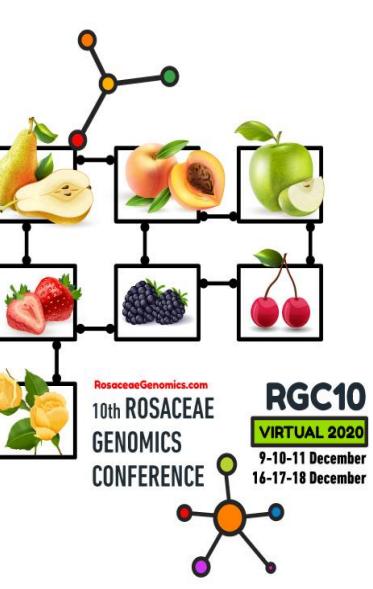# Exploring genome-wide diversity in the National Peach (*Prunus persica*) Germplasm Collections at CITA and EEAD-CSIC (Zaragoza, Spain)

Jorge Mas-Gómez[1], Celia M Cantín[2], María Ángeles Moreno[3], Mar Gómez-Abajo[1], María J Muñoz-Morales[1], Angela S Prudencio[1], Federico Jurado-Ruiz[1], Sara Jalili[4], Balazs Nagy[1], Luca Bianco[5], Michela Troggio[5], Pedro Martínez-Gómez[1], Manuel Rubio[1], Santiago García-Martínez[6], Pedro J Martínez-García[1]

[1]Department of Plant Breeding, CEBAS-CSIC, Murcia, Spain
[2]ARAID at Department of Horticulture, CITA-IA2 Aragón Zaragoza, Spain
[3]Department of Pomology, Estación Experimental de Aula Dei - CSIC Zaragoza, Spain
[4]Department of Horticultural Science, Tarbiat Modares University (TMU) Tehran, IRAN
[5]Department of Genomics and Biology of Fruit Crop (DGBPF), Fondazione Edmund Mach - Istituto Agrario San Michele All Adige San Michele All Adige, Italy
[6]Department of Applied Biology, Universidad Miguel Hernández Escuela Politécnica Superior de Orihuela, Orihuela, Alicante, Spain

INTRODUCTION

Peach [Prunus persica (L.) Batsch] is one of the most studied stone fruits. The preservation of plant germplasm collections is crucial to ensure genetic diversity in plants. Mainly due to the self-compatibility in peach and the low number of genotypes used in breeding programs, a low intraspecific genetic diversity is present in peach (Badenes et al., 2015). Genetic diversity provides the fundamental basis for new diseases and pests resistances in fruit trees and their adaptation to climate change. Molecular tools contribute to create core collections in genebanks, collecting the maximum genetic diversity and avoiding redundant genotypes (Wambugu et al., 2018).

In this work a set of 90 accessions have been sampled from the National Peach Germplasm Collection of the "Centro de Investigación y Tecnología Agroalimentaria de Aragón" (CITA) and the "Estación Experimental de Aula Dei" (CSIC), both located in Zaragoza (Spain). These 90 samples have been genotyped using the new version of the high-density Illumina peach SNP chip (18K). Genetic diversity and population structure have been analysed to obtain useful information for an efficient preservation of peach germplasm.

MATERIAL AND METHODS

Among the 90 accessions, 84 were Spanish and 6 were foreign ["Andora" (2 clones), "Vivian" (5266), "Aurelio" were from United States of America (US), "Pepita" from Brazil (Br) and "Aniversario" from Argentina (Ar)]. Spanish accessions were from Zaragoza (28) (Za), Región de Murcia (18) (Mu), Lérida (13) (Lé), Huesca (12) (Hu), Navarra (9) (Na), Teruel (2) (Te), La Rioja (1) (La) and Valencia (1) (Va). There were cultivars of yellow or white flesh, and some flat peaches too. Illumina Infinium Peach BeadChip 18K (18000 SNPs) was used to genotype the samples. This large number of SNPs were scored with the GenomeStudioTM Genotyping Module v2.0 from Illumina (Illumina, Inc.). Afterwards, Assist was used to classify SNPs.

Measures of genetic diversity (Ho, He, FST, etc.) were estimated with the DiveRsity R package to describe the populations and their genetic differences. Through Plink software, genotype data was used to provide an estimate of pairwise Identity-By-Descent (IBD), to detect clones. Genetic structure analysis was conducted with fastStructure. Clusters (K) were set from 1 to 10. For the choice of the most likely K, chooseK.py script was considered.
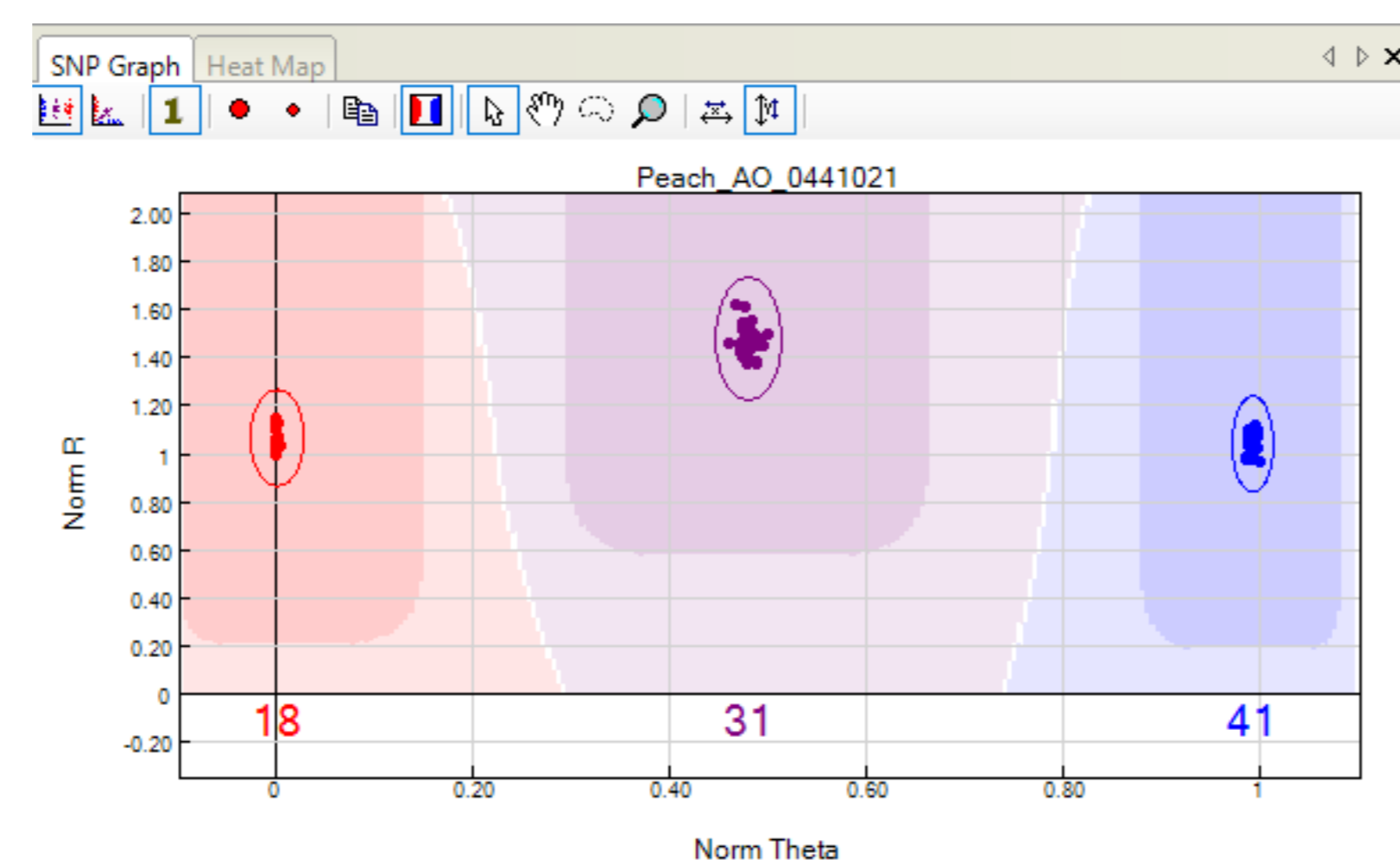

Figure 1. SNP Graph in GenomeStudio.

RESULTS AND DISCUSSION

A total of 16038 SNPs were scored in GenomeStudio. A total of 8955 (55.84%) were categorized as Approved by Assist, among them 3247 (20.25%) were robust. In addition, for 2780 (17.34%) SNPs one of the homozygous genotype were present only in 5% or less of our individuals and 2928 (18.26%) SNPs were described as distorted and with an unexpected segregation. Within of discarded SNPs (7083, 44.16%), 2724 were monomorphic, 1016 failed and in 3175 SNPs one of the homozygous genotypes were absent in our population.
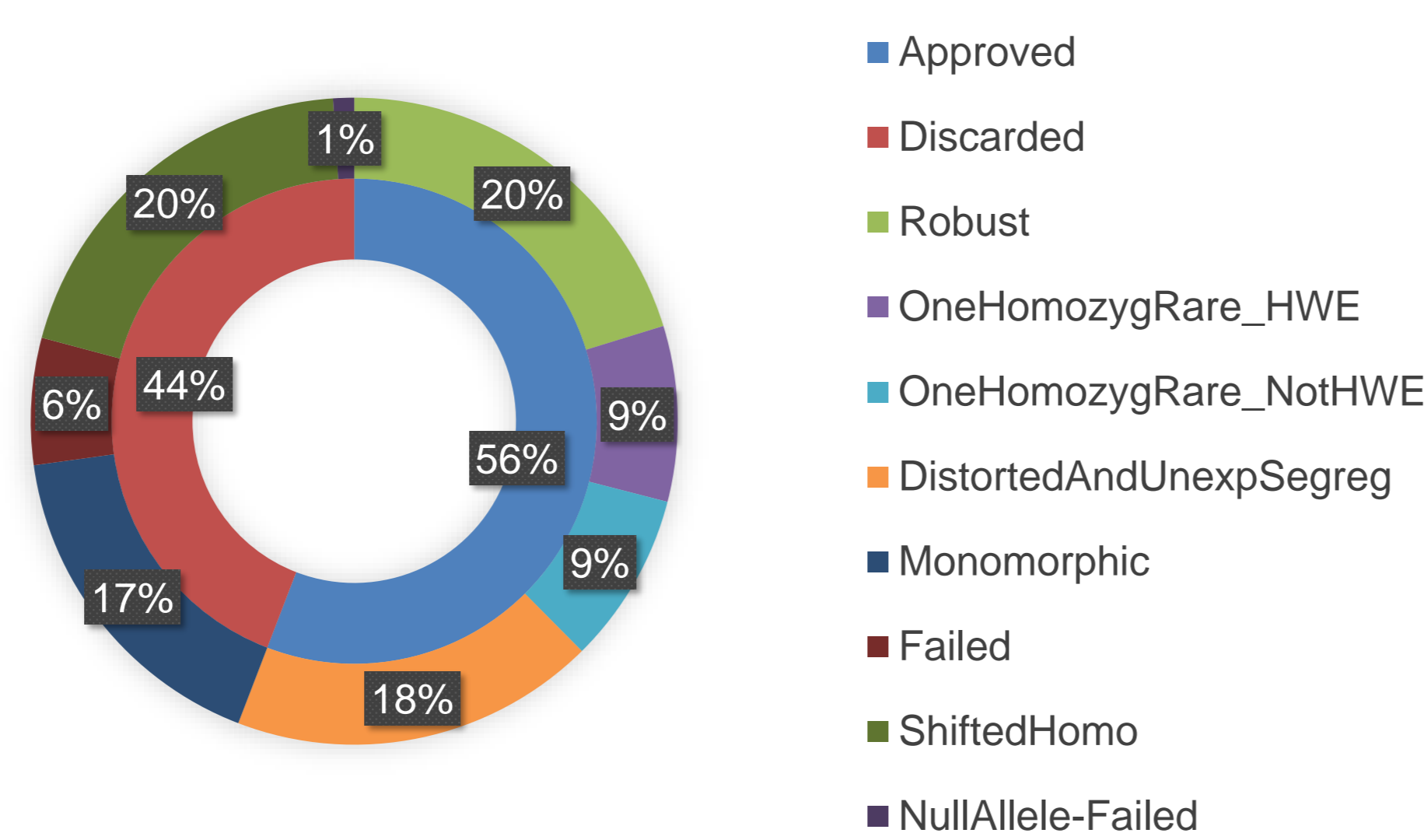

Figure 2. Classification of SNPs using Assist

- Approved
- Discarded
- Robust
- OneHomozygRare_HWE
- OneHomozygRare_NotHWE
- DistortedAndUnexpSegreg
- Monomorphic
- Failed
- ShiftedHomo
- NullAllele-Failed

Regarding of the relatedness analysis, the results obtained with Plink using robust SNPs, have showed that 44 accessions, representing 15 different groups, are individuals identical or clones. The 15 groups of clones can be observed in Figure 3 (each group have a different colour). All the individuals in the same group showed a value of PI_HAT=1, in the pairwise analysis. In the other hand, individuals with same ID such as "La Escola" (5093) and "La Escola" (5340) obtained a PI_HAT=0, indicating that were not really clones. Pérez et al. (2020) genotyped with 10 SSRs some of the individuals used in the present study. Several differences have been found in our study.
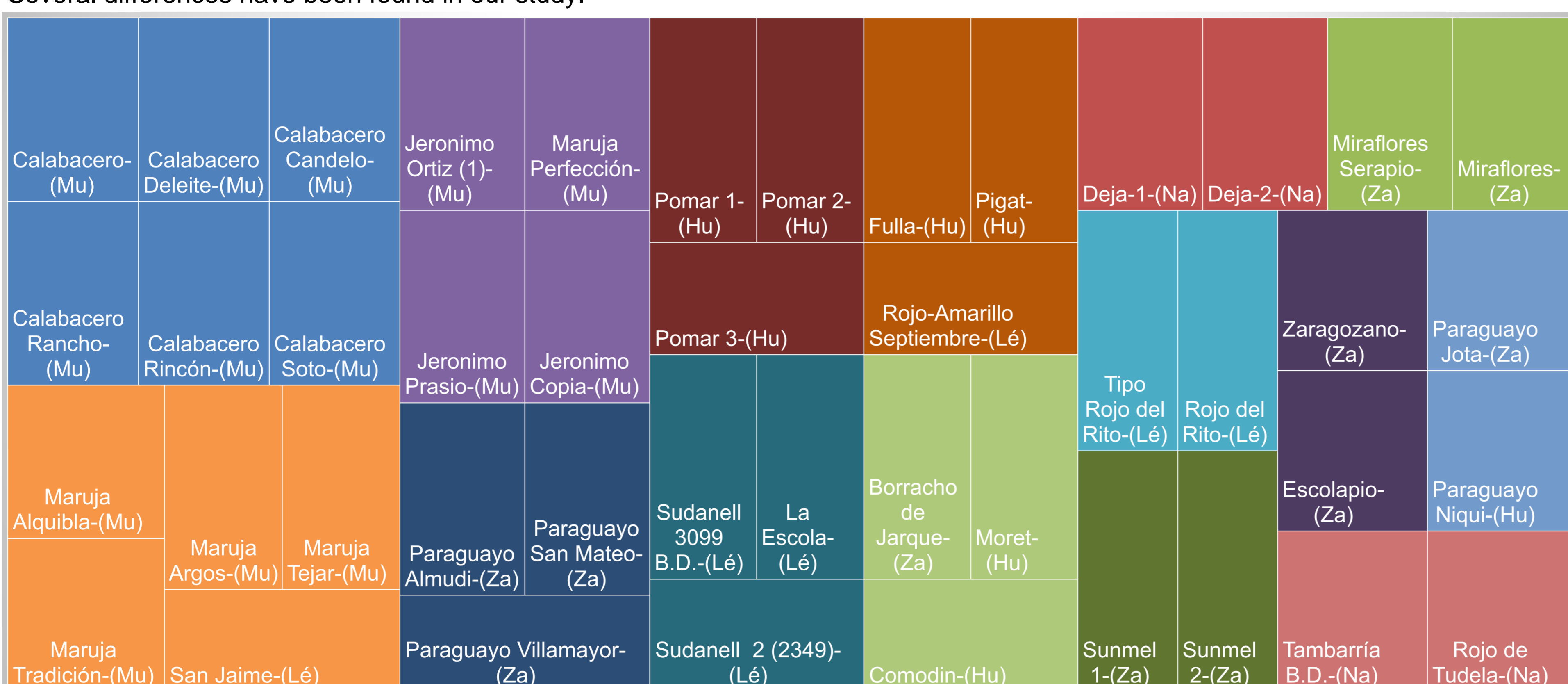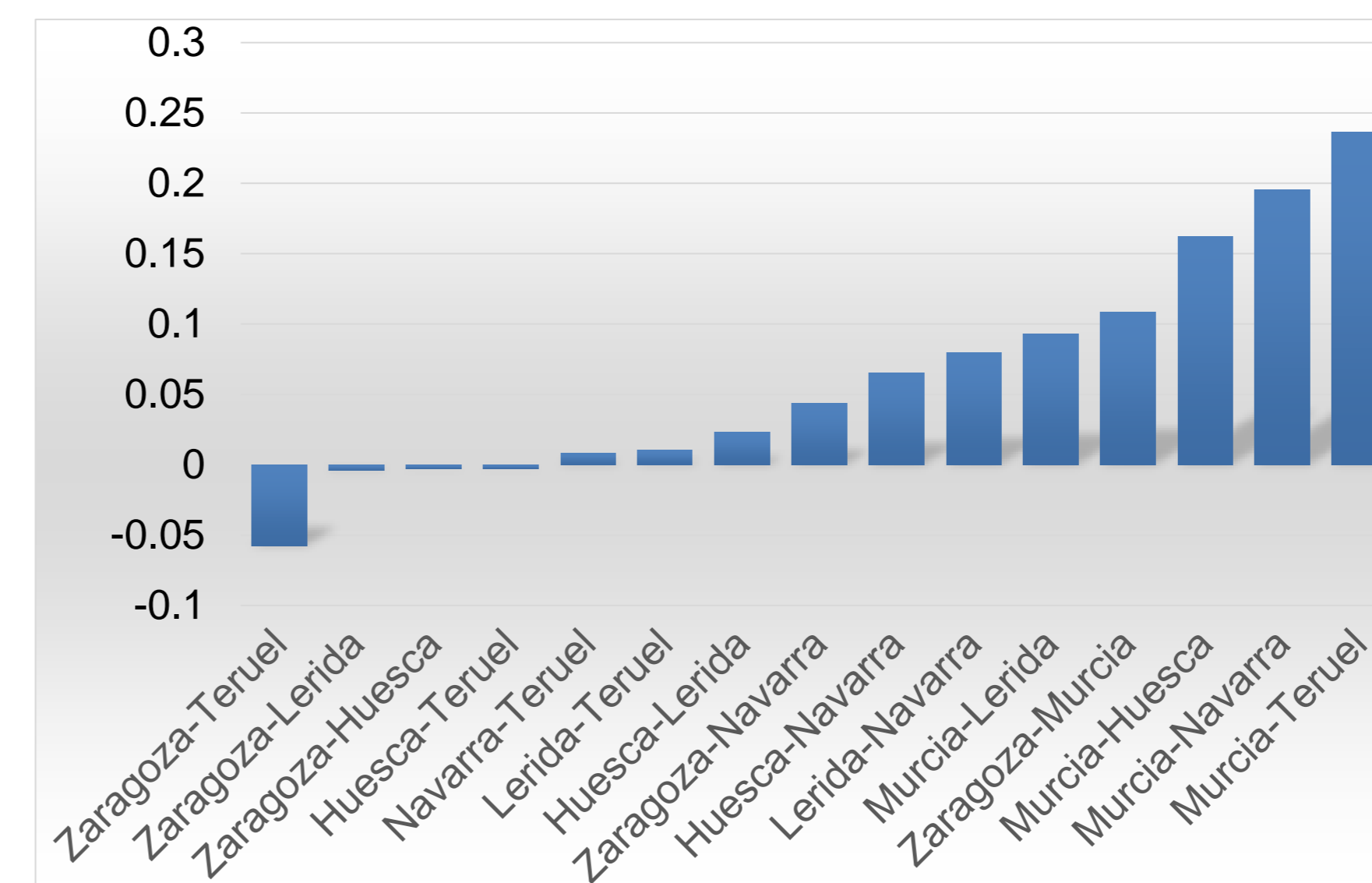

Figure 3. Clones detected through IBD analysis in Plink

Figure 4. F$_{ST}$ pairwise values obtained with DiveRsity R package.

According to the genetic diversity results, individuals from Murcia (South Spain) have showed the highest values of pairwise F$_{ST}$ with other populations (North Spain) (Figure 4). Large differentiation with Teruel (0.2365), Navarra (0.1956) and Huesca (0.1624) and moderate differentiation with Zaragoza (0.1085) and Lérida (0.0930) were observed. Lower values of F$_{ST}$ in the rest of pairwise analysis were obtained. These results suggest an exchange of plant material in the north of Spain and a different genetic background in accessions from Murcia region.

The results of the chooseK.py script showed K=5 as the best option (Marginal likelihood= -0.7905), but K=7 (Marginal likelihood= -0.7977) was selected to provide a best explanation of the genetic structure (Figure 5). These results shown a green cluster that grouped "Calabacero" type individuals, but also the individual "Campillo Rocho" from Zaragoza. "Campillo Rocho" also shows affinity with dark blue cluster, which is composed of accessions from Murcia. This fact suggests that "Campillo Rocho" is genetically close to Murcia cultivars. Overall, the population structure analysis revealed some differences regarding the geographic origin considered at the beginning of this work.
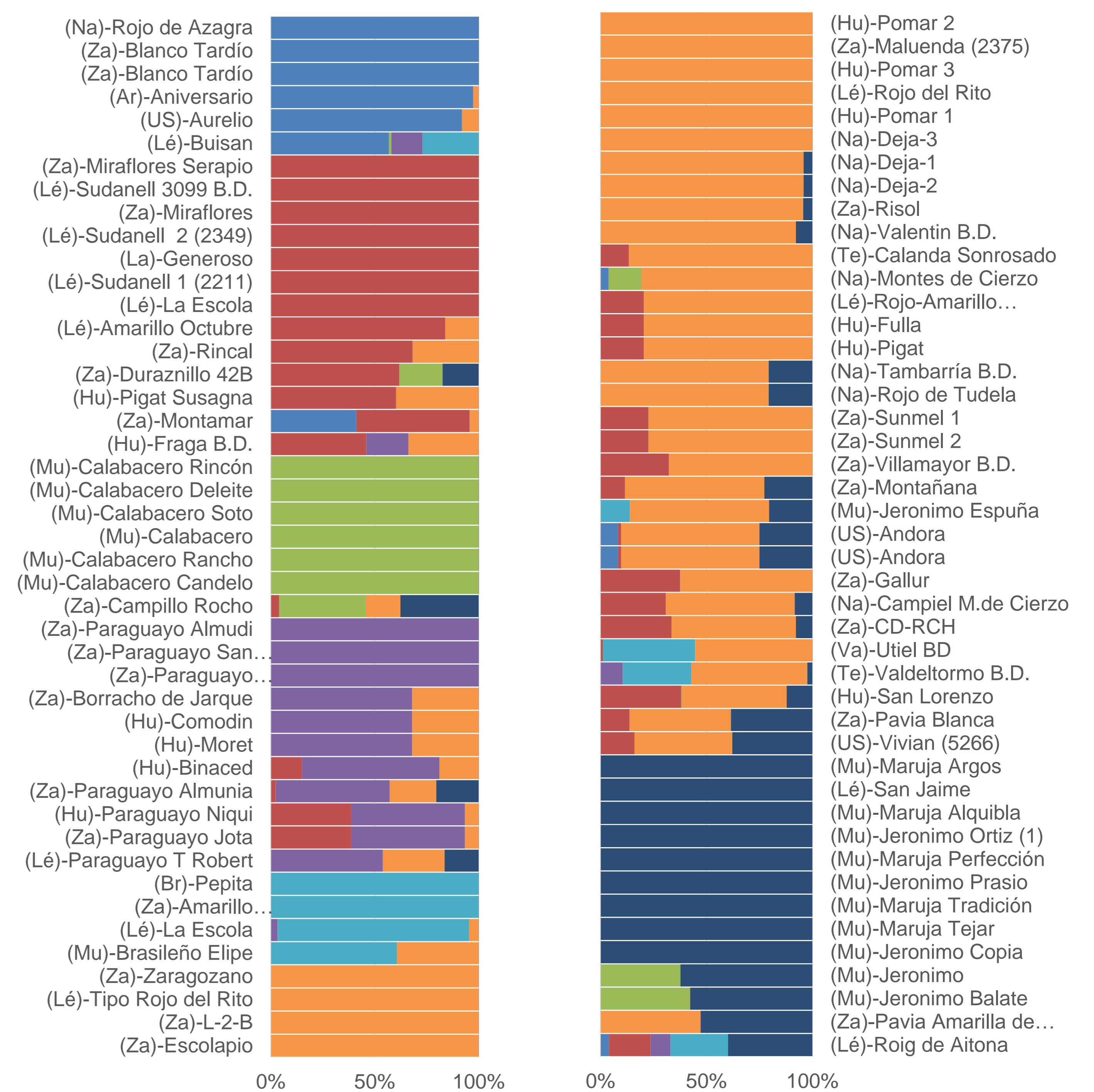

Figure 5. Genetic structure obtained in fastStructure (K=7).

CONCLUSIONS

According to our knowledge it is the first study using the Illumina Infinium Peach BeadChip 18K (18000 SNPs). The results have been useful to validate this new peach SNP chip version. The results obtained here established the genetic relationship among the different accessions, detected duplicates in the studied germplasm and possible mistakes of putative clones. Accessions from Murcia seem to have a different genetic background from northern accessions suggesting also that an exchange of genetic resources has happened among northern territories of Spain.

REFERENCES

Pérez, V., Larrañaga, N., Abdallah, D., Wünsch, A., & Hormaza, J. I. (2020). Genetic diversity of local peach (Prunus persica) accessions from La Palma island (Canary Islands, Spain). Agronomy, 10(4). https://doi.org/10.3390/agronomy10040457

Wambugu, P. W., Ndjiondjop, M. N., & Henry, R. J. (2018). Role of genomics in promoting the utilization of plant genetic resources in genebanks. In Briefings in functional genomics (Vol. 17, Issue 3). https://doi.org/10.1093/bfgp/ely014

Badenes, M. L., Cambra, M., López, M. M., Batlle, I., Iglesias, I., Aranzana, M. J., López, E., Vives, C., Garcia Brunton, J., Romeu, J., Rubio-Cabetas, M. J., Espiau, M., & Alonso, J. M. (2015). A peach germplasm collection for increasing the genetic diversity in European breeding programs. Acta Horticulturae, 1084. https://doi.org/10.17660/ActaHortic.2015.1084.15