An improved version of cultivated strawberry linkage map using the IStraw90 Axiom® Array for QTL analysis.

J.M. Hidalgo, P. Arús, A. Monfort.
IRTA, Centre de Recerca en Agrigenòmica CSIC-IRTA-UAB-UB

① Introduction

The cultivated strawberry, *Fragaria x ananassa*, is a very economically-important fruit species member of the Rosaceae family. Although being an important fruit, the octoploid nature of its genome (2n =8x = 56) presents a challenge to the development of molecular breeding tools. *F. vesca*, a closely related diploid species (2n = 2x =14), has been used as a model organism and the availability of its genome boosts the research in strawberry.

Previously, a genetic linkage map was developed from a F2 population (21-AF) comprising 96 seedlings derived from two *F. x ananassa* cultivars, 'Camarosa' and 'Dover', as a tool to identify loci near qualitative and quantitative traits. More recently, the RosBREED consortium developed the IStraw90™ SNP array to genotype strawberry cultivar collections and populations. We used this array with more than 90K SNPs to hybridize our F2 population. With this information we could develop an improved version of the genetic map that will be a valuable tool for further analyses like QTL mapping of nutritional traits or the genome assembly assessment.

② IStraw90 SNP Array

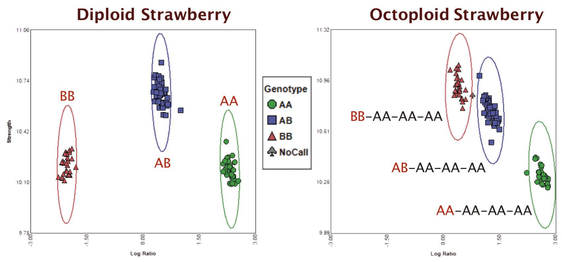
Affymetrix and RosBREED developed in late 2013 the IStraw90™ SNP array, with **approximately 94000 SNPs** obtained from sequences from *F. vesca, F. x ananassa* and *F.innumae,* another related diploid species.

The IStraw90™ SNP array was hybridized with **120 individuals and parentals** of our population and analyzed using GenotypeConsole and SNPolisher software. After checking the quality of the hybridization, with a **call rate above 97%** for all individuals, the SNPs were classified into 6 clusters: PolyHighResolution, MonoHighResolution, OTV or "null" allele, NoMinorHomozigous, CallRate Below Threshold and Other.

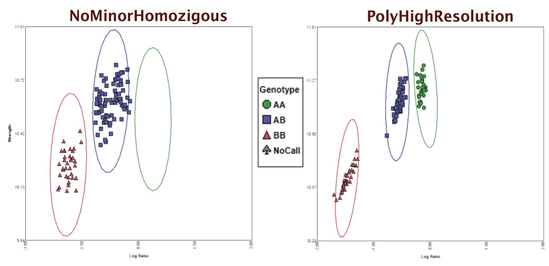
Out of the 6 types, only **PolyHighResolution** and **NoMinorHomozygous** were valid to create the genetic map.

SNP Type	Number	%
PolyHighResolution	**4939**	5,3%
NoMinorHomozygous	**9356**	10,0%
CallRateBelowThreshold	2881	3,1%
MonoHighResolution	70470	75,0%
OffTargetVariant	73	0,1%
Other	6218	6,6%
Total	**93937**	**100%**

A significant amount of SNPs were discarded due to the low CallRate produced by the ploidy level of *F. x ananassa*. The presence of 4 homologous chromosomes causes a bias in genotype identification towards the references genotype, reducing the discrimination capacity between the three possible genotypes.



Comparison of genotype cluster identification of diploid (left) and octoploid (right) samples. The graph is represented with both channels (A and B) intensity signal.



Cluster genotyping profile of NoMinorHomozigous (left) and PolyHighResolution SNPs (right).

Above all types of SNP classification, PolyHighResolution were the most valuable ones because those are the most reliable and the most informative markers given its **codominant** nature. NoMinorHomozigous SNPs were also used because, despite being less informative, are still reliable as a **dominant** marker.

③ Genetic Map

Back in 2013, a genetic map based on SSR markers was developed from the 21-AF population. With 131 SSR markers, generating 192 loci along the 28 expected Linkage Groups (LGs) with an approximate average of 7 loci and 83 cM per LG.

Using the **10.134** segregating SNPs obtained with the IStraw90™ array, a new genetic map has been constructed. **7168 SNPs** were mapped to the **28 LGs** with an approximate average of 265 loci and 59 cM per LG. The annotation of the LGs was conserved from the previous map by including the SSR markers in the construction of the new map.

This new map has a greater marker density and higher genome coverage than the previous, making it a better tool for further analysis like QTL mapping. The distribution of the SNPs across the homeologous groups will assess the assembly of contigs and scaffolds of the genome draft sequence covering 154x of one population individual obtained by Illumina paired-end sequencing.

Previous genetic map (left) and new genetic map (right) of two *Fxananassa* LGs. SSR markers are annotated, SNP markers are represented with black lines.

Number of loci per LG

LG	A	B	C	D
1	279	-	260	169
2	168	245	224	74
3	693*	339	384*	112*
4	223	267	148	232
5	573	328	-	329
6	236	120	389	442
7	297	263	129	245

cM per LG

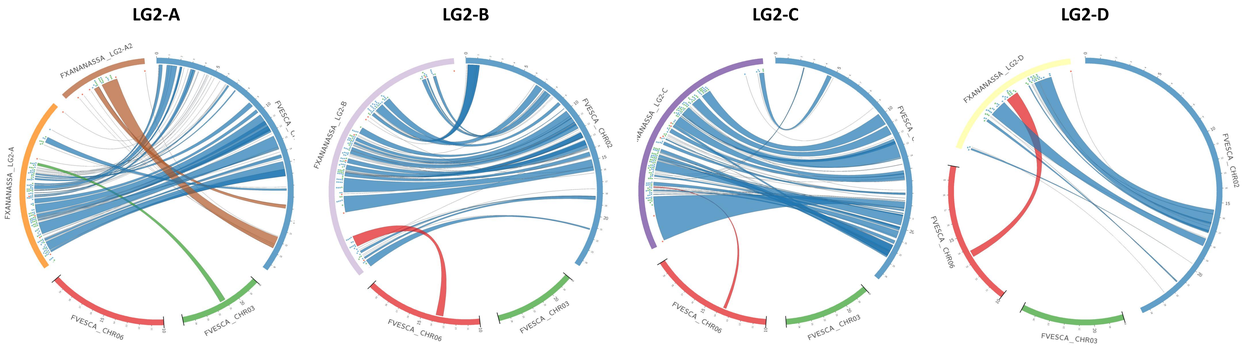
LG	A	B	C	D
1	81,16	-	69,58	60,94
2	56,31	92,92	78,15	38,519
3	-	70,09	25,94*	40,84*
4	75,95	70,62	55,84	49,81
5	-	78,97	-	101,29
6	-	-	-	-
7	76,97	39,67	25,29	55,14

* Not assigned with old SSR LGs

④ Genome synteny with *F. vesca*

Due to *F. vesca* genome simplicity compared to *F. x ananassa*, it has been used as a **model organism** for the strawberry research. Tools developed for the diploid species and the knowledge obtained can be a powerful tool to enhance the research on the octoploid strawberry.

The known position of the array SNPs in *F. vesca* genome and the development of a genetic map in *F. x ananassa* makes possible a synteny analysis between both species. As it was expected, there is a **high synteny** across both genomes. With this information, we can not only transfer the knowledge of gens or regions of interest in *F. vesca* but also we can observe the bigger genome rearrangements and differences between the **4 octoploid subgenomes**.

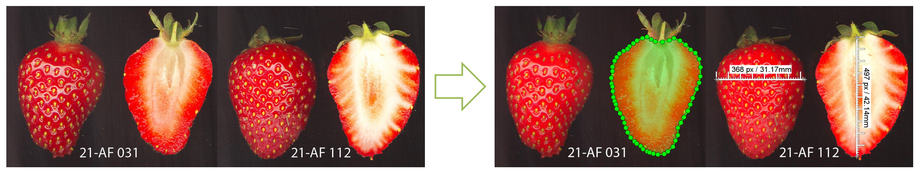


Circles representing the linkage between each homeologous subgenome of the LG2 in *F. x ananassa* and its physical position in the F. vesca chromosomes. Across the LGs, red, green and blue tiles represents, respectively, the mapping position of SSR, dominant and codominant SNPs.

⑤ Phenotyping

The 21-AF population is being maintained at the greenhouses and during the next seasons fruits will be collected to perform linkage analyses between the new genetic map and several traits of interest. The main objective of the analyses will be the study of characters related to the fruit quality of the strawberries, such as the size and shape, but especially those relevant to the **nutritional and aromatic profile** of the fruit. For this, fruits obtained from each individual of the population will be analyzed with **HPLC** for sugar content and total polyphenol content and the polyphenolic profile will be determined with **LC-MS** .

Due to the poor segregation of aromatic compounds in the 21-AF population, a better suitable population will be used to study the aromatic profile with **GC-MS** and its linkage to mapped SNPs.



Strawberry fruits halved and scanned for phenotyping purposes, some traits can be determined with this method.

⑥ Acknowledgments

• AGL2010-21414 (Nutrifresa) by Science and Technology Spanish Minister and PLANASA, Innovation in plant varieties.

Download the PDF version of this poster.