# Molecular characterization of apricot genotypes by SSR markers

S. Herrera[1,2], J. Lora[3], J.I. Hormaza[3], G. Ylla[4] and J. Rodrigo[1,2]

[1] Unidad de Hortofruticultura. Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Zaragoza, Spain

[2] Instituto Agroalimentario de Aragón – IA2 (CITA-Universidad de Zaragoza), Zaragoza, Spain

[3] Instituto de Hortofruticultura Subtropical y Mediterránea La Mayora (IHSM La Mayora-UMA-CSIC), Algarrobo-Costa, Málaga, Spain

[4] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

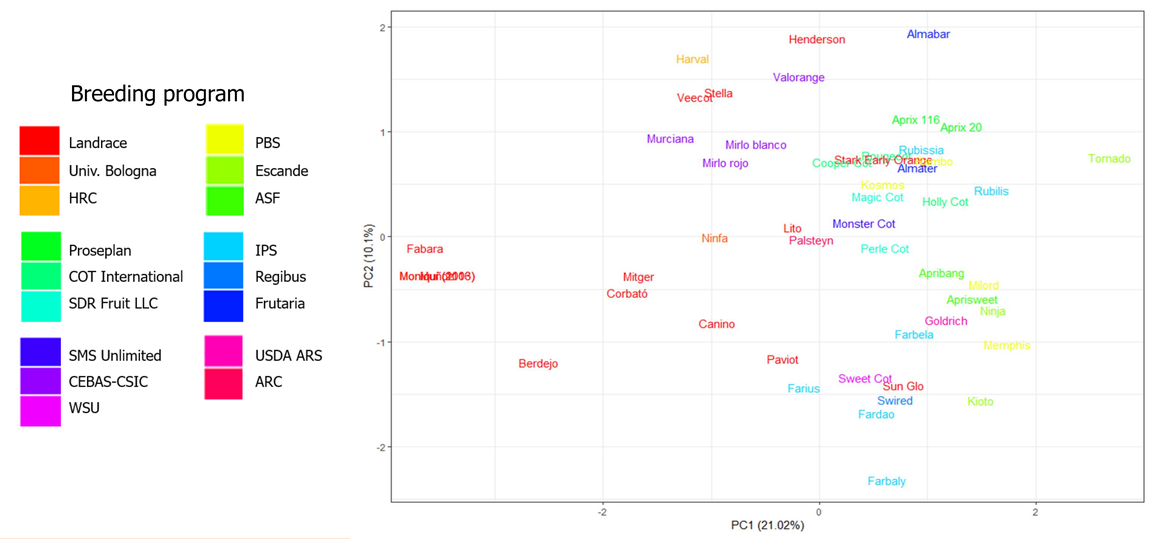
**In the last few years, the release of an increasing number of new apricot cultivars from different breeding programs is resulting in an important renewal of plant material worldwide. The new cultivars are replacing traditional and local varieties in many situations. However, local varieties constitute a source of genetic traits of interest and conserving and analyzing this genetic pool allows preserving genetic material for future breeding programs. In order to study the current diversity and determine genetic relationships among genotypes, in this work fifty apricot cultivars have been analyzed, including traditional and new cultivars from breeding programs of different countries.**

| Locus | Linkage group | SSR motive | Reference | $N_a$ | Range (bp) | $H_o$ | $H_e$ | $F_{IS}$ | $F_{ST}$ |
|---|---|---|---|---|---|---|---|---|---|
| ssrPaCITA7 | 1 | $(AG)_{22}$ | Lopes et al. (2002) | 10 | 189 - 224 | 0,76 | 0,77 | -0,19 | 0,07 |
| ssrPaCITA10 | 3 | $(CT)_{26}$ | Lopes et al. (2002) | 9 | 155 - 180 | 0,82 | 0,79 | -0,27 | 0,15 |
| ssrPaCITA23 | 3 | $(AC)_2(AG)_{18}$ | Lopes et al. (2002) | 9 | 140 - 156 | 0,88 | 0,82 | -0,25 | 0,10 |
| ssrPaCITA27 | | $(TC)_8(TA)_6(TG)_{17}$ | Lopes et al. (2002) | 6 | 228 - 267 | 0,54 | 0,68 | 0,03 | 0,22 |
| UDAp_415 | 1 | $(GA)_{21}$ | Messina et al. (2004) | 6 | 149 - 171 | 0,78 | 0,75 | -0,19 | 0,12 |
| UDAp_420 | 6 | $(CT)_{20}$ | Messina et al. (2004) | 6 | 158 - 181 | 0,84 | 0,76 | -0,19 | 0,06 |
| pchgms3 | 1 | $(CT)_{19}$ | Sosinski et al. (2000) | 6 | 180 - 200 | 0,56 | 0,60 | -0,28 | 0,18 |
| **Mean** | | | | **7** | | **0,74** | **0,74** | **-0,19** | **0,13** |

Molecular characterization was carried out using 7 microsatellite loci (SSRs), which produced polymorphic repeatable amplification patterns. A total of 52 alleles were found in the 50 analyzed cultivars, ranging from 6 to 10 in the different loci ($N_a$).

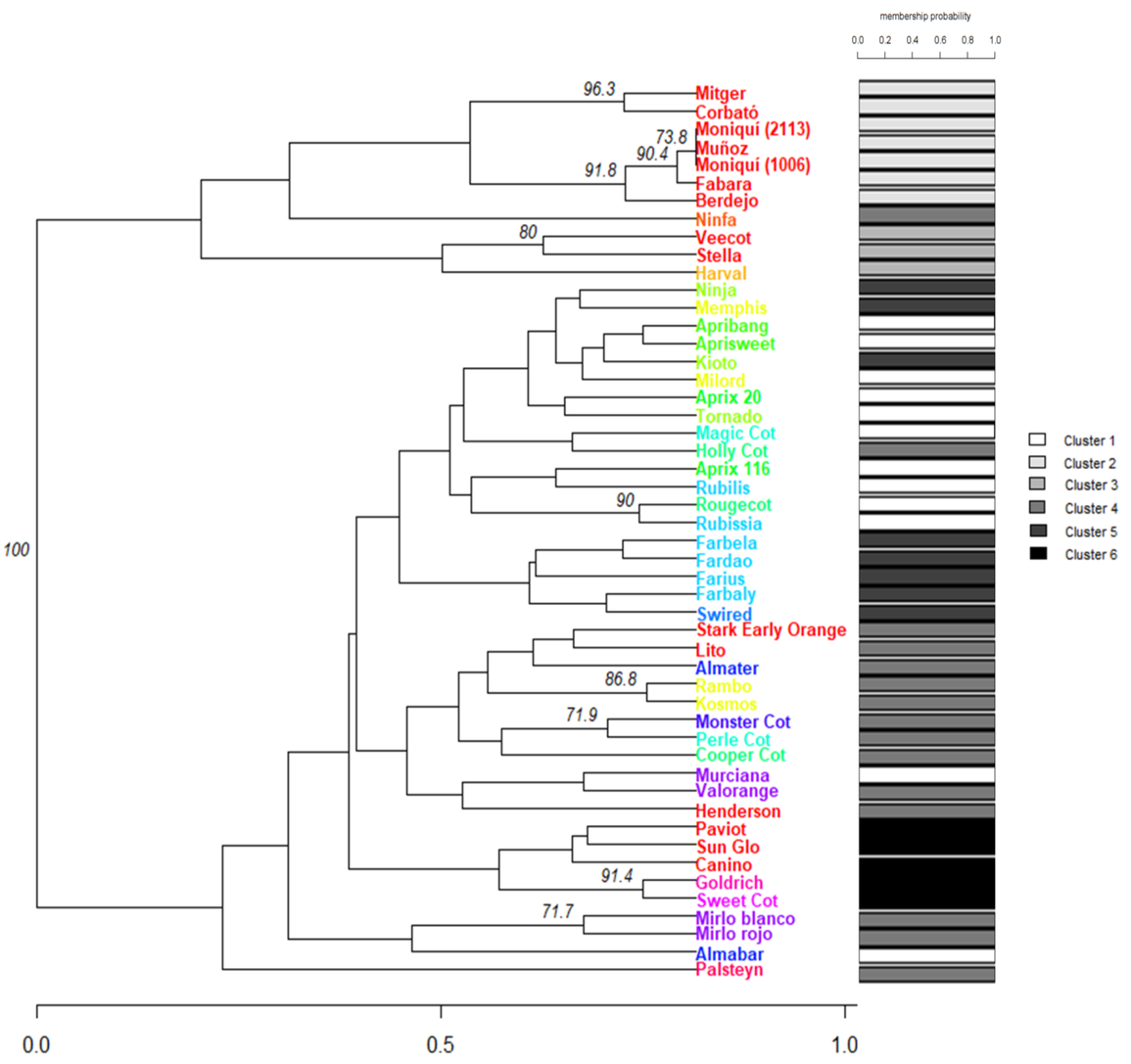
The analysis of genetic diversity showed that the mean observed heterozygosity ($H_o$) was equal to the mean expected heterozygosity ($H_e$). The mean $H_o$ was higher (0.74) than previous studies in which only landraces were analyzed. This could be due to the different origin of the accessions analyzed herein and the diversity of the parents used for breeding.

The negative values for inbreeding coefficient ($F_{IS}$) indicated a high number of heterozygotes. The mean fixation index ($F_{ST}$) was 0.13, indicating a low genetic differentiation probably due to the fact that most of the populations share some genotypes as parents in breeding programs.




Statistical analysis were performed in the R programming environment in order to evaluate the genetic structure. Principal Component Analysis (PCA) was performed as a first approximation. The optimal number of main components selected was 20 using the Eigenvalue values since they explained 90% of the accumulated variance.

The first axis (PC1, 21,02%) reflected population differentiation corresponding to the breeding origin forming two main clusters. Landraces were located on the left along the x axis, whereas commercial cultivars were placed to the right. In the second axis (PC2, 10,1%), no clear differentiated groups could be observed, probably due to the use of common parentals in the different breeding programs.




The obtained SSR profiles allowed us to distinguish 48 unique genetic profiles, identifying three synonymies. UPGMA cluster analysis based on Nei's genetic distance was used for analyzing the similarity relationships. The dendrogram revealed two main clusters supported by a strongly bootstrap value (100) in which the cultivars were classified according mainly to the geographical origin of the cultivars and/or the breeding program. The first group comprised traditional cultivars from Spain and North America. The second group included the majority of new commercial cultivars showing a tendency of grouping according to the breeding program.

Discriminant Analysis of the Principal Components (DAPC) was performed to study the genetic structure of the populations. Six groups were identified by the Bayesian information criterion (BIC). Spanish landraces were assigned into the cluster 1 and cluster 2 contained North American landraces. No clear population differentiation was observed among the new commercial cultivars that were classified in the clusters 3 to 6.

**These results reveal a clear differentiation between traditional cultivars and new commercial apricot cultivars developed from breeding programs. Moreover, the results showed a high level of inbreeding, which is consistent with the use of common parents in the different breeding programs since they share similar breeding objectives.**

**Thus, the introduction of new releases is increasing the number of cultivated apricot varieties but the use of common parents can result in an erosion of the genetic diversity resulting in the replacement of more diverse landraces by genetically related genotypes and, consequently, in a decrease of overall genetic diversity in cultivated apricot worldwide.**

ACKNOWLEDGEMENTS

This research was funded by Ministerio de Ciencia, Innovación y Universidades - European Regional Development Fund, European Union (INIA RFP2015-00015-00, INIA RTA2017–00003–00); Gobierno de Aragón - European Social Fund, European Union (Grupo Consolidado A12–17R).