high-depth resequenced lines, duplicated DNA samples from different origins, duplicated DNA samples, diverse cultivated and wild lines, and multiple F2 and recombinant inbred lines. Samples were prepared using the guidance provided for the Axiom® 2.0 assay.

The data analysis and clustering were automated using Affymetrix’ Genotyping Console™ Software and Affymetrix’ SNPolisher™ package. SNPs were filtered as per the Best Practice Supplement to Axiom® Genotyping Solution Data Analysis Users Guide (P/N 703083). A total of 222 of the 228 samples passed the Affymetrix best practices workflow. The SNP genotyping results were automatically classified into six SNP classes. The concordance rates between SNP calls from PolyHighResolution and OTV clusters and SNP calls from high-depth genome resequencing among 10 cultivated and 4 wild soybean lines ranged from 94.7 to 99.4%.

Ordering information

<table>
<thead>
<tr>
<th>Part number</th>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>550469</td>
<td>Axiom® Soybean Genotyping Array</td>
<td>Contains one plate with 96 arrays.</td>
</tr>
<tr>
<td>902234</td>
<td>Axiom® GeneTitan® Consumables Kit</td>
<td>Contains all GeneTitan® Instrument consumables required to process one array plate</td>
</tr>
<tr>
<td>902245</td>
<td>Axiom® 2.0 Reagent Kit</td>
<td>Includes all reagents (except isopropanol) for processing 384 DNA samples</td>
</tr>
</tbody>
</table>