DataSheet

affymetrix
Biology for a better world

Axiom® Apple Genotyping Array

The most comprehensive high-density apple genotyping array

Axiom® Apple Genotyping Array (Axiom_Apple480) was designed through the Expert Design Program at Affymetrix in collaboration with the FruitBreedomics consortium (www.FruitBreedomics.com). The sequencing and marker selection was conducted by experts from Fondazione Edmund Mach, INRA, Dalhousie University, Wageningen UR, University Di Bologna, and Universita' Degli Studi Di Milano.

Apple (*Malus domestica*) is one of the most cultivated plants in the world. The apple genome is an ancient tetraploid, with some varieties being either allotetraploid or triploid. The apple genome is highly polymorphic with approximately 1 single-nucleotide polymorphism (SNP) per 50 bp and has a rapid linkage disequilibrium (LD) decay (20–55 kb). Axiom Apple Genotyping Array includes 480,000 markers and together with Axiom™ Analysis Suite software overcomes the genotyping challenges associated with polyploidy, rapid LD decay, and high polymorphism observed in the apple genome.

The 96-format array includes markers identified using whole-genome sequence data from 63 *Malus domestica* cultivars and two double haploid accessions. Table 1 lists the number of cultivars and corresponding country of origin used in the SNP discovery process. The names of these cultivars are provided in Table 2.

Array highlights

- Very high diversity
  - Includes markers discovered in 63 worldwide *M. domestica* cultivars
- High resolution to address rapid decay of LD
  - 487,249 markers on the array
  - Bias towards common variants with minor allele frequency (MAF) >0.05
  - Includes 21,463 previously validated markers: 19,990 markers from an existing in-market 20K Fruitbreedomics apple array and 1,473 markers identified using genotyping-by-sequencing (GBS)
- Absence of paralogous variants through the use of double haploid accessions in SNP discovery

Applications

- Construction of high-resolution genetic maps
- Fine mapping of quantitative trait loci
- Genome-wide association studies
- Selection sweep analysis

Table 1: The country of origin and the number of cultivars used in the SNP discovery process for Axiom® Apple Genotyping Array.

Country of origin	Number of cultivars
Australia	1
Central Europe	39
Canada	1
Iran	1
Japan	1
New Zealand	1
Northern Europe	4
Russia	9
Tunisia	1
United States	5

Comprehensive coverage of world-wide diversity in apples

The 63 cultivars used in sequencing represent diverse apple germplasm and include some of the core European apple breeding founder varieties. These cultivars were chosen to maximize the genetic diversity in the SNP discovery phase. Two double haploid (DH) accessions, `X9273' and `X9748' derived from Golden Delicious, were included to identify pseudo-SNPs created from the erroneous assembly of paralogous regions of the apple genome.

Markers that have been previously validated and associated with desirable traits are very important in maintaining and breeding elite commercial populations. The inclusion of 19,990 markers from the existing in-market 20K Fruitbreedomics genotyping array ensures that the new Axiom® array can be used for comparison with data generated by previous studies. The data analysis with Axiom Analysis Suite overcomes the limitations associated with poor coverage and the challenging analysis[1] of genotype data observed in the 20K in-market apple genotyping array. The backwards compatibility also provides the ability to continue existing projects, while making use of the latest and most informative content, to extend the usefulness of the study. The performance of the existing in-market array markers is less than ideal because the markers represent a small set of core founder lines, the arrays have a very low density making it difficult to work with the rapid LD decay in the apple genome, and data analysis software is not suitable for polyploid analysis.

A key benefit of the Axiom array is the capability to genotype SNPs that may have neighboring markers as few as 20 bp away. This design feature is important in genotyping the highly polymorphic structure of apple. The array manufacturing technology from Affymetrix also guarantees 100% fidelity and ensures all markers are present on every manufacturing batch, unlike other technologies that experience batch-to-batch variability and SNP dropouts.

Genotyping is performed using Axiom Analysis Suite in a convenient 96 format. With one-click analysis, hands-on time for genotyping is reduced, minimizing costs and time to results. Axiom Analysis Suite genotypes and classifies the markers into six easy-to-visualize categories. The AxiomGT1 algorithm is the only algorithm that adapts to shifted clusters and cluster compression that is typically observed in polyploid species, eliminating the need for manual editing of the clusters and manual assignment of genotypes.

Array design

The markers on the array were identified from whole-genome data from 63 cultivars. The average number of reads for each cultivar was 95.2 million, which represents a mean sequencing depth of 25X. Sequencing reads were mapped as single ends on the reference genome.[2] A total of 15.5 million markers was identified from the whole-genome sequencing data. A putative list of 12,701,549 markers was submitted to Affymetrix to calculate *in silico* design scores. The putative list was generated by removing (i) markers with a low-quality phred score (<20), a high combined read depth (>4,000), and a low single-cultivar read depth (<8) in more than 50% of the sequenced cultivars; (ii) heterozygous markers identified in the DH cultivars because these are evidence of paralogous sequences; and (iii) insertion or deletions. The *in silico* design pipeline developed by Affymetrix identified the following additional markers that were then also removed: (i) markers with low *in silico* design score (<0.6), (ii) markers with 16-mer count >300 in the genome, (iii) multi-allelic markers, (iv) A/T or C/G transversions, and (v) markers with a SNP 35 bases up/downstream.

The remaining 2.8 million markers were used for choosing markers for array synthesis. The following criteria were applied to choose all the tag markers within a ±10 kbp window: (i) markers

Table 2: Cultivars used in SNP discovery efforts for the markers on Axiom® Apple Genotyping Array.

Grouping	Breeds
Australia	Lady Williams
Belgium	Belle et Bonne, Cabarette, Court-Pendu, Henry, Godelieve Hegmans, Président Roulin, Reinette Dubois
Bulgaria	Aivaniya
Canada	McIntosh
Czech Republic	Chodské, Hetilina, Košíkové, Malinové holovouské, Panenské české, Sonderskow
Denmark	Filippa
Filand	Maikki
France	Alfred Jolibois, Amadou, De L'Estre, Patte de Lous, Pepino Jaune, Reinette Clochard
Germany	Doctor Oldenburg
Greece	Fyriki
Iran	Precoce de Karaj
Italy	Abbondanza, Busiard, Durello di Forl, Gelata, Mela Rozza, Renetta Grigia di Torriana, Rosa (FI)
Japan	Fuji
New Zealand	Braeburn
Russia	Ag alma, Antonovka, Antonovka Pamtorutka, Aport Kuba, Borowitsky, Ijunskoe ranee, Jantarnoe, Ovčí hubička, Papirovka, Skry (Skryzhapel)
Serbia	Budimka
Sweden	Akerö, Heta, Kronprins, Spässerud
Tunisia	Ajmi
United States	Delicious, F2-26829-2-2, Golden Delicious, Jonathan, Macoun, Priscilla-NL, Young America
United Kingdom	Cox's Orange Pippin, Keswick Codlin, Worcester Pearmain

in genic regions with high MAF ≥0.1 and a Hardy-Weinberg Fisher's test p-value >10^{-8} with less than 32 missing genotypes, (ii) markers in intergenic regions with MAF >0.1 and Hardy-Weinberg Fisher's test p-value >10^{-8} with less than 14 missing genotypes, and (iii) genomic markers with 0.05 ≤ MAF <0.1 with Hardy-Weinberg Fisher's test p-value >10^{-8} with less than 14 missing genotypes. The 465,786 markers identified using this method were then combined with 21,463 previously validated markers from the in-market 20K Fruitbreedomics array and GBS data. The resulting 487,249 markers selected for array synthesis represented 40,192 sequence contigs and 562 Mb of the apple genome.

Automated genotyping and classification

Axiom Apple Genotyping Array was evaluated with a diverse set of cultivars to demonstrate the array's performance. A total of 1,200 samples were processed and analyzed using Axiom Analysis Suite, as per the *Axiom® Genotyping Solution Data Analysis Guide* (PN 702961 Rev. 3). Approximately 360,565 or 74% of the markers were automatically identified as high-quality markers under the polymorphic high-resolution (PolyHighResolution) category. The call rate of markers in this category was greater than 99%. The data was automatically clustered, assigned genotypes, and classified into six categories for easy visualization. SNP concordance with sequencing was carried out by genotyping 42 of the 63 accessions used in SNP discovery. A total of 347,805 markers in the PolyHighResolution category had 96% concordance, demonstrating the success of the array and the appropriate selection of the cultivars for SNP discovery.

Figure 1: Plots showing markers from each of the four recommended SNP categories on Axiom® Apple Genotyping Array. The plots from left to right represent a marker from the following SNP categories: 1) PolyHighResolution markers, 2) NoMinorHomozygous markers, 3) OffTargetVariant markers, and 4) MonoHighResolution markers.

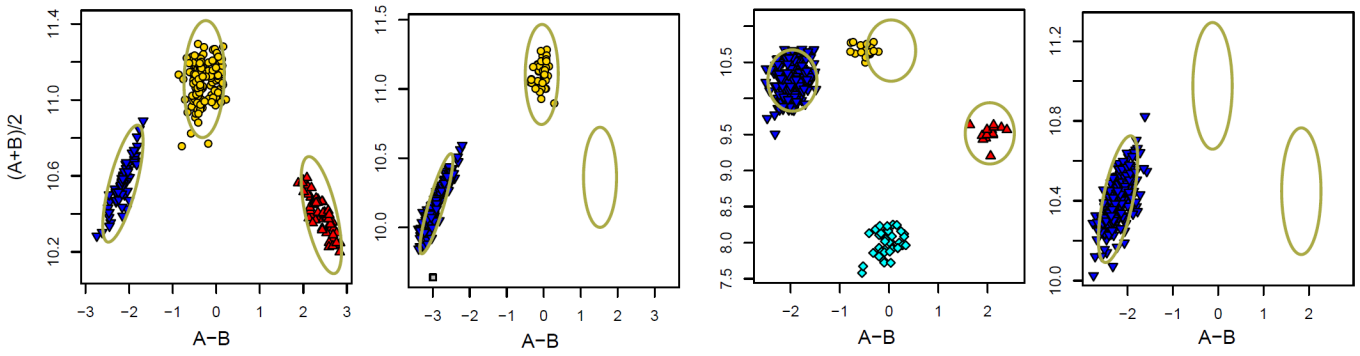

Table 3: Axiom® Apple Genotyping Array results assigned into six categories. The third column displays the classification of the markers that are available on the legacy apple array with 20,000 markers. The markers in the recommended categories include: 1) PolyHighResolution markers: markers demonstrating three clusters with good cluster resolution and at least two examples of the minorallele; 2) NoMinorHomozygous markers: markers exhibiting two clusters with no examples of the minor allele; 3) MonoHighResolution markers: markers demonstrating a single cluster; 4) OffTargetVariant markers: reproducible yet uncharacterized variants caused by double deletion, sequence non-homology, or DNA secondary structure.

SNP classification	Percentage of all markers in the different SNP categories (%)	Percentage of markers that were previously validated on 20K in-market array and GBS (%)
All markers	100%	100%
Recommended markers		
▪ PolyHighResolution	74	57
▪ NoMinorHomozygous	2	5
▪ MonoHighResolution	1	4
▪ OffTargetVariant	1	3
Unexpected heterozygosity	2	9
High variance	6	2
CallRateBelowThreshold(<97%)	6	6
Other	8	14

affymetrix
Biology for a better world

References

1. Bianco L., *et al.* Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus* × *domestica* Borkh). *PLoS ONE* **9**(10): e110377 (2014).

2. Velasco R., *et al.* The genome of the domesticated apple (*Malus domestica* Borkh.) *Nature Genetics* **42**(10):833{9} (2010).

Ordering information

Part number	Description	Details
550573	Axiom® Apple Genotyping Array	Contains one plate with 96 arrays. Reagents and GeneTitan® MC consumables must be quoted separately
901606	Axiom® GeneTitan® Consumables Kit	Contains all GeneTitan® Instrument consumables to process one 96-format array plate
901758	Axiom® 2.0 Reagent Kit	Includes all reagents (except isopropanol) to process one Axiom® 96-format array plate

Affymetrix, Inc: **(US)** +1-888-362-2447, +1-408-731-5000 ■ **(EU)** +44-(0)1628-552550 ■ **(JP)** +81-(0)3-6430-4020 ■ **(CN)** +86-21-63915511
eBioscience Products: (US) +1-888-999-1371, +1-858-642-2058 ■ **(EU)** +43 1 796 40 40 305 ■ **(JP)** +81-(0)3-6430-4020
USB Products: **(US)** +1-800-321-9322, +1-216-765-5000 ■ **(EU)** +44-(0)1628-552600

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

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

P/N GGNO06531 Rev. 1