UNIVERSITA' DEGLI STUDI DI MILANO

DISAA – Department of Agricultural and Environmental Sciences, Production, Landscape, Agroenergy of the University of Milan

PhD course in Agriculture, Environment and Bioenergy – XXIX cycle

COST-EFFECTIVE USE OF MOLECULAR MARKERS IN THE PRACTICAL RESOLUTION OF COMMON HORTICULTURAL CHALLENGES

Ph.D student: Dr. **Marco Passaro**

Supervisor: Dr. **Filippo Geuna**
Co-supervisor: Prof. **Daniele Bassi**

Doctoral Board Coordinator: Prof. **Daniele Bassi**

A.A. 2015/1016

Ai miei genitori…

"Grazie a tutti i miei colleghi per il supporto ed i preziosi consigli
A Gabriella che mi ha convinto a restare
A Marco, senza cui questo lavoro non avrebbe mai visto la luce"

INDEX

1 ABSTRACT

Genetic molecular markers (DNA markers) represent genetic differences between individual organisms or species placed directly into DNA sequence. They are widely used as powerful scientific instruments to accomplish different tasks, from genes mapping to forensic discrimination. The tremendous advance in DNA genotyping tools has lead to the development of impressive high-throughput technologies, such as Next Generation Sequencing platforms, that may revolutionize horticulture research and applications. However the cost of such technologies not always make them the most rationale approach, particularly when working on minor crop species or with large number of samples. The present work aims to the exploring a multi-purpose and cost-effective use of different kinds of molecular markers, for assisting fruit tree plants breeding and valorization. For this scope, three cases of study were presented, spanning from cultivar discrimination and phylogeny reconstruction to marker assisted selection (MAS) for Sharka resistance.

D.NA markers such as SSR and AFLP, were successfully used to discriminate the 'common' Chinotto from 'Chinotto di Savona', an uninvestigated traditional *Citrus* species cultivated in Liguria (italy) that is gaining increasing interest for the production of high-quality niche food and beverages. New polymorphisms on candidate genes, that could explain some of observed differences between the two accessions, were suggested.

SSR markers were used for the first time to the large-scale application of MAS on apricot (*Prunus armeniaca*) to boost the conventional breeding programmes. They were found new resistant breeding selections against the most important viral disease of stone fruits, Sharka, caused by Plum Pox Virus (PPV). Novel candidate accessions were also characterized for PPV-resistance, enriching and complementing the apricot germplasm available for breeding. Moreover the number of significant markers required for this task was reduced from seven to two, decreasing the overall cost, in terms of time and resources, usually required for the conventional breeding programmes.

A further reduction of resources for the application of MAS in apricot was achieved developing new SNP markers linked to Sharka resistance, and able to be screened using fluorescence on Real Time PCR machine with or without High Resolution Melting (HRM) technology.

The performed works demonstrate that the correct choice of molecular instruments together with the implementation of new techniques could easily led to cost-effective, time-saving, and reliable results even without the facility and resources reserved for main crops research and applications.

2 INTRODUCTION

The markers have been used over the years for the classification of plants. Markers are any trait of an organism that can be identified with confidence and relative ease, and can be followed in a mapping population. In other words, they can be defined as heritable entities associated with the economically important trait under the control of one or more genes (Beckman and Soller, 1986). For this reason they are usually called as genetic markers. A genetic marker can be defined as a chromosomal landmark or allele that allows for the tracing of a specific region of DNA, or more in details, as a way to mark a chromosome, a locus or a gene often associated with a valuable attribute, transmitted by the standard laws of inheritance from one generation to the next (Semagn et al., 2006). Genetic markers can be divided into two widely different classes: those based on visually assessable traits (morphological and agronomic traits), and those based on gene product or DNA assay (non-morphological markers).

2.1 Genetic molecular markers.

Genetic molecular markers (DNA markers) usually represent genetic differences between individual organisms or species placed directly into DNA sequence. DNA of an individual is unique and thus determines its identity. Differences between individuals lie in the nucleotide sequence of their DNA and these differences led to the development of such type of markers. Generally, they do not represent the target genes themselves but act as ‘flags’ able to reveal their positions and/or presence within a target genome. Moreover they do not affect the phenotype of the trait of interest because they are located only near to genes controlling that trait (they are linked to a specific trait) or, often, they just underline a specific position inside a chromosome. Moreover, a DNA marker can be polymorphic (reveals differences between individuals of the same or different species) o monomorphic (does not discriminate between genotypes). Polymorphic markers may also be described as co-dominant or dominant basing on the capacity of the markers to discriminate between homozygotes and heterozygotes genotypes. In all cases the specific genomic regions occupied by genetic molecular markers are called ‘loci’ (singular ‘locus’).

DNA markers are the most widely used type of marker predominantly due to various reasons in particular they are:

- unlimited in number and present in all the living organisms (Winter & Kahl, 1995)
- highly polymorphic
- dominant or co-dominant

- unaffected by pleiotropism, epistatic interactions, environments/abiotic stresses and developmental stage of the plants (Winter & Kahl, 1995)
- have a Mendelian inheritance
- easily reproducible by different laboratories
- selectively neutral because usually located in non-coding regions of DNA

DNA markers arise from different classes of DNA mutations such as substitution mutations (point mutations also called Single Nucleotide Polymorphism, or SNP), rearrangements (insertions or deletions) or errors in replication of tandemly repeated DNA (Paterson, 1996a).

One of the most important uses of genetic molecular markers regards the development of linkage maps (a map that indicates the position and relative genetic distances between markers along chromosome) to perform Quantitative Trait Locus analysis (QTL) to discover the loci involved (and thus the genes) in a trait of interest (Paterson, 1996a). However they have also numerous applications in plant breeding and cultivar/species discrimination by assessing the level of genetic diversity among different plants (Baird et al., 1997; Henry, 1997; Jahufer et al., 2003;Weising et al., 1995;Winter & Kahl, 1995). Moreover, genetic markers associated to genes or loci that carry agronomical important traits are often used as substitutes to the phenotypic selection in the plant breeding processes, enabling the breeders to make the conventional programs more efficient, quick and cost-effective, opening the way to marker assisted selection (MAS) (Rafalski & Tingey, 1993; Ribaut & Hoisington, 1998).

DNA markers can be divided into three classes based on the method of their detection: hybridization-based, polymerase chain reaction (PCR) based and DNA sequence-based (Jones et al., 1997; Joshi et al., 1999;Win- ter & Kahl, 1995).

2.1.1 Hybridization-based markers.

2.1.1.1 Restriction Fragment Length Polymorphism (RFLP)

RFLP markers were first used in 1975 to identify DNA sequence polymorphisms for genetic mapping of a temperature-sensitive mutation of adeno-virus serotypes (Grodzicker et al., 1975). They was then used for human genome mapping (Botstein et al., 1980), and later adopted for plant genomes (Helentjaris et al., 1986; Weber and Helentjaris, 1989). As most part of the markers, RFLP is based on differences in the DNA sequence occurred during its replication due to the action of many mechanisms (Joshi et al , 2011). Mutations are usually inherited to progenies and fixed into populations as different alleles, in particular when those mutations occur into the non-coding

regions of genome. Consequently, two individuals of the same species will always differ at a few nucleotides due to one or more of the following causes: point mutation, insertion/deletion, translocation, inversion and duplication. Some of the differences in DNA sequences at the restriction sites can result in the gain, loss, or relocation of a restriction site. Hence, digestion of DNA with restriction enzymes results in fragments whose number and size can vary and thus represents different alleles among individuals, populations, and species. RFLP is based on these principles and follows 5 steps:

- digestion of the DNA with restriction enzyme (one ore more)
- separation of the restriction fragments on agarose/acrylammide gel
- transfer of the fragments from the gel to a filter by Southern blotting (Southern, 1975)
- detection of individual fragments by hybridization with labelled probe(s)
- autoradiography (Perez de la Vega, 1993; Terachi, 1993; Landry, 1994) or non-radioactive methods (Holtke et al., 1995; Mansfield et al., 1995).

The choice of which restriction enzymes to use (digestion step), represents a central aspect when using this type of markers. The greatest resolution is obtained by using 'four-cutters' (enzymes recognizing a four base pair sequence) because there are many such sites in the genome and the fragments produced are small, numerous and provides a better chance of identifying single base mutations. Using ‘eight-cutters’ (enzymes recognizing a eight base pair sequence) will produce fewer fragments with bigger length providing information only about large alterations of DNA. Six-cutters (enzymes recognizing a six base pair sequence) are the most used compromise producing fragment in the range of 200-20000 bp, which can be easily separated on gel providing both type of information (Potter and Jones, 1991).

The probes used for hybridization are preferably single locus, species-specific (Staub and Serquen, 1996) and are generated from genomic clones (fragments of nuclear DNA) or cDNA clones (DNA copies of mRNA molecules). Genomic-derived probes are easy to construct but contain many duplicates due to the repetitive nature of genome thus they will hybridize into many fragments on the filter producing very complex patterns. cDNA probes are not easy to produce but usually provide fewer bands representing only expressed genes. Therefore, the selection of appropriate source for RFLP probes represents another factor that must be taken into account.

The major strength of RFLP markers is due to their high reproducibility, co-dominant inheritance and good transferability between laboratories. However there are also several limitations, such as the requirement of high amount of starting DNA (Potter and Jones, 1991; Roy et al., 1992; Young et al., 1992), the dependence from specific probe libraries for the species and

the large requirements in terms of time and costs. RFLP markers were successfully used on genetic linkage mapping studies and QTL analysis in fruit crops (Rajapakse et al., 1995; Wang et al., 2000).

2.1.2 PCR-based markers

The discovery of polymerase chain reaction (PCR) method of DNA amplification by Mullis et al. (1986) opened the way to the 'modern era' of molecular biology. PCR is an in-vitro method for enzymatic amplification of a specific DNA segment from the genomic DNA. Two oligonucleotide primers, flanking the genomic region of interest (up to 10 Kb), allow the specific amplification of the amplicone by a series of repeated cycles of heat denaturation, annealing of the primers to the complementary sequence and their extension through the action of a thermophilic DNA polymerase (Taq). Due to the nature of the process, each PCR cycle double the amount of the target DNA synthetized in the previous cycle resulting in an exponential accumulation of the target of interest, allowing its visualization on gels or by fluorescence. If the amplicone selected contains differences between different individuals or species (in terms of length or base composition), it's possible to score them following various methods, from electrophoresis to DNA sequencing.

2.1.2.1 Random Amplified Polymorphic DNA (RAPD)

RAPD markers use PCR to amplify DNA fragments of any species without prior knowledge of sequence information. Indeed RAPD technique are normally executed using 10 bp long primers that must contains at least 40-50% of GC base pairs composition avoiding the presence of palindromic sequences (Williams et al. 1990). Due to the arbitrary nature of primers, low GC% contents and palindromic sequences prevent them to stay linked to the stamp DNA during the 72° C extension step enabling the PCR reaction (Williams et al. 1990). The resulting PCR products are usually resolved on 2.0% agarose gels and stained with ethidium bromide (EtBr), polyacrylamide gels in combination with either silver staining (e.g., Huff et al., 1993; Vejl, 1997; Hollingsworth et al., 1998), radioactivity (e.g., Pammi et al., 1994), or fluorescently labeled primers or nucleotides (e.g., CorleySmith et al., 1997; Weller and Reddy, 1997). Generally, the output of RAPD depends on the fact the one polymorphism could enabling the correct annealing of primers in a single locus resulting in a loss of bands visualized through the gel. Samples are thus characterized by the presence or absence of bands in the same locus and so RAPD are considered dominant markers. However, bands of different intensity may result from copy number of the considered allele (Devos and Gale, 1992) and may serve to distinguish homozygote dominant individuals from heterozygotes. However, some authors (Thormann et al., 1994) found no correlation between copy

number and band intensity. Main factors that influence RAPD methods are the quality and quantity of template DNA, PCR buffer, concentration of $MgCl_2$ and the annealing temperature.

Advantages associated with RAPD analysis include:

-needs of small amount of DNA

-fast and efficient

-no radioactive assay needed

-no specific probes needed

-no blotting required

However, many negative aspects have to be taken into account:

-dominant nature of RAPD cause a loss of information compared with co-dominant markers

-shortness of primer could generate a lot of false positive and false negative

-low temperatures needed during the annealing steps of PCR prevent its reproducibility especially when transferring it between population or laboratory (Liu et al. 1994)

Moreover a pairwise comparison of RAPD fragments along samples begins with the assumption that co-migrating bands represent homologous loci. However the assumption that equal length means equals homology may not be necessarily true (Thormann et al., 1994; Pillay and Kenny, 1995).

RAPD were successfully used in many fruits studies (for example peach and almond) such as cultivars discrimination (Lu et al., 1996), genetic diversity (Warburton et al., 1996) and genetic relatedness among breeding lines (Bartolozzi et al., 1998).

2.1.2.2 Sequence Characterized Amplified Region (SCAR)

SCAR markers often derived by cloning and sequencing the two ends of a previously known PCR marker (for example RAPD markers). The specific DNA fragment identified is scored by PCR amplification using a pair of specific oligonucleotide primers (Paran and Michelmore, 1993; McDermott et al., 1994). This method has been often used when a RAPD marker appeared to be diagnostic for specific purposes, as for example prediction of a status against a disease, allowing the screening of that locus in other samples avoiding all the issues linked with RAPD technology (Paran and Michelmore, 1993).

2.1.2.3 Amplified Fragment Length Polymorphism (AFLP)

AFLP represents a successful combination of the RAPD and RFLP methods (Farooq and Azam, 2002) mainly for its capacity to give a “whole genome representation” (the simultaneous

screening of representative DNA regions distributed randomly throughout the genome) of any species without prior knowledge of sequence information. It is thus considered a powerful DNA fingerprinting technology and it is first developed by Vos et al. in 1995. AFLP technique is based on PCR amplification of a set DNA fragments previously digested with a pair of specific restriction enzymes, one of them usually being a frequent cutter (generates small DNA fragments that will amplify well and are in the optimal size range for separation on gel) and the other a rare cutter (reduces the number of fragments to be amplified). The restricted DNA is ligated with small double stranded oligonucleotides adaptors that recognize the specific cutting ends of the restricted fragments basing on the enzymes used. The known adaptors sequences allow the two subsequent PCR amplifications of the restricted fragments. The first PCR step (pre-amplification) is performed with primer combinations containing the adaptors sequence plus a single bp extension chosen randomly, resulting in the first selective amplification of the fragments previously generated. The PCR products from this pre-amplification step are diluted and used as template for the second PCR amplification (selective) that use primer pairs with up to 3-bp extension. A primer extension of one, two or three bases reduces the number of amplified fragments by factors of 4, 16 and 64, respectively allowing their visualization on gels or by capillary electrophoresis. Because of the high selectivity, primers differing by only a single base in the AFLP extension amplify a completely different subset of fragments resulting in a different fingerprinting pattern. The choice of how many base pair extension to use depends on the size of the species of interest. However the optimal numbers of bands have to be sufficient to provide adequate polymorphisms without causes smears or high levels of co-migrating bands. As mentioned before AFLP fragments are visualized either on agarose gel or on denaturing polyacrylamide gels with autoradiography, $AgNO_3$ staining or automatic DNA sequencers.

The advantages of AFLP are more or less similar of that of RAPD excluding the fact that they are more laborious. Other advantages of this method are:

- highly reliability and reproducibility (Mueller et al., 1996; Lin et al., 1996; Powell et al., 1996; Jones et al., 1997).
- allows the analyses of a large number of polymorphic loci simultaneously with a single primer combination on a single gel (Powell et al., 1996; Milbourne et al., 1997; Russell et al., 1997).

Common disadvantages include:

- the multi-step requirement of the procedure.
- the use of polyacrylamide gel in combination with silver staining or fluorescent methods of detection, which will be more expensive and laborious than agarose gels (see RAPD).

- AFLP loci are usually dominant, reducing the accuracy on genetic population and genetic mapping studies.

AFLP were successfully used in many studies such as cultivar identification (Geuna et al., 2003) and linkage mapping (Vilanova et al., 2003) in apricot.

2.1.2.4 Sequence Tagged Site (STS).

STS is a short, unique DNA fragment whose exact sequence is found nowhere else in the target genome. STS can be derived from any clone previously isolated with other markers as for example RFLP (Blake et al., 1996) or AFLP (Shan et al., 1999; Prins et al., 2001). As previously described for SCAR, STS are developed when a particular unique region appeared to be diagnostic for specific purposes (e.g., an AFLP band present in a cultivar but absent in another one).
It is scored by PCR amplification using a pair of specific oligonucleotide primers. STS markers are codominant, highly reproducible, suitable for high-throughput and automation, and technically simple for use (Reamon-Buttner and Jung, 2000).

2.1.2.5 Cleaved Amplified Polymorphic Sequence (CAPS)

CAPS are based on a combination between PCR-based markers and restriction-based markers. They were originally developed by Maeda et al. (1990). The method works on the same principle seen for RFLPs, allowing an easy recognition of nucleotide polymorphisms that inactivate a target restriction site. CAPS involve a prior amplification of a target DNA through PCR, followed by digesting with restriction enzymes (Konieczny and Ausubel, 1993; Jarvis et al., 1994; Michaels and Amasino, 1998). The result of digestion it is then visualized on agarose gel (usually 4%). CAPS basically bring together most of the advantages of the RFLP and PCR-based markers avoiding some of the common problems linked with restriction enzymes as the use of time-consuming methods for scoring polymorphisms (blotting or autoradiography) and the problem linked to the co-migration of bands. Moreover, they are co-dominant and are inherited in a co-dominant manner (Matsumoto and Tsumura, 2004). However CAPS have also some problematic aspects as:
- the polymorphisms rate on genome is not as high as SSRs and AFLPs
- the development is only possible where mutations disrupt or create a restriction site
To overcome this last aspect some researchers developed an alternative marker called derived-CAPS (dCAPS) that eliminate this problem by generating mismatches in a PCR primer, which are subsequently used to place the mutation in the context of a restriction site (Michaels and Amasino,

1998; Neff et al., 1998). Moreover EST-CAPS was extensively used for comparative mapping study (see EST).

2.1.2.6 Microsatellites

The genomes of higher organisms contain tree types of multiple copies of simple repetitive DNA sequences (satellite, minisatellites, and microsatellites) arranged in arrays of differing size (Armour et al., 1999; Hancock, 1999). Microsatellites represent the smallest class among the repetitive DNA distributed in the genome, and are usually divided into three different classes: Simple Sequence Repeats (SSR) (Tautz et al., 1986), Short Tandem repeats (STRs) and Simple Sequence Length Polymorphisms (SSLPs) (McDonald and Potts, 1997). Is not easy to find a standard definition of microsatellites even if most of authors define it as genome region of simple repetitive DNA composed by a tandem repetition of a small base pairs motif. Some authors (e.g. Armour et al., 1999) define microsatellites as 2–8 bp repeats, others (e.g., Goldstein and Pollock, 1997) as 1–6 or even 1–5 bp repeats (Schlotterer, 1998), however Chambers and MacAvoy (2000) suggest to consider as the standard definition of microsatellites a repetition of 2-6 bp.

Usually microsatellites are born into genomic region already over represented by repetitive DNA motif (Tautz et al., 1986) by the well-studied mutation mechanism called "slipped-strand mispairing" that occurs during DNA replication (Levinson and Gutman, 1987; Eisen, 1999). When slipped-strand mispairing occurs within a microsatellite array, it can result in the gain or loss of one, or more, repeat units. These mutations represent therefore the allelic differences underlined by SSR markers that usually shows high levels of inter- and intra-specific polymorphism, particularly when tandem repeats number is ten or greater (Queller et al., 1993). SSR detection requires a simple PCR reaction using a primer pairs designed upstream and downstream the repetitive DNA. Thus different alleles will results in amplicons with different length easily visualized through agarose gel (usually 3%), acrylamide gels or automatic sequencer if fluorescent probes were previously used. This method is more or less shared between different protocols developed in the last decades (Bruford et al., 1996; McDonald and Potts, 1997; Hammond et al., 1998; Schlotterer, 1998).

Unlike most of the previous methods described, SSRs require prior sequence information about the loci of interest. Thus the development of microsatellites markers is not easy and involves several distinct steps:

- microsatellite library construction
- identification of unique microsatellite loci
- identification of a suitable area for primer design

During the '90s, SSR primers were developed by cloning random segments of DNA from the target species into *Escherichia coli*. Colonies are then screened with simple sequence oligonucleotide repeat probes that permit to isolate positive clones. Following DNA sequencing it is possible to design PCR primers flanking microsatellites regions to specifically tag a genomic locus. This process involves significant trial and error as microsatellite repeat sequences must be predicted and primers that are randomly isolated may not display polymorphism (Queller et al., 1993; Jarne and Lagoda, 1996). Indeed the SSRs obtained by this procedure show a very low conversion rate from the development of primers to a successful identification of a functional and polymorphic locus. Today the great amount of sequence information of various species, together with the new bio-informatics tools, allows a more efficient development of SSR markers (see below).

Researchers often prefer to work with SSR markers that underline loci containing tri- and tetra-nucleotide repeats rather than di-nucleotide repeats because the former frequently give fewer 'stutter bands' (multiple near-identical fragments of PCR products which are one or two nucleotides shorter or longer than the full-length product) causing frequent errors during the allele-sizing process (Hearne et al., 1992; Diwan and Cregan, 1997). However di-nucleotide SSRs are more frequent into genomes and not always tri- or tetra-nucleotide SSR are available.

SSRs are now the marker of choice in most areas of molecular genetics for the following reasons:

- high polymorphism
- low amount of DNA required
- high transferability between populations and laboratories
- co-dominance, proving excellent for studies of population genetics and mapping (Jarne and Lagoda, 1996)
- ease of preparation for high-throughput screening using fluorescent primers in combination with automatic analyzers.

However, differences in SSR allele size is often difficult to resolve on agarose gels and high resolutions can be achieved through the use of polyacrylamide gels or, better, automated capillary electrophoresis that is not always affordable for small or medium research facilities.

2.1.2.7 Inter-Simple Sequence Repeat (ISSR)

ISSR involves amplification of DNA segments present at an amplifiable distance between two identical microsatellite repeat regions oriented in opposite direction and it was developed by Zietkiewicz et al. (1994). ISSR marker uses the microsatellite sequence itself as primers allowing multiple loci amplification inside the genome in a single PCR reaction. The microsatellites

sequence used could be di-, tri- tetra- or penta-nucleotide and moreover the primers used could be either unanchored (Meyer et al., 1993; Gupta et al., 1994; Wu et al., 1994) or more usually anchored at the 3` or 5` end with 1-4 degenerate bases extended into the flanking sequences (Zietkiewicz et al., 1994). Anchored primers led to a more selective reaction with a consequent reduction of the migrating band in the final pattern. ISSR primers (15-30) are longer than RAPD primers allowing higher annealing temperature and increasing their reproducibility. They also do not require any prior sequence information and are easier and quicker to use compared to AFLP. Like RAPDs, reproducibility, dominant inheritance and homology of co-migrating amplification products are the main limitations of ISSRs (Gupta et al., 1994; Tsumura et al., 1996; Ratnaparkhe et al., 1998; Wang et al., 1998; Wu et al., 1994; Akagi et al., 1996; Wang et al., 1998; Sankar and Moore, 2001). Even if it seems that ISSR could easily replace AFLP as a whole genome fingerprinting method, it must be taken into account that all the genomic regions without microsatellite repeat sequences are excluded from the ISSR assay and not from AFLP.

2.1.2.8 Expressed Sequence Tags (EST)

EST is not properly a molecular marker itself but represented one of the most useful tools for markers development. The production of ESTs started with the construction of cDNA libraries. Complementary DNA, or cDNA, is a double-stranded DNA synthesized from a single stranded RNA (messenger RNA or microRNA) template in a reaction catalysed by the enzyme reverse transcriptase. Once cDNA, that usually represents an expressed gene, has been isolated, a few hundred of nucleotides from either the 5' or 3' end could be sequenced to create 5' expressed sequence tags (5' ESTs) and 3' ESTs, respectively (Jongeneel, 2000). A 5' EST is obtained from the portion of a transcript (exons) that usually codes for a protein and thus this kind of regions tend to be conserved across species and gene families. The 3' ESTs include non-coding (introns) and untranslated regions (UTRs), and therefore tend to exhibit less cross-species conservation.

Surprisingly, even if ESTs were originally intended as a way to identify gene transcripts they rapidly became a valuable instrument for the development of EST-based molecular markers as for example EST-RFLPs, EST-SSRs, EST-SNPs and EST-CAPSs. EST-based RFLPs have been extensively used for the construction of high-density genetic linkage maps (e.g., Harushima et al., 1998; Davis et al., 1999). Moreover different bioinformatics tools were developed in order to find suitable SSRs or SNPs inside ESTs databases, that today count more than 20 millions sequences from different species (https://www.ncbi.nlm.nih.gov/genbank/dbest/dbest_summary/). 1 to 5% of the ESTs in various plant species have been found to have SSRs of suitable length (20 bp or more) for marker development (Kantety et al., 2002). Moreover EST-SSRs also have a higher probability

of being functionally associated with differences in gene expression than the genomic SSRs (Gao et al., 2004) and hence they are expected to be more transferable to closely related genera (Cordeiro et al., 2001; Hempel and Peakall, 2003; Decroocq et al., 2003). Most of the EST-SNPs were found by comparing computationally the 3' UTRs of ESTs coming from different cultivars to maximize the probability to find differences.ESTs were also extensively used as DNA probes for the development of microarrays involved in gene expression studies in different species.

2.1.2.9 Single Nucleotide Polymorphism (SNP)

Starting from the 80's the public accessibility to the genome sequences of several organisms (e. g. ESTs) has enabled the study and identification of sequence variations, as for example SNPs, between individuals, cultivars, and species. Many studies revealed that SNP and other minor DNA mutation such as insertion and deletions of single nucleotides (InDels) are highly abundant (virtually infinite) and distributed throughout the genome in various species including plants (Garg et al., 1999; Drenkard et al., 2000; Nasu et al., 2002; Batley et al., 2003a) making them an attractive tool for mapping, marker-assisted breeding and map-based cloning (Gupta et al., 2001; Rafalski, 2002a; Batley et al., 2003b).

As suggested by the acronym, SNP marker consists of just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. For this reason, in contrast with the methods previously described, the SNP identification and allele discrimination usually cannot be based on size differences on a gel. Over the past years, a large number of different SNP identification/genotyping methods have been developed starting from methods based on hybridization to those involving the high throughput sequencing of DNA (Semagn et al., 2006; Ganal et al., 2009).

2.1.2.10 SNP identification

There are several SNP identification techniques that are used for the identification of large numbers of SNPs in particular in plants.

As previously seen (see 'EST'), a large number of ESTs have been generated for many plants, including models such as *Arabidopsis thaliana* and crop species. The number of available ESTs ranges from 2 millions for the main crops species (maize) to 15000 for the little investigated crops (apricot) (https://www.ncbi.nlm.nih.gov/genbank/dbest/dbest_summary/). In the most part of the cases, ESTs libraries were created for gene identification and expression studies, however the presence of ESTs from different lines or closely related species provided the opportunity to

generate many molecular markers including SNP. Also the ESTs coming from heterozygous highly polymorphic individuals were used for SNP identification using bioinformatic analysis methods (Pavy et al., 2006). In some cases, ESTs have been specifically generated for SNP identification in different lines as for example on *Arabidopsis thaliana* (Schmid et al., 2003). However ESTs do not represent the best tool for SNP identification for many reasons. Their sequence quality is usually very low and this represents a central issues because sequencing errors could be mistakenly considered as true SNP mutations. To overcome this, several ESTs from each of the compared lines must be available for the same gene in order to reliably identify a SNP and this is not always possible, because the number of ESTs that can be compared with each other is limited because of the different expression level of genes. As a consequence the number of identified SNP from ESTs is relatively low for many species with a validation rates usually under 50%.

Another approach for SNP identification that is also based on ESTs, involves the use of arrays containing oligonucleotides derived from large numbers of genes. This kind of arrays, like ESTs, were originally created for comparative expression studies of individual genes, but were soon used also for the identification of SNPs when the hybridization patterns generated with cDNA or DNA samples from different individuals are being compared. They are in this case termed single feature polymorphisms (SFPs). Examples for the identification of SFP in large numbers have been published for Arabidopsis using various arrays (Borevitz et al., 2003; Singer et al., 2006), rice (Kumar et al., 2007), barley (Cui et al., 2005; ,Rostoks et al., 2005), and maize (Krist et al., 2006). Moreover this kind of arrays can also be used for closely related species as for example demonstrated through the use of a soybean genome array for the identification of SFPs in cowpea (Das et al., 2008). However, as for the ESTs, this approach has a high false discovery rate.

A better but more laborious method to identify SNP consists in the single amplicon re-sequencing. In summary, it involves the design of primers for the amplification of DNA fragments derived from genomic sequences trough PCR reaction. In this case the fragments amplified from different lines or species were aligned and compared using bioinformatics tools and thus all the mutations were discovered. Moreover using this approach the sequence of each investigated sample is determined trough double-strand sequencing allowing a double control. With amplicon re-sequencing SNPs can be identified in a very reliable way with a false discovery rate usually significantly below 5%. However this approach requires an enormous effort for the analysis of many loci because for each fragment, specific primers have to be developed in multiple lines or species.

Finally, sequenced genomes can be used in several ways for the identification of large numbers of SNPs. In case of heterozygous species, SNPs can be directly mined in the genomic

sequence since in fact two genome sequences have been generated. In other cases a sequenced genome must be compared by other sequenced genomes (e. g. other cultivars or other species) or by genomic libraries (see 'NGS').

2.1.2.11 SNP genotyping assays

SNP based genotyping was usually divided into four kinds of assays (Sobrino et al., 2005):

- allele specific hybridization
- primer extension
- oligonucleotide ligation
- invasive cleavage

Allele-Specific Oligonucleotide (ASO), also known as "specific oligonucleotide hybridization" is based on the differential hybridization of two allelic-specific probes on the target DNA (Wallace et al., 1979). The two probes sequences contain the SNP usually in their central position and, under optimized assay conditions, only the fully complementary probe will hybridize in a stable manner giving a signal, instead hybrids with one-base mismatch are unstable and thus will not give any signal. The common way to score this kind of assays consists in the use of the Dot Blot. This technique, also known as "slot blot" is used to detect DNA and other biomolecules and represents a simplification of the Northern, Southern, or Western Blot methods. Unlike these methods, in a Dot Blot the DNA to be tested (either genomic, cDNA or a PCR product) are not separated by electrophoresis. Instead, the molecule to be detected is fixed directly on a membrane as a dot, and then it is spotted through hybridization by circular templates directly onto the membrane or paper substrate. In the Reverse Dot Blot technique, it is a oligonucleotide probe that is immobilised on the membrane. However, hybridization techniques are error prone and need carefully designed probes and hybridization protocols (Pastinen et al., 1997). The latest improvement of this family of techniques is represented by DNA chips (collection of microscopic DNA spots attached to a solid surface), on which the probes are directly synthesised (Pease et al., 1994).

Primer extension is based on the ability of DNA polymerase to incorporate specific deoxyribonucleotides (dNTPs) complementary to the sequence of the template DNA and there are at least three different main methods that use that principle, the 'mini-sequencing', the 'allelic-specific extension' and the 'pyro-sequencing'.

In the 'mini-sequencing' techniques a primer anneals to its target DNA immediately upstream to the SNP and is extended with a single nucleotide complementary to the polymorphic base. These formats use a wide range of detection techniques that include Mass spectrometry or the

incorporation of either fluorescently labeled dideoxynucleotides (ddNTP) or fluorescently labeled dNTPs.

The allelic-specific extension is based on the fact that only a perfect match between primers and the target sequence allows its extension by the DNA polymerase. In this case one of the primer for the PCR reaction is placed across the SNP mutation. Only the samples with no mismatch between the primer used and the target sequence will results in a correct amplification discriminating the different alleles.

The last method consists in the Pyro- (Ronaghi et al., 1996) or Sanger-sequencing (Sanger et al., 1977) of the target sequence for its direct scoring.

Allele specific oligonucleotide ligation assay (OLA) is a method for SNP typing based on the ability of ligase to covalently join two oligonucleotides when they hybridize next to one another on a DNA template. Two allelic-specific probes and one common ligation probe are designed for each SNP. The common ligation probe hybridized adjacent to the allelic-specific probe. When there is a perfect match of the allelic-specific probe, the ligase joins together both allelic-specific and common probes. In the other case the ligation does not occur.

The invader assay is based on the specificity of recognition, and cleavage, by a flap endonuclease, of the three-dimensional structure formed when two overlapping oligonucleotides hybridize perfectly to a target DNA (Kaiser et al., 1999; Lyamichev et al., 1999). This method requires two oligonucleotides called respectively invader probe and allelic-specific probes. The invader probe anneal to the target DNA with an overlap of one nucleotide in the exact position of the target SNP mutation. When the allelic-specific probe is complementary to the SNP, overlaps the 3' end of the invader oligonucleotide, forming the structure that is recognized and cleaved by the Flap endonuclease, releasing the 5' arm of the allelic-specific probe. On the other case the formed structure is not recognized by the endonuclease and thus there is not any cleaved fragment.

Independently from the chosen SNP assay,

There are several detection methods for analyzing the products of each type of allelic discrimination reaction:

- gel electrophoresis
- fluorescence resonance energy transfer (FRET)
- fluorescence polarization,
- arrays or chips,
- luminescence,
- mass spectrophotometry

2.1.3 Next Generation Sequencing (NGS)

Sanger dideoxy sequencing (Sanger, 1977) dominated the DNA sequencing field for nearly 30 years and in the past 10 years the length of Sanger sequence reads has increased from 450 bases to more than 1 kb. However, when the new high-throughput sequencing technologies started to be available, Sanger method showed its own limitation:

- the process involves capillary electrophoresis to separate the elongation products making it unable to handle high throughput technologies;
- before sequencing it is needed to produce clonal populations of DNA using *Escherichia coli*, or to produce PCR products, which is labor-, robotics- and space-intensive for large-scale operations;
- performing the sequencing reactions in reduced reaction volumes can reduce the cost per sample, but the fundamental restrictions on reducing the cost of Sanger sequencing are at their technological limits.

Advances made in different scientific field as nanotechnologies and informatics, allow the development of alternative methods to increase the rapidity and/or throughput of DNA sequencing, giving birth to the so-called NGS technologies. Next generation sequencing relies on massively parallel sequencing and imaging techniques to yield several 100s of millions to several 100s of billions of DNA bases per run (Shendure and Ji, 2008). Today commercially available NGS technologies are 'Roche/454' (www.454.com), 'Solexa/Illumina' (www.illumina.com) and AB SOLiD (www.appliedbiosystems.com). Currently, Roche/454, Solexa and AB SOLiD are the technologies that are predominantly used in crop genetics and breeding applications.

All NGS strategies follow a similar protocol for DNA template preparation, where universal adapters are ligated at both ends of randomly sheared DNA fragments. They also rely on the cyclic interrogation of millions of clonally amplified DNA molecules immobilized on a synthetic surface to generate up to several billions of sequences. Sequencing is performed in an iterative manner, where the incorporation of one or more nucleotides is followed by the emission of a signal and its detection by the sequencer (Metzker, 2010). As example, considering Illumina technology, DNA molecules and primers are first attached on a slide and amplified with polymerase so that local clonal DNA colonies are formed (up to 1000 copies of the original molecule for each single colony). To determine the sequence, four types of proprietary reversible fluorescent terminator deoxyribonucleotides, defined as RT-bases, are added and non-incorporated nucleotides are washed away. A camera takes images of the fluorescently labeled nucleotides, then the dye, along with the terminal 3' blocker, is chemically removed from the DNA, allowing for the next cycle to begin. The DNA chains are extended one nucleotide at a time and image acquisition can be performed at a delayed moment, allowing for very large arrays of DNA colonies to be captured by sequential

images taken from a single camera (Mardis, 2008). In other words, the Illumina sequencing technology combines clonal amplification of a single DNA molecule with a cyclical sequencing-by-synthesis approach producing ultra-high throughput sequence data compared to Sanger sequencing (Pareek et al., 2011). However contrarily to Sanger methods, NGS can generate fragments that range from 50 to 300 bp.

The increased ability to sequence in a cost-effective manner large numbers of individuals within the same species has altered the concept of variant discovery and genotyping in mapping studies, especially in plant species with complex genomes or limited public resources available. A new concept, namely genotyping-by-sequencing (GBS), has emerged, where the detection of SNPs in a large segregating or mutant population is combined with scoring, thus allowing a rapid and direct study of its diversity without any prior sequence knowledge. GBS uses restriction enzyme to reduce genome complexity, then the resulting fragments are subjected to a PCR reactions to increase theirs concentrations and finally they are sequenced using next generation sequencing technologies, usually resulting in about 100 bp single-end read, and analysed using bioinformatics tools.

A similar concept is shown by the 'Restriction Site Associated DNA sequencing' technology often defined as RAD-seq. Rad-sec is based on the process of isolating RAD tags, which are the DNA sequences that immediately flank each instance of a particular restriction site of a restriction enzyme throughout the genome. Once RAD tags have been isolated, they can be used to identify and genotype DNA sequence polymorphisms mainly in form of SNPs. Polymorphisms that are identified and genotyped by isolating and analysing RAD tags are referred to as RAD markers.

NGS technologies together with GBS approaches represented a revolutionary step into molecular biology, allowing the generation of larger data set for mapping and diversity studies in each type of organism at any complexity level (inter-, intra-specific) in a rapid, effective, and low-cost manner. Until now, NGS technologies have been used for whole genome sequencing and for re-sequencing projects, for SNPs and InDels identification, for exploring the diversity, constructing haplotype maps performing genome-wide association studies and Marker Assisted Selection (MAS) (Elshire et al., 2011).

3 AIM AND SCOPES

Aim of this thesis is a multi-purpose and cost-effective use of molecular markers for assisting fruit tree plants breeding and valorization. Molecular markers were used for different purposes: SSR and AFLP markers for cultivars identification and phylogeny reconstruction within *Citrus myrtifolia* species (better known with the Italian name 'Chinotto' and 'Chinotto di Savona'); SSR markers for the screening of a major locus conferring Sharka resistance in a wide panel of apricot (*P. armeniaca*) cultivars, accessions and selections; SNP-based genotyping using HRM technology for quick and cost-effective Marker-Assisted Selection (MAS) in apricot.

4 CHARACTERIZATION OF 'CHINOTTO DI SAVONA' *CITRUS*: PRELIMINARY RESULTS

4.1 Introduction

Citrus myrtifolia, better known with the Italian name 'chinotto', belongs to the *Citrus* genus of the *Rutaceae* family. Classified in 1961 (Tanaka, 1961), *C. myrtifolia* is a flowering tree, about 3 m tall, with small leaves that resemble those of common myrtle, hence the Latin name (Hanelt et al., 2001). The flowers are white and very fragrant. Unripe fruits look small and green (2-3 cm diameter) and are mainly used in the food industry for candies and jam. Mature fruits develop a deep orange color, and are used as an essential flavor component in many soft drinks because of their peculiar bitter and sour taste. Native from China *C. myrtifolia* has been imported in Italy during the 15th century in the town of Savona (Liguria, Italy). Currently, it is cultivated in several Italian regions (Liguria, Tuscany, Calabria, and Sicily) as well as in France and in the "*Citrus* region" of USA.

Despite the growing diffusion of 'Chinotto'-derived food and beverages, only few studies about its agronomical and pomological characteristics are available in literature. An early study reported the presence of at least 51 chemical compounds in the essential oil of *C. myrtifolia* dried peel (Chialva et al., 1990). 'Chinotto' juice is characterized by a high content of aspartic acid and proline, and a lower acidity than the sweet orange (*C. sinensis*) juice (Cautela, 2004). Fingerprinting analysis of flavonoids compounds in *C. myrtifolia* juice revealed a close similarity with *C. bergamia* (bergamot), supporting their common origins from *C. aurantium* (sour orange) (Tanaka, 1961, Hodgson, 1967), remarking also wide differences from other *Citrus*, such as *C. limon* (lemon) and *C. sinensis* (Cautela et al., 2004; Barreca et al., 2010). 'Chinotto' fruit is known for its nutraceutical properties (Protti et al., 2015), the very high antioxidant activity, in particular of the albedo and flavedo tissues (Barreca et al., 2011), and for the antiproliferative action against some human cancerogenic cells lines (Camarda et al., 2007). The increasing evidences about healthy value of 'Chinotto'-derived foods and beverages are supporting its commercial valorization in traditional productive areas.

The genetic origin of *C. myrtifolia* is still controversial, as well as the taxonomy of *Citrus* genus, as a consequence of a large morphological diversity, sexual interspecific compatibility (also between related genus), partial apomixis and several centuries of cultivation (Garcia Lor. et al., 2012). There are two major systems to classify *Citrus* species: the Swingle and Reece (1967) and the Tanaka (1977) classifications. The first counts 16 species and it is widely used by the scientific

community because in good agreement with molecular data; the second is less used and considers 156 species. However, it is possible to trace the origin of all *Citrus* to only three "basic" true species: *C. medica* (citron), *C. reticulata* 'Blanco' (mandarin) and *C. maxima* (pummelo) (Scora, 1975, Barret & Rhodes, 1976). This hypothesis gained support from various biochemical and molecular studies, using for examples RAPD, SCAR, cpDNA (Nicolosi et al., 2000) and SSR markers (Barkley et al., 2006, Garcia-Lor et al., 2012). A comparative analysis of the chloroplast genomes of 34 *Citrus* genotypes suggested the existence of three main clades (the citron/Australian, the pummelo/Micrantha, and the papeda/mandarins) from which the *Citrus* ancestor were probably generated in a succession of speciation events occurring between 7.5 and 6.3 Ma (Carbonell-Caballero et al., 2015). However, a comparison of various *Citrus* nuclear genomes, confirm the identification of pummelos as a single *Citrus* species but denied the role of true species to the traditional cultivated mandarins (as 'Blanco') suggesting the unknown small-fruited wild mandarins as the real *C. reticulata* ancestor (Wu et al., 2014). 'Chinotto' is closely related to sour orange and grouped within *C. reticulata* (mandarin) cluster, together with *C. sinensis* (Herrero et al., 1996). Moreover, a recent study suggests that 'Chinotto' is more closely related to *C. aurantium* compared to other relatives, forming a separate cluster from others sour orange cultivars (Polat et al., 2012), and supporting the hypothesis that *C. myrtifolia* derives from a sour orange mutation (Hodgson 1967). Furthermore, cpDNA analysis inferred the putative hybrid origin of the bitter orange from pummelo and mandarin, and thus the common origin of chinotto and bergamot from *C. aurantium* (Bayer et al., 2009).

Because of the lack of exhaustive studies about *C. myrtifolia*, it is not surprising the scarce knowledge about the existence of different cultivars. Hodgson reported, "*at least four forms or varieties of myrtle-leaf orange are recognized and there are doubtless several clones of each*": the 'Boxwood Leaf Chinotto' (Chinois à Fouilles de Buis), 'Crispifolia Chinotto' (Crinkle-Leaf Chinotto), 'Large Chinotto' and 'Dwarf Chinotto' (Hodgson 1967). Of particular interest is the case of the 'Chinotto di Savona', a tree cultivated in Liguria region (northern Italy), differentiated to the common chinotto for easily recognizable phenotypic differences. The bitter taste, intense aroma and seedless fruits make it a valuable chinotto variety and the preferred choice for the production of typical Italian candy, jam and beverages. Despite this, the genetic identity and the peculiarity of the 'Chinotto di Savona', including the main horticultural and morphological characteristics have not been investigated yet.

Aim of this work is the preliminary characterization of 'Chinotto di Savona' in order to provide evidences about its genetic identity, paving the way for further analysis.

4.2 Materials and Methods

4.2.1 Morphological and pomological analysis

4.2.1.1 Plant material

Twelve trees of 'Chinotto di Savona' and an equal number of *C. myrtifolia* were selected respectively in the farm "Ottone Luca" (Pietra Ligure, Liguria, Italy) and the farm "L'aquila" (Finale Ligure, Liguria, Italy) during February 2016. Shoots, leaves and 10 mature fruits for each plant were sampled. In order to perform the reflectance analysis, fruits with different flavedo colours were also collected. All the fresh samples were analysed immediately after sampling and stored at 4°C.

4.2.1.2 Determination of morphological measurement, total soluble solid, pH and titratable acidity

Leave length and fruit dimension were measured with ImageJ software (Abramoff et al., 2004) using a dimensional standard system. Fruit weight was obtained using a PE200 precision balance (Mettler Toledo, Columbus, USA).

For qualitative analysis, the juice was extracted by a manual juicer press and filtered. The determination of total soluble solids content (SSC) was performed for each fruit by an Atago PAL-COFFEE digital refractometer (Atago Co., LTD., Tokyo, Japan) and expressed in °Brix. For the determination of titratable acidity, after reading of the pH of the juice by a digital pH-meter Crison Compact Titrator D (Crison Instrument SA, Barcelona, Spain), the neutralization was performed by a solution of 0.1 N sodium hydroxide (NaOH) at pH 8.3 and with a starting volume of 7.5 ml of fresh juice using a Crison Compact Titrator D (Crison Instrument SA, Barcelona, Spain). The titratable acidity was then expressed as grams per litre of citric acid equivalent.

4.2.1.3 Total Phenolic Content

Total phenolic content (TPC) was determined spectrophotometrically according to the Folin–Ciocalteu's method (Singleton et al., 1999), modified as follows. Juices were centrifuged at 5000 rpm for 5 minutes and 500 µl of Folin–Ciocalteu's reagent was added to 500 µl of juice previously diluted 1:10 with distilled water. Two ml of 10% Na_2CO_3 solution were added to 7 ml of distilled

water. Samples were stored in the dark and after 90 minutes, the absorbance of blue coloration was measured at 700 nm against a blank sample. The measurements were expressed in absorbance unit.

4.2.1.4 Fruit reflectance analysis

Overall 537 reflectance spectra were obtained by a Jaz System (Ocean Optics, B.V., Dunedin, USA) spectrometer, completed with a channel with a DPU module and ILX511b detector, OFLV-3 filter, L2 lens and 50 μm slit as installed options. A reflection probe QR600-7-VIS125 was coupled to the spectrophotometer. The instrument was set up with a NIR/vis light source 4095 power setting, and the integration time was automatically corrected by the instrument. Collected spectra ranged between 340 nm and 1025 nm with a stepwise of about 0.3 nm. The spectra were calculated as percentage of reflectance (%R) in comparison with a reference blank spectrum obtained by a PTFE diffuse reflectance standard (Ocean Optics, B.V. Dunedin, USA). Spectral modifications during color evolution were shown after normalization at 800 nm (N800) (Rustioni et al., 2015). The 450-800 nm region was only taken into account.

4.2.1.5 Data analysis

Statistical analyses were performed by PAST3 software (Hammer et al., 2001; http://folk.uio.no/ohammer/past). Reflectance spectra were analysed by SPSS software (IBM Corp. 2013. IBM SPSS Statistics for Windows, version 22.0. Armonk, NY: IBM Corp; http://www-01.ibm.com/software/it/analytics/spss/).

4.2.2 Genetic analysis

4.2.2.1 Plant material

Fourteen species and cultivars belonging to the ancestral and secondary species of *Citrus* were selected. Nine of these (*C. aurantium*, *C. bergamia* 'Femminello', *C. medica* 'Ethrog', *C. latifolia* 'Bears', *C. limettioides*, *C. reticulata* 'Ponkan', *C. maxima* 'Pigmented', *C. sinensis* 'Tarocco Giallo' and *C. limon* 'Femminello') were sampled from the germplasm collection of Dipartimento di Scienze delle Produzioni Agrarie e Alimentari (University of Catania, Sicily, Italy). *C. sinensis* 'Navel', *C. medica*, *C. reticulata* and 'Chinotto di Savona' were sampled from the farm "Ottone Luca" (Pietra Ligure, SV, Italy) and *C. myrtifolia* from the farm "L'aquila" (Finale Ligure, SV, Italy).

4.2.2.2 DNA extraction

High molecular weight genomic DNA was extracted from leaf samples by the DNeasy Plant Mini Kit (Qiagen S.A, Madrid, Spain) according to the manufacturer's instructions. Nucleic acids were resolved on a 2% (w/v) agarose TAE gels and visualized by ethidium bromide staining under UV light (365 nm). Size and quantity of nucleic acid bands were estimated by comparison to HindIII-cut Lambda (Thermo Scientific, Waltham, USA) and 100 bp (Thermo Scientific, Waltham, USA) molecular markers.

4.2.2.3 SSR analysis

Twenty-one primers were chosen from the literature according to their power to provide polymorphic molecular data among a wide range of *Citrus* species (Kijas et al., 1997; Cuenca et al., 2011; Garcia-Lor et al., 2012; Ahmad et al., 2003; Froelicher et al., 2008). Seventeen primers pairs were labelled on the 5' position of the forward primer with an HEX, FAM or NED fluorophore (Table 1). The PCR mixture consisted of 1 ng/µl template DNA, 0.16 mM dNTPs, 0.3 mM forward primer, 0.3 mM reverse primer, 1X PCR buffer, 1.5 mM $MgCl_2$ and 0.075 U/µl *Taq* DNA polymerase (Eurx LTD, Danzica, Poland) in a final volume of 20 µl. Four primers pairs were marked according to the Tail PCR technique (Schuelke, 2000), with PET or VIC fluorophore, as shown in Table 1. The PCR mixture consists of 1 ng/µl template DNA, 0.2 mM dNTPs, 0.075 mM forward primer, 0.2 mM reverse primer, 0,2 mM marked primer, 1X PCR buffer, 1.5 mM $MgCl_2$ and 0.05 U/µl *Taq* DNA polymerase (Eurx LTD, Danzica, Poland) in a final volume of 20 µl. The reactions were carried out with the following program according to the annealing temperature of each primer pair (Table 1): 2 min at 94° C, 35 cycles of 30 s at 94° C, 30 s at 45–60° C and 1 min at 72° C with a final extension of 10 min at 72° C. The size of the PCR products was scored through an automated capillary analyser AB3730 (Thermo Scientific, Waltham, USA) at the Genomic Platform of Parco Tecnologico Padano, Lodi, Italia.

Table 1. Oligonucleotides used in the genetic analyses

Name	F	Ta	Forward primer (5'-3')	Reverse primer (5'-3')	Reference
CAC19	FAM	52	ACAACCTTCAACAAAACCTAGG	AAGACTTGGTGCGACAGG	Kijas 1997
TAA15	FAM	55	GAAAGGGTTACTTGACCAGGC	CTTCCCAGCTGCACAAGC	Kijas 1997
TAA41	FAM	55	AGGTCTACATTGGCATTGTC	ACATGCAGTGCTATAATGAATG	Kijas 1997
CAC15	HEX	55	TAAATCTCCACTCTGCAAAAGC	GATAGGAAGCGTCGTAGACCC	Kijas 1997
cAGG9	HEX	55	AATGCTGAAGATAATCCGCG	TGCCTTGCTCTCCACTCC	Kijas 1997
CAC33	HEX	55	GGTGATGCTGCTACTGATGC	CAATTGTGAATTTGTGATTCCG	Kijas 1997
CAC39	FAM	55	AGAAGCCATCTCTCTGCTGC	AATTCAGTCCCATTCCATTCC	Kijas 1997

TAA27	HEX	52	GGATGAAAAATGCTCAAAATG	TAGTACCCACAGGGAAGAGAGC	Kijas 1997
CAC23	FAM	45	ATCACAATTACTAGCAGCGCC	TTGCCATTGTAGCATGTTGG	Kijas 1997
TAA3	FAM	55	AGAGAAGAAACATTTGCGGAGC	GAGATGGGACTTGGTTCATCACG	Kijas 1997
TAA1	HEX	55	GACAACATCAACAACAGCAAGAGC	AAGAAGAAGAGCCCCCATTAGC	Kijas 1997
TAA33	HEX	45	GGTACTGATAGTACTGCGGCG	GCTAATCGCTACGTCTTCGC	Kijas 1997
TAA45	HEX	55	GCACCTTTTATACCTGACTCGG	TTCAGCATTTGAGTTGGTTACG	Kijas 1997
TAA52	FAM	45	GATCTTGACTGAACTTAAAG	ATGTATTGTGTTGATAACG	Kijas 1997
mCrCIR03B07	VIC	55	CACCTTTCCCTTCCA	TGAGGGACTAAACAGCA	Cuenca 2011
mCrCIR02G02	PET	55	TGGTAGAGAAACAGAGGTG	CAATAAGAAAACGCAGG	Cuenca 2011
Ci08C05	PET	55	TCCACAGATTGCCCATTA	CCCTAAAAACCAAGTGACA	Froelicher 2008
CMS-26	VIC	55	TGATGTCTTGATCCACACTTCC	ACTCAAAGCTCCGCTACAGTG	Ahmad 2003
CMS-47	FAM	55	GGATCCTCCACCATCTCGTA	TTCTTCTTCCATGCCGACTT	Ahmad 2003
MEST56	FAM	55	GGTGCAAAAGAGAGCGAGAG	AGTCCGCCTTTGCTTTTTCT	Garcia-Lor 2012
MEST488	NED	55	CACGCTCTTGACTTTCTCCC	CTTTGCGTGTTTGTGCTGTT	Garcia-Lor 2012
EcoRI_adapter1	/	/	CTCGTAGACTGCTACC	/	/
EcoRI_adapter2	/	/	AATTGGTACGCAGTC	/	/
MseI_adapter1	/	/	GACGATGAGTCCTGAG	/	/
MseI_adapter2	/	/	TACTCAGGACTCAT	/	/
E01_preamp_A	/	56	GACTGCGTACCAATTCA	/	/
M02_preamp_C	/	56	/	GATGAGTCCTGAGTAAC	/
E40_sel_AGC	/	56	GACTGCGTACCAATTCAGC	/	/
M47_sel_CAA	/	56	/	GATGAGTCCTGAGTAACAA	/
M48_sel_CAC	/	56	/	GATGAGTCCTGAGTAACAC	/
M49_sel_CAG	/	56	/	GATGAGTCCTGAGTAACAG	/
M50_sel_CAT	/	56	/	GATGAGTCCTGAGTAACAT	/
M51_sel_CCA	/	56	/	GATGAGTCCTGAGTAACCA	/
M52_sel_CCC	/	56	/	GATGAGTCCTGAGTAACCC	/
M53_sel_CCG	/	56	/	GATGAGTCCTGAGTAACCG	/
M54_sel_CCT	/	56	/	GATGAGTCCTGAGTAACCT	/

F: fluorophore; Ta: annealing temperature

4.2.2.4 AFLP analysis

DNA samples of *Citrus aurantium*, *Citrus myrtifolia* and 'Chinotto di Savona' were analysed by Amplified Fragment Length Polymorphism (AFLP). For each sample 0.5 µg of genomic DNA was mixed with 4 µl of TANGO buffer 10X (Thermo Scientific, Waltham, USA), 1 µl of EcoRI enzyme at 10 U/µl (Thermo Scientific, Waltham, USA) and H_2O in a total volume of 22 µl. The restriction solutions were incubated for 30 minutes at 37° C followed by 30 minutes at 67° C after the addition of 1 µl of MseI enzyme at 10 U/µl (Thermo Scientific, Waltham, USA) in each solutions. The adaptors were prepared by mixing 10 µl of EcoRI_adapter1 10 µM with 10 µl of EcoRI_adapter2 10 µM and 10 µl of MseI_adapter1 10 µM with 10 µl MseI_adapter2 10 µM (Table 1). The two adapter solutions were heated for 5 minutes at 94° C and then allowed to cool at room temperature.

For each sample 4 µl of the previously restricted DNA was mixed with 2 µl of T4 DNA Ligase Buffer 10X (Thermo Scientific, Waltham, USA), 1 µl T4 DNA Ligase 5 U/µl, 1 µl of EcoRI adapter 10 µM, 1 µl MseI adapter 10 µM and H_2O for a total volume of 20 µl. The ligation solutions were incubated at 22° C for 1 hour.

The pre-selective PCR reaction was performed for each sample using the primers E01 and M02 (Table 1). The pre-amplification PCR mixture consists of 5 µl of previously ligated DNA, 0.16 mM dNTPs, 0.3 mM primer E01, 0.3 mM primer M02, 1X PCR buffer (Eurx LTD, Danzica, Poland), 1.5 mM $MgCl_2$ and 0.075 U/µl *Taq* DNA polymerase (Eurx LTD, Danzica, Poland) and H_2O in a final volume of 20 µl. The reactions were carried out with the following program: 2 min at 94° C, 20 cycles of 30 s at 94° C, 30 s at 56° C and 1 min at 72° C. The selective PCR reaction was performed for each sample using all the eight possible combinations of the primer E40 with the others (Table 1). The selective PCR mixture consists of 1 µl of previously pre-amplificated DNA, 0.16 mM dNTPs, 0.3 mM primer E40, 0.3 mM primer 'M' (from M47 to M54), 1X PCR buffer (Eurx LTD, Danzica, Poland), 1.5 mM $MgCl_2$ and 0.075 U/µl *Taq* DNA polymerase (Eurx LTD, Danzica, Poland) and H_2O to a final volume of 20 µl. The reactions were carried out with the following touch-down-PCR profile: 2 min at 94° C, 10 cycles of 20 s at 94° C, 30 s at 66° C, 2 min at 72° C with each cycle scaling down of 1° C the annealing temperature and then 35 cycles of 20 s at 94° C, 30 s at 56° C and 2 min at 72° C, with a final extension of 10 min at 72° C. The resulting amplified products were visualized on 6% acrylamide gels using silver nitrate staining and 10 bp DNA Ladder (Thermo Scientific, Waltham, USA) as size marker.

4.2.2.5 Genetic data analysis

The output files were visualized and scored using the software Geneious 9.0 (Kearse et al., 2012) available at http://www.geneious.com. The resulting genotyping data were transformed into a binary matrix file, 1 for presence of the considered allele and 0 for absence, and a classic co-dominant matrix file. Cluster analysis was performed by using the PAST3 software (Hammer et al., 2001; http://folk.uio.no/ohammer/past), calculating the Sørensen–Dice similarity index (Sørensen 1948, Dice 1945) and both the neighbour-joining algorithm (Saitou and Nei 1987) and the unweighted pair-group method (UPGMA). Statistics on the SSR data was performed using GenAlEx (Peakall and Smouse, 2006; 2012) available at http://biology-assets.anu.edu.au/GenAlEx/Download.html) and PICcalc (Nagy et al., 2012) available at http://w3.georgikon.hu/pic/english/default.aspx).

4.2.2.6 Sanger sequencing of polymorphic DNA fragments

Acrylamide bands were cut from the gel using a scalpel and then incubated in H_2O at 65 °C for 24 hours. The surnatants were collected and quantified using a Qubit Fluorometric Quantitation system (Thermo Scientific, Waltham , USA) and then the DNA was amplified using the experimental conditions previously described for the AFLP analysis. The resulting amplicons were purified using a microCLEAN PCR purification kit, resolved on a 2% (w/v) agarose TAE gels and visualized by ethidium bromide staining under UV light (365 nm) and quantified using a Qubit Fluorometric Quantitation system (Thermo Scientific, Waltham , USA). The sequencing reactions were performed through an external service provider (BMR Genomics, Padova, Italy) and the consensus sequences were elaborated using the software Geneious 9.0 (Kearse et al., 2012). The MEGA-Blast algorithm was used to find matches against the *Citrus sinensis* and *Citrus clementine* reference genomes and the whole NCBI Genomic Reference Sequences database (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

4.3 Results and Discussion

4.3.1 Phenotypic analysis

Literature information about main phenotypic characteristics of *C. myrtifolia* are scarce and very little is known about 'Chinotto di Savona'. In Table 2 are reported the values of some morphological traits detected in this work with the aim to provide a phenotypic characterization about the two investigated accessions.

Table 2. Average values of different botanical traits.

Organ	Chinotto di Savona	*C. mirtyfolia*	t-test
Leaf length (cm)	6.11 ±1.42	2.54 ± 0.51	*
Leaf width (cm)	2.74 ±0.60	1.07 ± 0.37	*
Internode (mm)	8.04 ±0.66	4.95 ± 0.18	*
Fruit weight (g)	26.63 ±8.83	34.88 ± 7.05	
Fruit dimension (ed/ld)	1.22 ± 0.08	1.14 ± 0.05	
Fruit BRIX (°)	9.11 ± 0.99	9.54 ± 0.76	
Fruit pH	2.96 ± 0.25	4.84 ± 0.07	*
Fruit acidity (g/l)	31.68 ± 2.30	2.68 ± 0.25	*
Fruit TPC (A)	0.43 ± 0.06	0.50 ± 0.03	

*: Positive t-test (p-value lower than 0.01); A: absorbance units. ed/ld: equatorial diameter/longitudinal diameter.

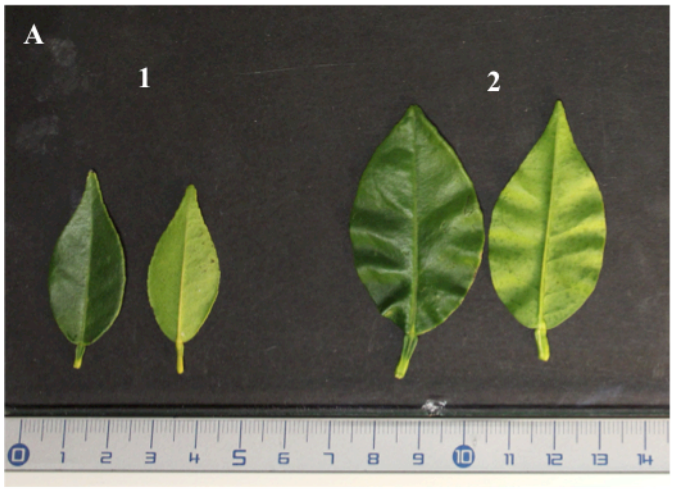
4.3.1.1 Shoots and leaves

Significant differences between *C. myrtifolia* and 'Chinotto di Savona' were detected for leaf length and width, as well as for the internode length (Table 2). The 'Chinotto di Savona' leaves are 2-3 times larger than those of *C. myrtifolia* (Figure 1A), whereas internodes are longer, resulting in a less dense vegetation (Figure 1B). As a consequence, the two accessions are easily recognizable in the field.

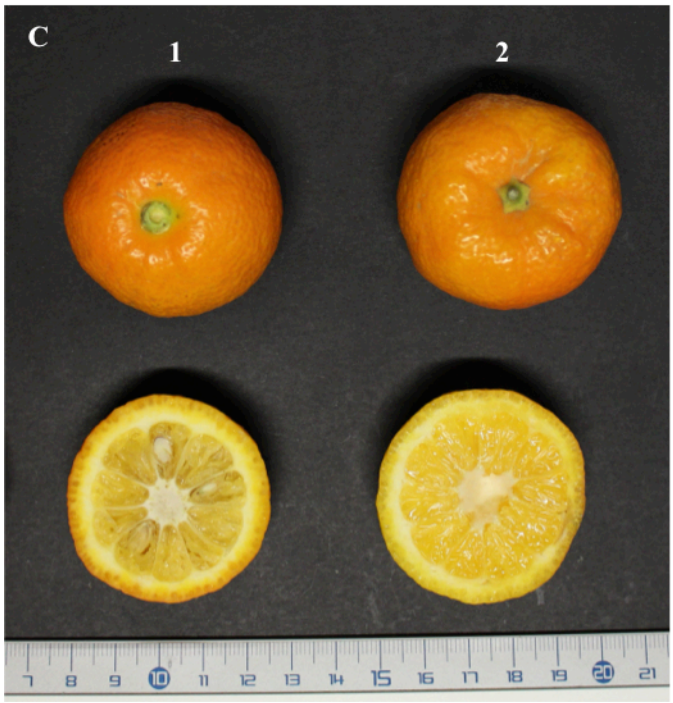
4.3.1.2 Fruit morphology and juice

Fresh weight and size of fruits showed no significant differences between *C. myrtifolia* and 'Chinotto di Savona' even if both samples show high within-plant variability (Table 2)**.** Longitudinal sections revealed a similar number of segments between the two accessions (8.14 for 'Chinotto di Savona' and 8.54 for *C. myrtifolia*) but they largely diverge for the number and the morphology of the seeds. While *C. myrtifolia* produces viable seeds, (on average 3 per fruit), 'Chinotto di Savona' fruits are seedless, except a few cases with aborted and sketchy seeds (Figure

1C-1D). The presence of aborted seeds could be due either to a triploidy or to a genetic defect in the regulatory pathway of seed development. Seedless fruit could be a consequence of parthenocarpy, a trait rather common in several *Citrus* species. Seedless fruits or, at most, the occurrence of fruits with a few aborted seeds is one of the reasons that make the fruits of 'Chinotto di Savona' the preferred choice for the candy and jam industry. The juice colour ranges from yellow to orange. No differences were found for SSC, averaging at around 9 °Brix, and for the total phenolic content (Table 2). On the contrary, the two accessions widely differed for pH and acidity. 'Chinotto di Savona' showed a lower pH (2.9 vs. 4.8) and a ten-fold higher acidity (31.6 vs. 2.7) compared to *C, myrtifolia*. Carboxylic acids play a crucial role in the organoleptic quality of fruit beverage representing an important aspect for the commercial development of 'Chinotto di Savona'.


A
1
2


B
1
2


C
1
2


E
0
1
2
3


D
1
2

Figure 1. Phenological traits between *Citrus myrtifolia* and 'Chinotto di Savona'

(**A**) Representative leaves of *Citrus myrtifolia* (1) and of 'Chinotto di Savona' (2). (**B**) Representative shoots of *Citrus myrtifolia* (1) and of 'Chinotto di Savona' (2). (**C**) Representative fruits of *Citrus myrtifolia* (1) and of 'Chinotto di Savona' (2). (**D**) Representative seeds of *Citrus myrtifolia* (1) and of 'Chinotto di Savona' (2). (**E**) Colour evolution of 'Chinotto di Savona' fruit: green stage (0), yellow stage (1), orange stage (2) and deep-orange stage (3).

4.3.1.3 Reflectance spectroscopy

Spectroscopy analyses were performed on fruits skin in order to find differences about the pigment composition in the two investigated accessions. Both spectra showed a similar trend (Figure 2A). However, a wide hyperchromic (higher absorption intensity) and a bathochromic (shift of the absorption band to low energetic wavelengths) effects were detected in the 'Chinotto di Savona'. The hyperchromic effect suggests a higher content of pigments in the flavedo (Rustioni et al., 2014 A; Rocchi et al., 2016). However, the additional bathochromic effect could be related to interactions involving pigments and other molecules (Rustioni et al., 2014 b). Moreover, the main absorbance band of both accessions ranged from 450 nm to 550 nm, suggesting that carotenoids and chlorophylls are the most represented pigments in flavedo (Zur et al., 2000). Based on the spectra, the red β-citraurin and the orange cryptoxanthin could be the main responsible for the orange colour of mature 'Chinotto', as also reported for Tangerine (*C. tangerine*) (Gross J. 1981). The confidence intervals of the two spectra do not overlap, suggesting a significant difference between the two accessions.

In order to evaluate pigmentation of *Citrus* fruit, semi-objective scores have been developed by using a colour chart (Iwahori et al., 1986). Searching for more objective determination of skin colour of 'Chinotto di Savona', colorimetric parameters using reflectance spectroscopy have been estimated in fruits (Figure 1E) representative of different colour evolution stages (from 0, green-coloured fruits to 4, deep-orange fruits). Fruits were analysed also regarding their acidity and total phenolic content for each different skin colour. During colour change the fruit acidity remains unchanged (Table 2), whereas total phenolic content gradually decreases from the green to the deep-orange stage (from 0.83 A to 0.43 A). In green fruits, the main pigment composition of flavedo is represented by chlorophyll a, (680 nm), and chlorophyll a, b and carotenoids, all together at 450-550 nm (Figure 2B) (Zur et al., 2000; Rocchi et al., 2016; Rustioni et al., 2014a). Subsequently, From "0" to "1" stage a large decrease of chlorophyll a is detected together with a small decrease of the peak at 450-550 nm suggesting that in this wavelength range the decrease in chlorophylls is counterbalanced by carotenoids accumulation (Zur et al., 2000). From "1" to "2"

stage, chlorophyll a undergoes a further decrease but instead the other peak shows a hyperchromic and bathochromic effect, supporting the ongoing shift from chlorophylls to carotenoids. In the final spectrum, the corresponding peak of chlorophyll a completely disappears, and the peak in the spectral range of 450-550 nm undergoes a further hyperchromic/bathochromic shift, positioning itself at the same wavelength range previously reported for 'Chinotto di Savona' (Figure 2A), thus showing the complete shift from chlorophylls to carotenoids and explaining the change in flavedo color (Figure 2B, Figure 1E). As also observed in several fruit of *Citrus* (Lado et al., 2014) the changes in 'Chinotto di Savona' fruit color during ripening evolution is mainly due to different relative proportions between pigments.

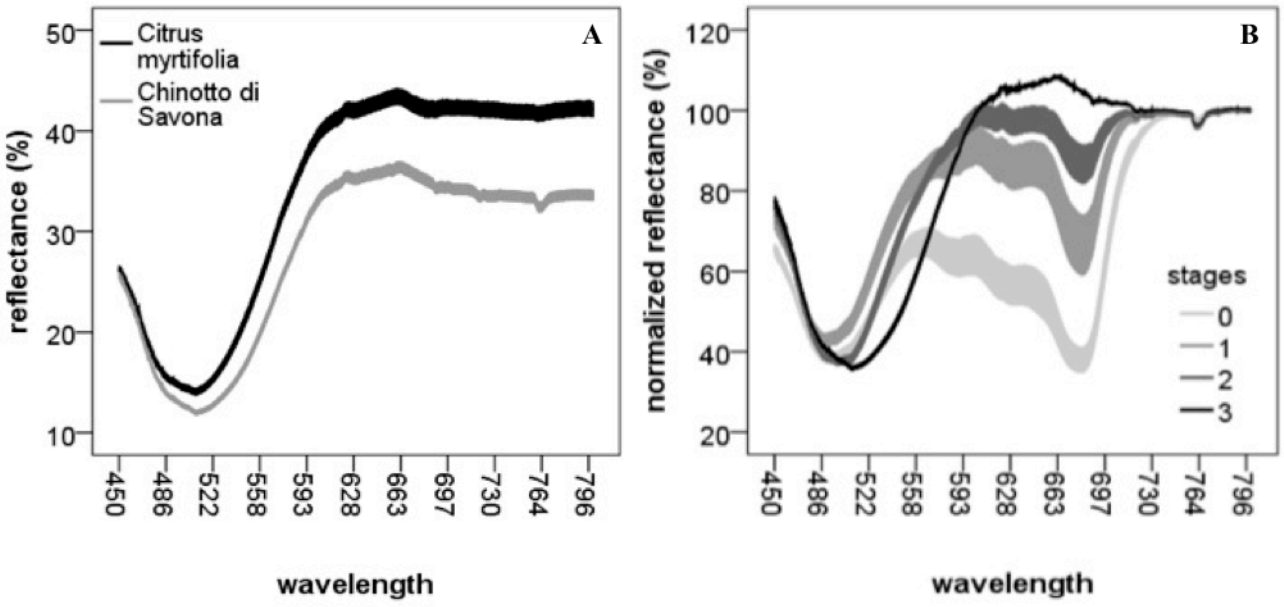


Figure 2

(**A**) Confidence intervals (per P=0.95) of the average spectra (%R) for mature fruits of *Citrus myrtifolia* and 'Chinotto di Savona'. (**B**) Confidence intervals (per P=0.95) of the average normalized spectra (N800) of fruit skin. Green stage (0), yellow stage (1), orange stage (2) and deep-orange stage (3).

4.3.2 Genetic analysis

4.3.2.1 SSR analysis

Since many years, 'Chinotto di Savona' accession, has been considered a variety of *C. myrtifolia* in the traditional growing areas in Liguria region (Northern Italy). Nevertheless, genetic evidences about the true identity of this variety are still lacking.

For this purpose twenty-one SSR markers (Kijas et al., 1997; Garcia-Lor et al., 2012; Ahmad et al., 2003, Froelicher et al., 2008), widely used for genetic studies in *Citrus*, were selected to clarify the genetic origin of 'Chinotto di Savona' and the relationship with different *Citrus* species, including *C. myrtifolia*. SSR markers represent a useful tool for the rapid and relatively inexpensive assignment of new, unknown accessions or cultivar to their genetic group in *Citrus* genus (Garcia et al., 2012).

Eighteen of the 21 SSR marker loci resulted to be polymorphic, producing well-resolved fragments (Supplementary Material Table 1). The marker TAA45 shows a multi-locus pattern as also previously reported for CMS-47 (Ahamad at al., 2003) . Marker CAC33 showed a complex PCR amplification pattern, as also observed by Kijas et al., (1997) and was excluded from the analysis, together with CMS-46 and CAC19. In contrast with the same work (Kijas et al., 1997), markers CAC15 and CAC39 evidenced polymorphism. Genetic diversity parameters were calculated for each locus marker (Table 3). A total of 82 alleles were detected, ranging from 2 (for molecular markers cAGG9, CAC15) to 9 (Ci08C05, TAA41) and showing mean alleles numbers of 5.125, although the effective alleles are lower. Markers TAA15 and Ci08C05 show the larger discrepancy between the two values and, therefore, the greater amount of low-frequency alleles. The observed heterozygosity (OH) was calculated as a measure of marker diversity (Table 3). The percentage of detected heterozygotes per marker ranges from 23.1% (cAGG9) to 92.3% (mCrCIR02G02) while the mean observed heterozygosity across all markers is 61.2%. Moreover MEST488 shows a high value of expected heterozygosity as also reported by Garcia et al. (2012) (data not shown). The polymorphism information content (PIC) value for the marker set range from 0.262 (cAAG9) to 0.842 (TAA41) with a mean value of 0.593. Most of markers (10 out of 16) show values of OH and PIC higher than 0.5, providing a good discrimination potential for genetic analysis. In particular, cAAG9 is the least informative marker and TAA41 is the most informative one as also reported by other works (Barkley et al., 2006; Garcia et al., 2012).

Table 3. SSR markers discrimination power

SSR loci	AS	OA	EA	OH	PIC
cAGG9	99-121	2	1.451	0.231	0.262
TAA1	157-180	4	3.798	0.846	0.689
TAA52	76-113	4	1.858	0.250	0.425
TAA33	107-125	4	1.823	0.500	0.412
CAC39	147-180	2	1.600	0.357	0.305
mCrCIR03B07	264-278	6	3.798	0.769	0.703
TAA15	141-204	7	2.497	0.643	0.576
CAC15	144-180	2	1.704	0.583	0.328
mCrCIR02G02	110-138	6	4.899	0.923	0.764
TAA3	112-255	6	4.612	0.571	0.749
Ci08C05	148-182	9	6.627	0.615	0.832
CAC23	105-270	5	3.449	0.769	0.664
MEST488	133-164	6	4.829	0.846	0.762
TAA41	123-185	9	7.042	0.769	0.842
MEST56	129-145	7	4.899	0.692	0.767
TAA27	165-242	3	1.806	0.429	0.402
MEAN	/	5.125	3.543	0.612	0.593

AS: Allele size in base pairs; OA number of observed alleles; EA number of effective alleles; OH: observed heterozygosity; PIC: polymorphism information content value.

The cluster analysis (Figure 3) supports the common differentiation of the three main *Citrus* basic taxa as shown by many studies (Scora 1975, Barrhet and Rodes 1976, Nicolosi et al., 2000, Barkley et al., 2006, Garcia-Lor et al., 2012): *C. reticulata* and *C. maxima* tends to share a clade, instead *C. medica* clearly separates from the others as also reported (Barkley et al., 2006, Nicolosi et al., 2000, Garcia-Lor et al., 2013). *C. medica* clusters together with *C. limettoides*, *C. latifolia* 'Bears' and *C. limon* 'Femminello', supporting the common origin of these secondary species from citron (Nicolosi et al., 2000, Barkely et al., 2006, Garcia-Lor et al., 2012). The same evidences are true for *C. sinensis* and *C. aurantium,* clustering together with mandarins and pummelo, accordingly to their origin (Nicolosi et al., 2000, Barkely et al., 2006). *C. bergamia* was supposed to be an hybrid between *C. aurantium* and *C. medica* and it was reported to cluster in the citron group (Nicolosi et al., 2000; Barkley et al., 2006). However, it clusters with *C. aurantium* in the pummelo/mandarins group. As reported in the literature, SSR markers are not always able to distinguish between cultivars arisen by spontaneous mutation such as among sweet oranges (Barkley et al., 2006; Fang and Roose 1997; Breto et al., 2001). Surprisingly the selected markers clearly differentiated the two sweet oranges cultivars 'Tarocco' and 'Navel' and the two mandarins, but failed to discriminate *C. myrtifolia*, *C. aurantium* and 'Chinotto di Savona' and this could be possibly due to the supposed origin of 'Chinotto' from a somatic mutation of *C. aurantium* (Hodgson, 1967). Even if this cluster analysis did not provide a genetic discrimination of the two

accessions of interest, it groups the 'Chinotto di Savona' into the pummelo/mandarins cluster suggesting that this accession could be originated from a somatic mutation of either *C. myrtifolia* or *C. aurantium*. All these evidences are supported by bootstrap values higher than 70% (Figure 3). The same results are obtained for the UPGMA tree (data not shown).

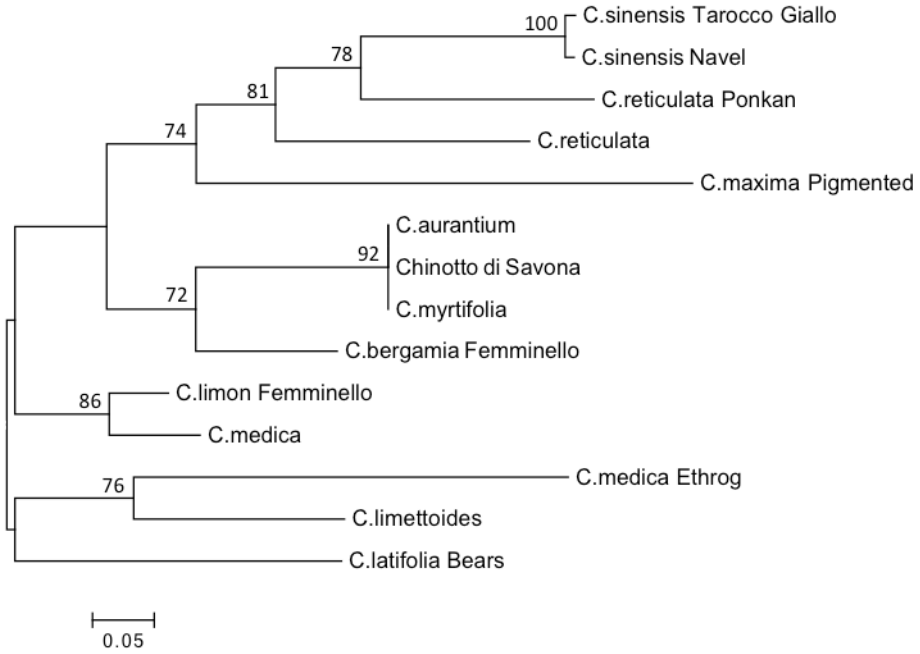


Figure 3. Cluster analysis of the main *Citrus* species by 18 SSR markers.
Clustering by the neighbour-joining algorithm on similarity index calculated by the PAST3 software. Bootstrap values higher than 70% are shown at each node as a percentage of 1000 repetitions.

4.3.2.2 AFLP analysis

Since SSR approach was not able to distinguish between *C. myrtifolia*, *C. aurantium* and 'Chinotto di Savona', the accessions were further tested by using an AFLP approach. All primer combinations generated a recognizable fragments pattern for each sample and the most part of them were shared between all the samples. However, five clear bands uniquely discriminate *C. myrtifolia* from 'Chinotto di Savona' (primer pair 1, 350 bp; primer pair 3, 270 bp; primer pair 3, 130 bp; primer pair 5, over 400 bp; primer pair 7, 200 bp) showing a genetic difference between the two accessions (Figure 4). However, the set of primer combination used for AFLP analysis were still unable to genetically distinguish 'Chinotto di Savona' from *C. aurantium*. The output obtained by using AFLP suggests that 'Chinotto di Savona' is more related to sour orange than the common chinotto and further supports the hypothesis of *C. myrtifolia* as originated from a somatic mutation

of *C. aurantium*. To this regards, 'Chinotto di Savona' could represent the first step of *C. myrtifolia* differentiation from *C. aurantium*, configuring itself as the real ancestor of the 'common' chinotto. On the contrary, the two accessions could have been arisen from independent somatic mutations of *C. aurantium*.

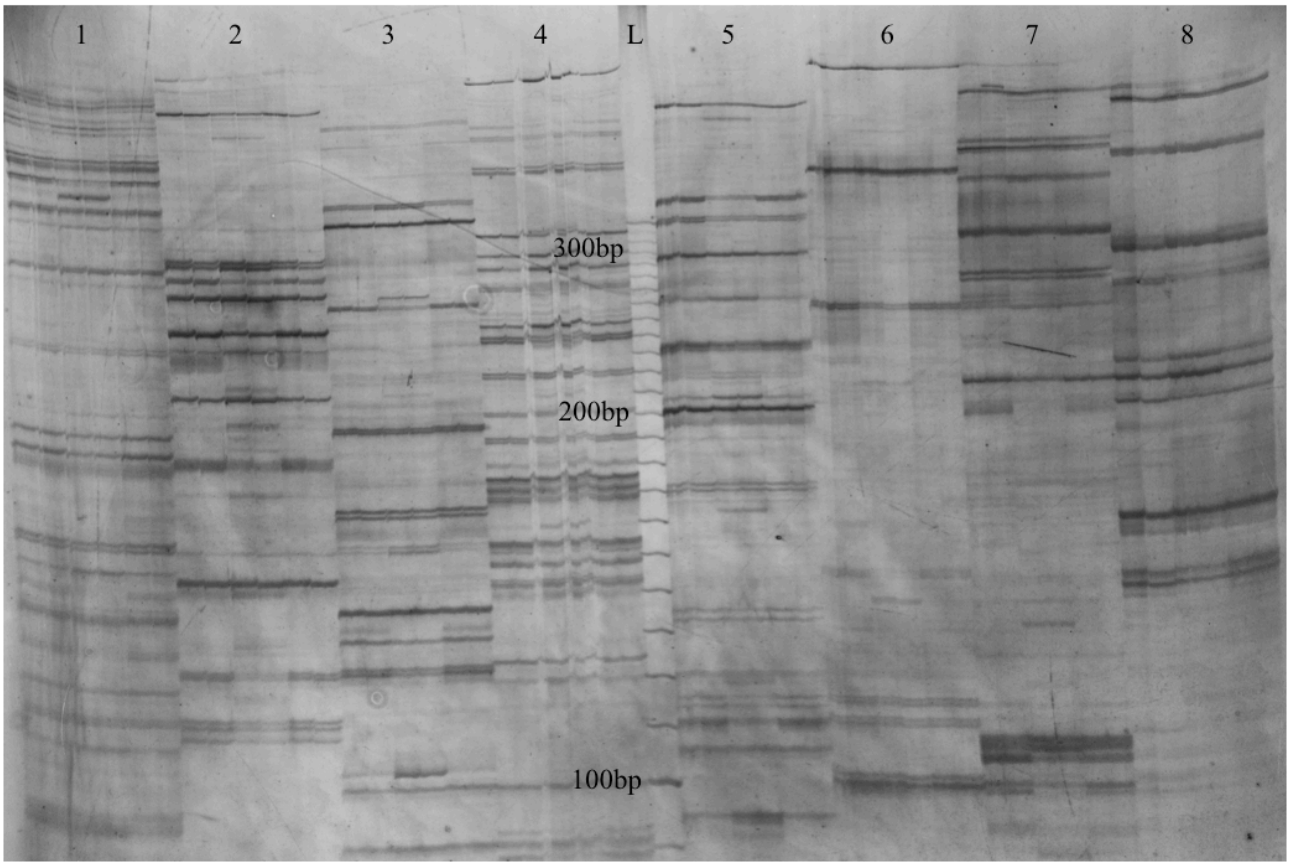


Figure 4. AFLP analysis of Citrus species.

AFLP on 6% acrylamide gel with silver staining. Primer combinations: 1 (E40 + M47), 2 (E40 + M48), 3 (E40 + M49), 4 (E40 + M50), 5 (E40 + M51), 6 (E40 + M52), 7 (E40 + M53), 8 (E40 + M54), L: 10bp DNA Ladder. Within each primer combination the samples are arranged with the following order: *C. aurantium* (two biological replicates), *C. myrtifolia* (two biological replicates) and 'Chinotto di Savona' (two biological replicates).

4.3.2.3 Sequencing of gel-excised bands

The five bands able to uniquely discriminate the *C. myrtifolia* accessions were excised from the acrylamide gel, suspended in water and re-amplified using the appropriate primers combinations. The resulting amplicons were sequenced and blasted on genomics databases to search for relevant matches. Three bands were excluded from the analyses because constituted by

aspecific amplification products. The ‘E40 + M51’ band (400 bp) matches with an ENHANCED DISEASE RESISTANCE 2-like predicted protein on *Citrus sinensis* reference genome and on the homologue hypothetical proteins on *Citrus clementina* reference genome. The observed polymorphism in the ‘E40 + M49’ (130 bp) band is of particular interest because it is located on the CDS region of the gene for a phosphoenolpyruvate carboxykinase (and related homologue on *C. clementine*), a protein usually involved in the metabolism of tricarboxylic acids (TCA cycle). This evidence suggests that it could have a role in the protein function during the ripening of fruits and thus explaining the observed differences on acidity and pH of the two ‘Chinotto’ accessions.

4.4 Conclusions

'Chinotto di Savona', a traditional *Citrus* cultivated in Liguria region (Italy) is gaining a renewed interest for the production of high-quality niche food and beverages. Although it is widely considered a valuable variety of chinotto (*C. myrtifolia*), no scientific studies support this assertion. In this work, molecular and phenotypic analysis of 'Chinotto di Savona' highlighted the presence of significant differences with respect to *C. myrtifolia*. 'Chinotto di Savona' shows some favorable pomological traits, including seedlessness and a high juice acidity that makes it preferable to the common chinotto for the food industry.

Molecular analysis confirmed the genetic similarity of 'Chinotto di Savona' to the *C. aurantium* group. Moreover, a discrete number of AFLP markers allowed discriminating 'Chinotto di Savona' from *C. myrtifolia* suggesting the presence of a polymorphism inside the coding region of a phosphoenolpyruvate carboxykinase that could be involved into the observed differences on acidity and pH of the two Chinotto accessions. Further sequencing of the protein coding region coming from the three accessions, could reveal functional mutations to support this hypothesis. This finding confirms the local convinction about 'Chinotto di Savona', opening the door to the objective characterization of this accession. Moreover, by using AFLP, 'Chinotto di Savona' appeared undistinguishable from *C. aurantium*, giving birth to new questions about the origin and the identity of this species. In this sense, novel powerful genomic tools, such as the Citrus SNP array (Fujii et al., 2013) or NGS approaches may provide more in depth information, helping to unravel this question.

4.5 Supplementary Materials

Supplementary Table 1. Allele combination of each sample for the 18 selected markers.

Sample	*C. sinensis* 'Navel'	*C. medica*	*C. reticulata*	'Chinotto di Savona'
cAGG9	114/114	102/114	114/114	114/114
TAA1	159/162	168/176	159/162	159/168
TAA52	77/115	115/115	115/115	115/115
TAA33	114/123	114/114	114/114	114/123
CAC39	167/167	167/173	167/167	167/167
mCrCIR03B07	280/295	280/284	278/280	280/295
TAA15	162/185	162/162	189/201	162/195
CAC15	156/156	148/156/160	148/156/160	156/160
mCrCIR02G02	132/142	132/158	136/142	136/158
TAA3	138/142	146/146	142/142	138/148
Ci08C05	192/192	173/173	186/194	177/182
CAC23	244/248	242/242	134/244/248	134/244
MEST488	141/155	135/149	141/145	141/149
TAA41	135/151	142/147	144/144	130/142
MEST56	152/165	154/154	154/156	154/160
TAA27	190/190	166/190	190/190	190/213
TAA45	130/135/138	122/128/130/135	130/135/138	128/130/135
CMS47	188/190/209/212	184/188/194/206	184/188/206	178/184/188/194/206
Sample	*C. aurantium*	*C. bergamia* 'femminello'	*C. medica* 'Ethrog'	*C. myrtifolia*
cAGG9	114/114	114/114	102/102	114/114
TAA1	159/168	159/168	176/176	159/168
TAA52	115/115	0/0	0/0	115/115
TAA33	114/123	114/114	114/114	114/123
CAC39	167/167	167/173	173/173	167/167
mCrCIR03B07	280/295	280/295	282/282	280/295
TAA15	162/195	162/162	162/162	162/195
CAC15	156/160	156/160	156/156	156/160
mCrCIR02G02	136/158	130/158	0/0	136/158
TAA3	138/148	140/140	146/148	138/148
Ci08C05	177/182	177/182	155/155	177/182
CAC23	134/244	244/244	242/242	134/244
MEST488	141/149	149/149	135/135	141/149
TAA41	130/142	142/147	139/139	130/142
MEST56	154/160	154/160	0/0	154/160
TAA27	190/213	190/190	166/166	190/213
TAA45	128/130/135	128/130/135/144	138/138	128/130/135
CMS47	178/184/188/194/206	184/188/194	184/188	178/184/188/194/206
Sample	*C. latifolia* 'Bears'	*C. limettioides*	*C. reticulata* 'Ponkan'	*C. maxima* 'Pigmented'
cAGG9	102/105/114	102/114	114/114	114/114
TAA1	164/168/176	159/176	159/162	162/162
TAA52	77/77	115/115	91/115	75/75
TAA33	114/116	114/114	114/114	116/120
CAC39	167/173	167/173	167/167	167/167
mCrCIR03B07	280/288/292	278/284	280/280	290/290
TAA15	162/165	162/201	167/185	162/162

CAC15	156/156	156/160	156/156	156/160
mCrCIR02G02	132/140	142/142	132/136	130/142
TAA3	132/146	146/146	142/142	138/138
Ci08C05	155/171/182	186/186	177/177	161/163
CAC23	242/256	242/244	244/248	244/248
MEST488	135/143/149	135/153	145/153	141/153
TAA41	130/142/166	130/139	151/151	125/133
MEST56	156/156	163/163	152/163	160/186
TAA27	190/190	166/190	190/190	190/190
TAA45	122/128/130/135/150	130/144	130/135/138	96/128/168/170
CMS47	184/188/194	182/184/188/209/212	188/206/212	188/190

Sample	*C. sinensis* 'Tarocco Giallo'	*C. limon* 'Femminello'
cAGG9	114/114	102/114
TAA1	159/162	168/176
TAA52	77/115	115/115
TAA33	114/123	114/114
CAC39	167/167	167/173
mCrCIR03B07	280/295	280/284
TAA15	162/185	162/162
CAC15	156/156	156/160
mCrCIR02G02	132/142	130/158
TAA3	138/142	138/146
Ci08C05	173/192	173/182
CAC23	244/248	242/244
MEST488	141/155	135/149
TAA41	135/151	142/147
MEST56	152/165	154/154
TAA27	190/190	166/190
TAA45	130/135/138	128/130/135/144
CMS47	188/190/209/212	184/188/194/206

The manuscript has been submitted to Plant Physiology and Biochemistry

4.6 References

Abramoff, M. D., Magalhaes, P. J., Ram, S.J., 2004. "Image Processing with ImageJ". Biophotonics International. 7, 36-42.

Ahamd, R., Struss, D., Southwick, S. M., 2003. Development and Characterization of Microsatellite Markers in *Citrus*. J. Amer. Soc. Hort. Sci. 128, 584-590.

Barkley, A. N., Roose, M. L., Krueger, R. R., Federici, C. T., 2006. Assessing genetic diversity and population structure in a *Citrus* germplasm collection utilizing simple sequence repeat markers (SSRs). Theor. Appl. Genet. 112, 1519–1531.

Barreca, D., Bellocco, E., Caristi, C., Leuzzi, U., Gattuso, G., 2010. Flavonoid composition and antioxidant activity of juices from chinotto (*Citrus x myrtifolia* Raf.) fruits at different ripening stages. J. Agric. Food Chem. 58, 3031–3036.

Barreca, D., Bellocco, E., Caristi, C., Leuzzi, U., Gattuso, G., 2011. Elucidation of the flavonoid and furocoumarin composition and radical-scavenging activity of green and ripe chinotto (*Citrus myrtifolia* Raf.) fruit tissues, leaves and seeds. Food Chemistry. 129, 1504–1512.

Barrett, H. C., Rhodes, A. M., 1976. A numerical taxonomic study of affinity relationships in cultivated *Citrus* and its close relatives. Syst. Bot. 1, 105– 136.

Bayer, R. J., Mabberley, D. J., Morton, C., Miller, C. H., Sharma, I. K., Pfeil, B. E., Rich, S., Hitchcock, R., Sykes, S., 2009. A molecular phylogeny of the orange subfamily (*Rutaceae: Aurantioideae*) using nine cpDNA sequences. American Journal of Botany. 96, 668–685.

Breto, M. P., Ruiz, C., Pina, J. A., Asins M. J., 2001. The diversification of *Citrus clementina* Hort. ex Tan., a vegetatively propagated crop species. Mol. Phylogenet. Evol. 21, 285-293.

Camarda, L., Di Stefano, V., Fatta Del Bosco, S., Schillaci, D., 2007. Antiproliferative activity of *Citrus* juices and HPLC evaluation of their flavonoid composition. Fitoterapia. 78, 426–429.

Carbonell-Caballero, J., Alonso, R., Ibanez, V., Terol, J., Talon, M., Dopazo, J., 2015. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. Mol. Biol. Evol. 32, 2015-1035.

Cautela, D., Pirrello, A. G., Esposito, C., Minasi, P., 2004. Caratteristiche compositive del chinotto (*Citrus myrtifolia*). Essenze e Derivati Agrumari. 74, 49–55.

Chialva, F. and Doglia, G., 1990. Essential oil costituents of chinotto (*Citrus aurantium* L. var. *myrtifolia* Guill.). J. Ess. Oil Res. 2, 33-35.

Cuenca, J., Froelicher, Y., Aleza, P., Juarez, J., Navarro, L., Ollitrault, P., 2011. Multilocus half-tetrad analysis and centromere mapping in citrus: evidence of SDR mechanism for 2n megagametophyte production and partial chiasma interference in mandarin cv 'Fortune'. Heredity. 5, 462-470

Dice, L. R., 1945. Measures of the Amount of Ecologic Association Between Species. Ecology. 26, 297-302.

Fang, D. Q., Roose M. L., 1997. Identification of closely related *Citrus* cultivars with inter-simple sequence repeat markers. Theor. Appl. Genet. 95, 408-417.

Froelicher, Y., Dambier, D., Bassene, J. B., Costantino, G., Lotfy, S., Didout, C., Beaumont, V., Brottier, P., Risterucci, M., Luro, F., Ollitrault, P., 2008. Characterization of microsatellite markers in mandarin orange (*Citrus reticulata* Blanco). Molecular Ecology Resources. 8, 119-122.

Fujii, H., Shimada, T., Nonaka, K., Kita, M., Kuniga, T., Endo, T., Ikoma, Y., Omura, M., 2013. High-throughput genotyping in citrus accessions using an SNP genotyping array. Tree Genetics & Genomes. 9, 145–153.

Garcia-Lor, A., Curk, F., Snoussi-Trifa, H., Morillon, R., Ancillo, G., Luro, F., Navarro, L., Ollitrault, P., 2013. A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the 'true *citrus* fruit trees' group (*Citrinae, Rutaceae*) and the origin of cultivated species. Annals of Botany. 111, 1-19.

Garcia-Lor A., Luro, F., Navarro, L., Ollitrault, P., 2012. Comparative use of InDel and SSR markers in deciphering the interspecific structure of cultivated citrus genetic diversity: a perspective for genetic association studies. Mol Genet Genomics. 287, 77–94.

Gross, J., 1981. Pigment changes in the flavedo of Dancy tangerine (*Citrus reticulata*) during ripening. Zeitschrift für Pflanzenphysiologie.

Hammer, Ø., Harper, D. A. T., Ryan, P. D., 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica. 4, 9.

Hanelt, P., Büttner, R., Mansfeld R., 2001. Mansfeld's Encyclopedia of Agricultural and Horticultural Crops (except ornamentals). Institut für Pflanzengenetik und Kulturpflanzenforschung Gatersleben, Springer, Germany.

Herrero, R., Asins, M. J., Pina, J. A., Carbonell, E. A., Navarro, L., 1996. Genetic diversity in the orange subfamily *Aurantioideae*. II. Genetic relationships among genera and species. Theor Appl Genet. 93, 1327 1334.

Hodgson, R. W., 1967. Horticultural varieties of citrus. The Citrus Industry, vol. 1, second ed., University of California, Berkeley. 128-474.

Iwahori, S., Tominaga, S., Oohata, J. T., 1986. Ethychlozate accelerates coloration and enhances fruit quality of ponkan, *Citrus reticulata* Blanco. Scientia Horticulturae. 28, 243-250.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28, 1647-1649.

Kijas, J. M. H., Thomas, M. R., Fowler, J. C. S., Roose, M. L., 1997. Integration of trinucleotide microsatellites into a linkage map of *Citrus*. Theor. Appl. Genet. 94, 701-706.

Lado, J., Rodrigo, M. J., Zacarias, L., 2014. Maturity indicators and citrus fruit quality. Stewart Postharvest Review. 2, 3.

Nagy, S., Poczai, P., Cernák, I., Gorji, A. M., Hegedűs, G., Taller, J., 2012. PICcalc: an online program to calculate polymorphic information content for molecular genetic studies. Biochemical Genetics. 50, 670-672.

Nicolosi, E., Deng, Z., N., Gentile, A., La Malfa, S., Continella, G., Tribulato, E., 2000. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theor. Appl. Genet. 100, 1155–1166.

Peakall, R., Smouse P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics. 28, 2537-2539.

Peakall, R., Smouse P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes. 6, 288-295.

Polat, I., Kacar, Y. A., Yesilogli, T., Uzun, A., Tuzcu, O., Gulsen, O., Incesu, M., Kafa, G., Turgutoglu, E., Anil, S., 2012. Molecular characterization of sour orange (*Citrus aurantium*) accessions and their relatives using SSR and SRAP markers. Genet. Mol. Res. 11, 3267-3276.

Protti, M., Valle, F., Poli, F., Raggi, M. A., Mercolini, L., 2015. Bioactive molecules as authenticity markers of Italian Chinotto (*Citrus* × *myrtifolia*) fruits and beverages. Journal of Pharmaceutical and Biomedical Analysis. 104, 75–80.

Rocchi, L., Rustioni, L., Failla, O., 2016. Chlorophyll and carotenoid quantifications in white grape (*Vitis vinifera* L.) skins by reflectance spectroscopy. Vitis. 55, 11–16.

Rustioni, L., Maghradze, D., Popescu, C. F., Cola, G., Abashidzem E., Aroutiounian, R., Brazao, J., Coletti, S., Cornea, V., Dejeu, L., Dinu, D., Eiras Dias, J. E., Fiori, S., Goryslavets, S., Ibanez, J., Kocsis, L., Lorenzini, F., Maletic, E., Mamasakhlisashvili, L., Margaryan, K., Mdinaradze, I., Memetova, E., Montemayor, M. I., Munoz-Organero, G., Nemeth, G., Nikolaou, N., Pastore, G., Preiner, D., Raimondi, S., Risovanna, V., Sakaveli, F., Savin, G., Savvides, S., Schneider, A., Schwander, F., Spring, J. L., Ujmajuridzem L., Zioziou, E., Maul, E., Bacileri, R., Failla, O., 2014. First results of the European grapevine collections' collaborative network: validation of a standard eno-carpological phenotyping method. Vitis. 53, 219–226 (b).

Rustioni, L., Milani, C., Parisi, S., Failla, O., 2015. Chlorophyll role in berry sunburn symptoms studied in different grape (*Vitis vinifera* L.) cultivars. Scientia Horticulturae. 185, 145-150.

Rustioni, L., Rocchi, L., Guffanti, L., Cola, G., Failla, O., 2014. Characterization of Grape (*Vitis vinifera* L.) Berry Sunburn Symptoms by Reflectance. J. Agric. Food Chem. 62, 3043−3046 (a).

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-25.

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nature 18, 1-2.

Scora, R., W., 1975. On the history and origin of *Citrus*. Bull. Torr. Bot. Club. 102, 369-375

Singleton, V. L., Orthofer, R., Lamuela-Raventos, R. M., 1999. Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. Methods in Enzymology. 299, 152-178.

Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Kongelige Danske Videnskabernes Selskab. 5, 1-34.

Swingle, W. T., Reece, P.C., 1967. The botany of *Citrus* and its wild relatives. The Citrus Industry, vol. 1., University of California, Berkeley. 190– 430.

Tanaka, T., 1961. Citrologia Semi-Centennial Commemoration Papers on Citrus Studies. Citrologia Supporting Foundation, Osaka, Japan. 114.

Tanaka, T., 1977. Fundamental discussion of Citrus classification. Stud. Citrol. 14, 1–6.

Wu, A. G., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., Aurelio, Takit, M. A., Labadie, K., Poulain, J., Couloux, A., Jabbari, K., Cattonaro, F., Del Fabbro, C., Pinosio, S., Zuccolo, A., Chapman, J., Grimwood, J., Tadeo, F. R.,

Estornell, L. H., Munoz-Sanza, J. V., Ibanez, V., Herrero-Ortega, A., Aleza, P., Perez-Perez, J., Brunel, D. R. D., Luro, F., Chen, C., Farmerie, W. G., Desany, B., Kodira, C., Mohiuddin, M., Harkins, T., Fredrikson, K., Burns, P., Lomsadze, A., Borodovsky, M., Reforgiato, G., Freitas-Astua, J., Quetier, F., Navarro, L., Roose, M., Wincker, P., Schmutz, J., Morgante, M., Machado, M. A., Talon, M., Jaillon, O., Ollitrult, P., Gmitter, F., Rokhsar, D., 2014. Sequencing of diverse mandarin, pummel and orange genomes reveals complex history of admixture during citrus domestication. Nature Biotechnology. 32, 656-663.

Zur, Y., Gitelson, A. A., Chivkunova, O. B., Merzlyak, M. N., 2000. The spectral contribution of carotenoids to light absorption and reflectance in green leaves. Papers in Natural Resources. Papers 272. http://digitalcommons.unl.edu/natrespapers/272

5 APPLICATION AND OPTIMIZATION OF MARKER ASSISTED SELECTION FOR SHARKA RESISTANCE IN *PRUNUS ARMENIACA*

5.1 Introduction

Described by Atanasoff in 1933 (Atanasoff, 1933Sharka), Sharka disease was observed for the first time on plum trees in the south-western corner of Bulgaria in 1917. With a global estimated cost of 10 billion Euros in the last 30 years for its management, Sharka is considered one of the most devastating diseases among stone fruits trees (Cambra et al., 2006), and its pathological agent, the *Plum Pox Virus* (PPV), belongs to the top ten plant viruses (Scholthof et al., 2011). Since its discovery, the virus spread into most temperate fruit crop-growing areas (Capote et al., 2006) establishing himself in Europe, North and South America, Asia, and North Africa. Physiological disorder induced by PPV infection affects the sugar metabolism of plants (Baumgartnerova et al., 1998) resulting in increased fruit acidity (Sutic, 1971). Infected fruits may drop prematurely or may show various defects as deformation, presence of chlorotic spots and necrotic areas (hence the name of virus) (Nèmeth, 1986), making themselves unsuitable for consumption or industrial processing.
Sharka spread over long distances is mainly due to the transport of infected material or illegal exchanges (Cambra et al., 2006). But once the PPV is established in orchards, it is transmitted by aphids in a non persistent manner, thus chemical treatments are not effective in prevent virus propagation (Gildow et al., 2004; Labonne et al., 1995). Many other strategies exist to manage epidemics such as eradication of infected trees and containment of the virus. However these measures are often inefficient because of the time lapse between the virus infection and the appearance of the first symptoms (Martínez-Gómez et al., 2000), and thus Sharka is now epidemic in most of the infected regions. In this context the development of *Prunus* varieties that are resistant to PPV, could be the most promising long-term solution.

Today, high level of PPV resistance among cultivated stone fruit was only found in *Prunus armeniaca* (apricot), for the most part in cultivars released from North American breeding programs (Martínez-Gómez et al., 2000) that, however, seems to share the same mechanism of resistance introgressed from Chinese wild germplasm (Zhebentyayeva et al., 2008). Starting from the early 90s, these cultivars were used as donor of resistance in conventional breeding programs by crossing them with the best local cultivars susceptible to virus (Badenes and Llacer 2006; Bassi, 2006, Bassi and Audergon 2006; Karayiannis, 2006, Rubio et al., 2004). Although the introduction of PPV resistance through breeding is considered a good strategy, it is subject to several bottlenecks that

prevent the breeders to find promising cultivars in the short-therm. The long juvenile period and the space demands of traditional breeding are expensive and time-consuming and the phenotyping of PPV resistance is based on a test that requires three to four growing seasons for visual inspection, several replicates, and ELISA/RT-PCR tests (Lommel et al. 1982; Wetzel et al., 1991). The availability of a marker-based system called Marker Assisted Selection (MAS) for screening and prediction of resistance to PPV, has made a decisive contribution to boost the conventional breeding processes.

It was previously found, basing on several linkage maps and Quantitative Trait Locus studies, that Sharka resistance in apricot is a quantitative trait controlled by multiple genes (Lambert et al., 2007; Soriano et al., 2008; Lalli et al., 2008; Marandel et al., 2009, Pilarova et al., 2010; Dondini et al., 2011). However one main dominant locus located on the Linkage group 1 (LG1) and nominated *PPVres*, found consensus on all maps reported and explains from 25.4 up to 87% of the phenotypic variation among apricots (Pilarova et al., 2010; Soriano et al., 2008). *PPVres* region on LG1 was then fine mapped (Vera Ruiz et al., 2011) and PCR-based Short Sequence Repeat (SSR) markers linked with that locus were developed (Soriano et al., 2012). Subsequent sequencing and assembling of the region identified a 5-base pairs deletion on a candidate gene called "ppb022195m" (Zuriaga et al., 2013), which is putatively involved in virus translocation through the plant. It was supposed that the deletion results in a premature stop codon in the translated protein preventing the successful translocation of the virus between host cells thus conferring PPV resistance (Zuriaga et al., 2013). Based on this, a Single Sequence Length Polymorphism (SSLP) marker was developed (ZP002) and tested, together with three SSR markers (PGS1.21, PGS1.23 and PGS1.24) spanning the 200-kpb-long *PPVres* genomic region (Soriano et al., 2012), on 11 mapping populations, in order to validate its capacity to predict PPV resistance (Decroocq et al., 2014). It resulted that the marker set was not sufficient to unambiguously select PPV resistant cultivars (Decroocq et al., 2014). These markers are used by researchers and breeders to screen their populations because, even if they are not sufficient to completely predict resistance, they are necessary to mark susceptible plants.

In this work the markers, PGS1.21, ZP002, and PGS1.24 (Soriano et al., 2012, Zuriaga et al., 2013, Decroocq et al., 2014) were tested on 184 breeding selections in order to perform a screening for PPV resistance. In addition to this, these selections were also tested with a new set of four markers that enlarge the *PPVres* area to other genomic regions identified in Mariette et al. (2015) and Marandel et al. (2009), in order to test their reliability and possibly include them into the MAS standard protocol for Sharka resistance. Moreover, because of the common origin of the introgressed mechanism of PPV resistance (Zhebentyayeva et al., 2008), it is now important to

identify new sources of resistance in still-unknown germplasm. For this purpose 119 accessions and 14 well-known resistant/susceptible cultivars from an Italian germplasm collection, were tested with both marker sets. The outcomes of this work will establish the effectiveness of the new markers in MAS, will pick out the most promising breeding selections and will allow the breeders to enrich and complement with new original plant material, the current apricot germplasm available for breeding programs addressing PPVresistance.

5.2 Materials and Methods

5.2.1 Plant material

Young leaves of 183 breeding selections, 119 new accessions and 14 well-known varieties used as control, were sampled (Supplementary Tables 2 and 3) from CRPV (Centro Ricerche Produzioni Vegetali) experiment station (Faenza, Italy) using a leaf-disc sampler. The samples were then lyophilized and stored at room temperature until DNA extraction. Moreover, 22 samples of the 'Lito' x 'BO81604311' F1 apricot progenies (previously sampled and extracted by Dondini et al. 2007) were selected (Supplementry Table 1).

5.2.3 DNA Extraction

High-molecular weight genomic DNA was extracted from ground leaf samples using the Plant DNA Isolation Kit (BPI-TECH, USA) according to the manufacturer's instructions. Nucleic acids were resolved on a 1% (w/v) agarose, TAE-buffered gels and visualized by ethidium bromide staining under UV light (365 nm). Size and quantity of nucleic acid bands were estimated by comparison to HindIII-cut Lambda (Thermo Scientific, USA) and 100bp (Thermo Scientific, USA) molecular ladders.

5.2.4 SSR analysis

A set of four markers (PGS1.10, PGS1.21, ZP002, and PGS1.24) targeting the *PPVres* locus (Figure 1) (Soriano et al., 2012; Zuriaga et al., 2013; Decroocq et al., 2014) was chosen, together with new primers previously designed in the frame of the UE "MARS" project (Marker Assisted Resistance to Sharka, https://www6.inra.fr/mars) into the genomic regions identified by Mariette et al. (2015) and Marandel et al. (2009), AMPPG016, AMPPG021 and SRLG_11m52 (Table 1, Figure 1). PGS1.21 and PGS1.24 were labelled on the 5' position of the forward primer as shown in Table 1. In this case the PCR mixture consisted of 1 ng/µl template DNA, 0.10 mM dNTPs, 0.25 mM forward primer, 0.25 mM reverse primer, 1X PCR buffer (Eurx LTD, Poland), 2.5 mM $MgCl_2$ and 0.025 U/µl *Taq* DNA polymerase (Eurx LTD, Poland) in a final volume of 20 µl. The other primer pairs were marked according to the Tail PCR technique (Schuelke, 2000) (Table 1) and the PCR reaction was performed assembling a mixture of 1 ng/µl template DNA, 0.1 mM dNTPs, 0.1 mM forward primer, 0.25 mM reverse primer, 0.25 mM labelled primer, 1X PCR buffer (Eurix, Poland), 2.5 mM $MgCl_2$ and 0.025 U/µl *Taq* DNA polymerase (Eurx LTD, Poland) in a final volume of 20 µl. The reactions where carried out with the following programs according to the

annealing temperature for each primer pairs (Table 1): 5 min at 94°, 2 cycles of 30 s at 94°, 1.30 min at 55° and 1 min at 72°, 35 cycles of 30 s at 94°, 30 s at 55° and 30 s at 72° with a final extension of 10 min at 72°, for markers PGS1.21, PGS1.24 and ZP002 and 5 min at 94°, 35 cycles of 30 s at 94°, 1.30 s at 57° and 30 s at 72° with a final extension of 5 min at 72° for the other markers. The size of the PCR products was scored through capillary electrophoresis on a genetic analyzer AB3730 (Thermo Scientific, USA).

Table 1. Primers used for the SSR analysis.

Name	F	Type	Size	Position	Ta	Forward primer	Reverse primer
PGS1.10[1]	pet	SSR	232-274	7847912	57	TGTAAAACGACGGCCAGTGCC CTTTAATCCCAAGGAAG	GCAGGGCTTGCTC TATTCAC
PGS1.21[1]	pet	SSR	172-220	8527745	55	CCCTGGTGTTCTGCTCTCTC	CATCCACAAATGG GAAGCAT
ZP002[2]	fam	SSLP	107-112	8607078	55	TGTAAAACGACGGCCAGTAA CATTTTCTGATTCAATGCCA	TGTATCCTCCAGC TTCAAAGTC
PGS1.24[1]	ned	SSR	101-141	8668808	55	GTAAATGAGTGCCTGCGTGT	TGCGAGAGTTGTG ATTGATG
AMMPG016[3]	ned	SSR	165-225	9611309	57	TGTAAAACGACGGCCAGTTGG TGATGCTAATGGCAAGA	CATGGTCTCTTCC CGTGACT
AMMPG021[3]	fam	SSR	206-240	10060228	57	TGTAAAACGACGGCCAGTTTT TCTATGGTCGGCTTTGG	AGCCTCCAAAAA GCAGTGTC
SSRLG1_11m52[3]	vic	SSR	197-213	11972362	57	TGTAAAACGACGGCCAGTTAG ATAAGCCCACCAATTGTCA	GCATATACATCCA AAGGAAGCC

F: fluorophore. Ta: annealing temperature. [1]Soriano et al., 2012, [2]Decroocq et al 2014, [3] designed in the frame of the UE "MARS" project (Marker Assisted Resistance to Sharka). Tail sequences are underlined.

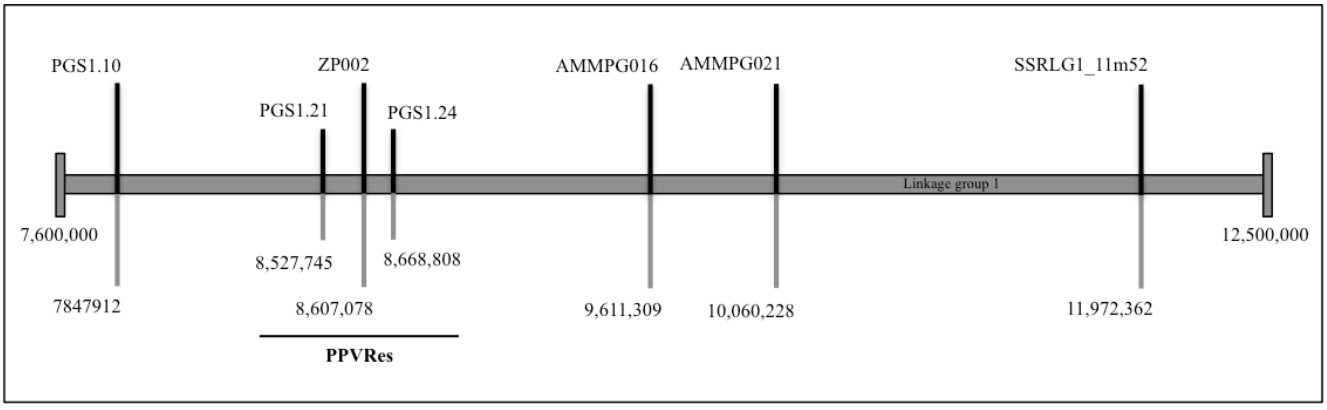


Figure 1. Schematic representation of the PPVRes locus together with the new regions identified by Mariette et al. (2015) and Marandel et al. (2009) on linkage group 1 of the apricot genome. Positions are expressed in base pairs.

5.2.5 Data analysis

The output files were visualized and scored using the software Geneious 9.0 (Kearse et al., 2012) available at: http://www.geneious.com. The resulting genotyping data were then formatted and analysed using GenAlEx (Peakall and Smouse 2006, 2012; http://biology-assets.anu.edu.au/GenAlEx/Download.html) PICcalc (Nagy et al., 2012; http://w3.georgikon.hu/pic/english/default.aspx) and Microsoft Excel (Office Suite). Statistical analyses were performed by PAST3 software (Hammer et al., 2001; http://folk.uio.no/ohammer/past) and PLINK software (Purcell et al., 2007; http://pngu.mgh.harvard.edu/purcell/plink/).

5.3 Results and Discussion

5.3.1 Genetic Analysis

The Sharka virosis is one of the most devastating disease among *Prunus* species and all the efforts in the last 30 years have scarcely been able to keep under control its spread and to prevent relevant economic losses globally (Cambra et al., 2006). Today the most promising strategies are based on the development of resistant cultivars by introgression of sources of resistance from wild germplasm (Zhebentyayeva et al., 2008) to local varieties by traditional breeding. However the application of breeding in tree species is subject to several bottlenecks, of which a reduction of overall operative times is a central aspect. Moreover, it is also important to identify new sources of resistance from wild germplasm for pyramiding different mechanism of resistance in new varieties.

For this purpose, three markers tightly linked to the *PPVres* locus (Soriano et. al., 2012, Decroocq et. al., 2014) plus four markers identified in the frame of two European projects (SharCo and MARS) and located in the new genomic regions identified in Mariette et al. 2015 and Marandel et al. 2009 (Table 1, Figure 1), have been used to screen for Sharka resistance on 317 apricot trees composed by selections of breeding and new accessions (Supplemental Tables 2 and 3).

Table 2. Primers used for the analysis.

Locus	AS	OA	EA	OH	PIC	GP
PGS1.10	232-274	13.00	3.455	0.66	0.66	7,398,550
PGS1.21	172-220	17.00	6.584	0.87	0.83	8,078,385
ZP002	103-112	3.00	1.719	0.47	0.33	8,157,718
PGS1.24	101-141	12.00	4.688	0.82	0.75	8,219,448
AMMPG016	165-225	18.00	4.647	0.78	0.76	9,161,952-9,162,152
AMMPG021	206-240	16.00	6.258	0.86	0.82	9,610,870-9,611,109
SRLG_11m52	197-213	8.00	4.060	0.81	0.71	11,524,130-11,524,349
MEAN		12.43	4.487	0.75	0.70	

AS: allele size; **OA**: observed alleles; **EA**: effective alleles; **OH**: observed heterozygosity; **PIC**: polymorphism information content; **GP**: *Prunus persica* genome position.

All the seven markers resulted to be polymorphic, producing well-resolved fragments. A total of 87 alleles were detected, ranging from a minimum of 3 for marker ZP002 (2 expected alleles plus 1 unexpected allele found only in the breeding selection 'BO04624042'), and a maximum of 18 for marker AMMPPG016, and showing a mean allele number of 12.43, although the mean of effective alleles is lower (Table 2). As expected, excluding the ZP002, the other markers show a large discrepancy between the two values and so a great amount of low frequency-alleles. It is possible to explain this trend considering the germplasm origin of the plant samples underlying the great variability between them. This evidence gains support also from the Observed Heterozigosity (OH)

and Polymorphism Information Content (PIC). Indeed the percentage of heterozygotes per marker detected among apricot samples range from 33.2% (ZP002) to 86.9% (PGS1.21) and the mean OH for all markers is 75.3%. Moreover the PIC index ranges from 0.33 (ZP002) to 0.83 (PGS1.21) with a mean value of 0.70 for the entire marker-set. Among all markers PGS1.21 is the most informative one (Table 2).

In order to perform MAS analysis on the sampled trees, the size of the alleles associated with resistance was obtained from literature (where available), or was experimentally determined by mapping each marker on 22 samples of the 'Lito' (resistant) x 'BO81604311' (susceptible) F1 apricot cross population previously characterized by Dondini et al., (2007) (Supplemental Table 1). Pearson product-moment correlation analysis shows a statistically significant correlation of the alleles 240 (PGS1.10), 221 (AMMPG016), 226 (AMMPG021), and 206 (SSRLG_11M52) with the genotypes that are resistant to Sharka and of the allele 199 (SSRLG_11M52) with those not resistant to the disease (Table 3). Markers PGS1.21, PGS1.24 and ZP002 show allele sizes that systematically diverge from those reported in the literature (Soriano et. al., 2012; Decroocq et. al., 2014) reflecting an error of the fragment analyser or of the tail PCR technique (Scheulke 2000) (data not shown). However due to the systematic differences in sizes, and the evidences shown by the 14 well-known cultivar used as control, it was still possible to associate the correct allele to the resistance against Sharka (Supplemental Table 2, 3).

Table 3 Pearson product-moment correlation analysis of the alleles obtained from the mapped markers on the F1 apricot population 'Lito' x 'BO81604311'.

Marker	Allele size	*p* value	Pearson correlation coefficient
PGS1.10	**240**	**0.002**	**0.845**
	260	0.505	-0.142
	262	0.742	-0.070
AMMPG016	197	0.541	-0.130
	211	0.156	-0.298
	221	**0.002**	**0.579**
AMMPG021	**226**	**0.002**	**0.845**
	228	1	0
SSRLG_11m52	**199**	**0.011**	**-0.507**
	206	**0.006**	**0.542**
	213	0.245	-0.246

Alleles with significant phenotype correlation ($p < 0.05$) are underlined in bold.

5.3.2 Breeding Selections

Introgression of PPV resistance for crop improvement is one of the most important goals in apricots breeding programs. Due to time-consuming protocols, phenotyping for Sharka is still the

major bottleneck in the breeding pipeline. In this context, screening of plants with MAS using markers linked to PPV resistance, provide the best solution for enhancing breeding efficiency.

A total of 184 young leaves of apricot trees, representing selection of breeding programs for Sharka resistance, were genotyped, together with 14 well-known varieties used as control, with the seven markers proposed for MAS in apricot (Supplemental Table 2, 3). 39 of 184 breeding selections carry the allele linked to the resistance to Sharka (in any possible configuration) for all the seven markers, 49 of 184 for the standard *PPV*res locus (PGS1.21, ZP002 and PGS1.24) and 110 of 184 considering only the 5 bp deletion on the candidate gene (Zuriaga et al 2013). On average, 6,13 % of samples, for each marker, are homozygotes for the allele in linkage with the resistance, 41,54 % are heterozygotes, 48,76 % does not carry the resistant allele and 3.57 % showed an illegible or null PCR amplification pattern (Table 4A). This last evidence is probably due to mutations into the sequences identified by the primers pairs or rearrangement inside the *PPV*res locus.

The final outcome of the MAS for Sharka resistance could be substantially different basing on which and how many markers will be considered. Considering the standard MAS protocols in apricot (PGS1.21, PGS1.24 and ZP002) a list of 49 candidates were extrapolated in order to establish which hybrids could be selected for further PPV resistance tests, saving time and resources.

Table 4

Number of samples carrying different combinations of the resistant allele (homozygote, heterozygote or absent) for the selections of breeding (**A**) and the germplasm accessions (**B**). %: average percentage of resistant allele combinations for each markers.

A

Allele configuration	PGS1.10	PGS1.21	ZP002	PGS1.24	AMMPPG016	AMMPPG021	SSRLG_11m52	%
Homozygous res.	17	5	13	7	8	13	16	6.13
Heterozygous res.	80	46	97	80	77	79	76	41.54
Absent	84	129	74	88	83	81	89	48.76
Null	3	4	0	9	16	11	3	3.57

B

Allele configuration	PGS1.10	PGS1.21	ZP002	PGS1.24	AMMPPG016	AMMPPG021	SSRLG_11m52	%
Homozygous res.	18	0	5	2	2	1	2	3.60
Heterozygous res.	48	35	39	35	23	30	34	29.29
Absent	50	84	75	74	78	77	78	61.94
Null	3	0	0	8	16	11	5	5.16

5.3.3 Accessions

Very few sources of resistance have been identified in *Prunus* species, with most occurring in *P. armeniaca*, and originated from North American breeding programs (Martin Gomez and Dicenta, 2000). The limited number of known resistant accessions prevents the breeder to the development of effective and strong resistant cultivars by pyramiding different mechanism of resistance. In this context, the MAS technology was used to rapidly screen and select new promising donor of resistance from the germplasm collection of CRPV experiment station (Faenza, Italy).

A total of 119 young leaves of apricot accessions were tested, together with 14 well-known varieties used as control, with the seven markers proposed for MAS in apricot (Supplemental Table 3). 17 of 119 individuals carry the allele linked to the resistance to Sharka (in any possible configuration) for all the seven markers, 31 of 119 for the standard *PPV*res locus (PGS1.21, ZP002 and PGS1.24) and 44 of 119 considering only the 5 bp deletion on the candidate gene (Zuriaga et al 2013). On average, 3.60 % of samples, for each marker, are homozygotes for the allele in linkage with the Sharka resistance, 29.29 % are heterozygotes, 61.94 % does not carry the resistant allele, and 5.16 % showed an illegible or null PCR amplification pattern (Table 4B). This last evidence, as previously seen, could be due to mutations or rearrangements of the *PPV*res locus.

In comparison with the breeding selections, the accessions show a higher number of null samples and a fewer number of genetically putatively resistant plant, reflecting the higher variability of the *PPV*res locus that is not only inherited by the common donors available for breeding. The result of MAS on the accessions differs basing on which markers set is considered. Based on the standard MAS protocol in apricot (PGS1.21, PGS1.24 and ZP002), 31 candidates have been selected for further test against PPV resistance, to enrich and complement with new original plant material, the current apricot germplasm available for PPV-resistance breeding programs.

5.3.5 Phenotype and genotype comparison

The phenotype of 78 out of 317 samples, including breeding selections and accessions, was determined by CRPV-ASTRA in collaboration with the Plant Protection Service of the Emilia Romagna region (data not published) and it is still in progress. The phenotypes of the samples are depicted in Supplementary Table 2 and 3 together with those of the 14 well-known varieties used as control. In many cases the genotype is quite consistent with the phenotype reported. However ‘Pisana’, that are evaluated as phenotypically resistant, carry the resistant allele for just one marker (PGS1.10 and SRLG_11m52, respectively) and ‘BO04624039’, phenotypically resistant, does not carry any resistant allele. On the contrary, ‘BO96621021’ and ‘Luna’ carry the resistant allele

(heterozygous configuration) on the entire marker set and 'BO06609003' on all the markers excluding PGS1.10, but they are scored as phenotypically susceptible. 'BO06609053' is evaluated phenotypically resistant but carry the resistant allele only for the marker SRLG_11m52, however its genotype is not consistent with its parents revealing this evidence as a consequence of experimental or sampling error. These incongruences could be the result of a sampling error or a further demonstration that the marker set was not sufficient to unambiguously select PPV resistant cultivars (Decroocq et. al., 2014). Despite this, 'Pisana', 'BO06609053' and 'BO04624039' must be taken into account for further analysis because they could hide different mechanisms of resistance against Sharka.

The association shown by marker SRLG_11m52 with the disease is not statistically significant for both tests (for $p < 0.01$), logistic regression and chi-square, considering the 78 already phenotyped samples (Table 5, Supplementary Table 2 and 3). Thus we do not report the presence of the "PPV1b" locus for Sharka resistance, previously discovered by Mariette et al. (2015), supposed to be in linkage with the SSRLG_11m52 SSR marker. It was also surprising that the region identified by marker PGS1.21, now commonly adopted in the standard MAS protocols in apricot, seems to be not statistically associated with the resistance (Table 5). Excluding the cases reported above, when the genomic region underlined by ZP002 and PGS1.24 carry the resistant alleles, the samples are phenotypically resistant and viceversa and this evidence is supported by the Odds Ratio values of logistic regression at both markers (Table 5), supporting them as the only suitable markers for the application of MAS in apricot. Markers PGS1.10, AMMPPG016 and AMMPG021, show low Odds Ratio values and thus they were excluded from further analysis (Table 5).

These results support the convenience to use MAS as a tool to predict the resistance to Sharka in apricot even if the system is not always able to unambiguously select PPV resistant cultivars. However we found that using 2 markers out of 7 is sufficient to assure a good prediction of Sharka resistance in unknown apricot trees.

Table 8. Logistic regression and Chi-square tests for the 78 accessions and selections of breeding already phenotyped.

Marker	OR	SE	L95	U95	*p* value	Chi-squared	*p* (chisq.)
PGS1_10	12.62	0.6728	3.376	47.19	0.0001643	9.984	0.00157
PGS1_21	3.57	0.5845	1.137	11.25	0.02923	5.182	0.02282
PGS1_24	44.16	0.7652	9.856	197.80	7.41E-04	18.66	0.0156
ZP002	43.92	0.7886	9.363	206.00	1.62E-03	17.47	0.0292
AMMPPG016	33.27	0.746	7.711	143.60	2.63E-03	19.97	7.87E-06
AMMPPG021	23.45	0.6894	6.072	90.57	4.73E-06	17.31	0.0318
SSRLG_11m52	2.95	0.507	1.092	7.96	0.03284	4.062	0.04385

OR: odds ratio; **SE**: standard error; **L95**: lower confidence interval; **U95**: upper confidence interval.

5.4 Conclusions

Although many published studies claim the development of new markers with a potential for MAS in only few cases a practical exploitation of this technology follows. Here it is presented the first large-scale application of markers developed and selected in the frame of two European projects for the screening for Sharka resistance in apricot. Forty-nine breeding selections and thirty-one accessions were selected to be further analyzed because positive for a major standard locus mutation (*PPVres*) saving the considerable resources and time otherwise needed for the traditional field- and greenhouse-based analysis of the discarded plants. In this way new putative candidates were preliminary selected to enrich and complement the current apricot germplasm available for PPV-resistance breeding programs.

The number of markers needed for MAS can be reduced to 2 out of 7, since not significant associations were observed for markers located on the PPV1b secondary locus previously discovered by Mariette et al (2015). Such limited number of effective markers paves the way to a reliable and cost-effective application of this technology also by small laboratories or nurseries.

This work represents a first step into the removal of the barriers which normally separate the basic knowledge from its practical, in-field application. Together with the optimization of the phenotpying methods, the data presented in this work could translate into a real boost into the effective development of resitant cultivars against PPV virus.

5.5 Supplementary Materials

Supplementary Table 1 Genotyping of 22 samples of the 'Lito' x 'BO81604311' F1 population with selected SSR markers

Pop.	PGS1.10	AMMPG016	AMMPG021	SSRLG_11m52
1	240/260	211/221	226/228	206/213
2	260/260	0/0	228/228	199/213
3	260/262	197/211	228/228	199/213
4	260/262	0/0	228/228	199/213
5*	240/260	0/0	226/228	206/213
6*	240/262	197/221	226/228	206/213
7	260/260	197/211	228/228	199/213
8	240/262	211/221	226/228	206/213
9	260/260	211/211	228/228	199/213
10*	240/260	0/0	226/228	206/213
11	260/260	211/211	228/228	199/213
12	260/262	197/211	228/228	199/213
13*	240/260	211/221	226/228	206/213
14*	240/262	211/221	226/228	199/213
15	240/260	211/221	228/228	206/213
16*	240/262	211/221	226/228	206/213
17	260/262	197/211	228/228	199/213
18*	240/262	211/221	226/228	206/213
19	260/262	197/211	228/228	199/213
20*	240/260	197/221	226/228	206/213
21	260/262	0/0	228/228	199/213
22*	240/262	197/221	226/228	206/213
L*	240/260	211/221	226/228	199/206
S	260/262	197/211	228/228	213/213

L: 'Lito'; S: 'BO81604311'; *: samples phenotypically resistant to Sharka

Supplementary Table 2. Genotypes and phenotypes reported for the 184 apricot selections and the whole SSR marker set tagging the resistance region.

Offspring code	PGS1.10	PGS1.21	ZP002	PGS1.24	AMMPG016	AMMPG021	SSRLG_11m52	Phenotype
BO96621002	240/240	220/220	107/107	101/101	221/221	226/226	206/206	resistant[1]
BO03615049	240/240	220/220	107/107	101/101	221/221	226/226	206/206	resistant[1]
BO03615025	240/240	220/220	107/107	101/101	221/221	226/226	206/206	resistant[1]
BO03615070	240/240	220/220	107/107	101/101	221/221	226/226	206/206	resistant[1]
BO03615034	240/240	220/220	107/107	101/101	221/221	226/226	206/206	resistant[1]
BO04626008	240/240	218/220	107/107	101/101	221/221	226/226	206/213	
BO89606006	240/240	218/220	107/107	101/101	221/221	226/226	206/213	
BO04639050	240/240	188/220	107/107	101/109	197/221	224/226	206/206	
BO04639319	240/240	188/220	107/107	101/109	197/221	224/226	206/206	
BO04639270	240/240	188/220	107/107	101/109	197/221	224/226	206/206	
BO04639076	240/240	188/220	107/107	101/109	197/221	224/226	206/213	
BO04639056	240/240	188/220	107/107	101/109	197/221	224/226	206/209	

BO03605009	240/240	188/210	107/107	101/109	175/197	224/226	206/213	
BO03621023	240/240	174/220	107/112	101/103	197/221	224/226	206/213	
BO02614009	240/240	200/220	107/112	101/140	221/225	226/234	206/213	
BO02614036	240/260	198/220	107/112	101/105	175/221	226/226	206/206	
BO02629005	240/260	174/220	107/112	101/103	207/221	226/226	206/206	
BO02615033	240/260	198/220	107/112	101/105	175/221	226/226	199/206	
BO03624004	240/260	198/220	107/112	101/105	175/221	226/226	199/206	
BO04639109	240/262	190/220	107/112	101/105	197/221	226/228	206/213	
BO04628009	240/260	174/220	107/112	101/103	197/221	224/226	206/213	
BO04630014	240/260	174/220	107/112	101/103	197/221	224/226	206/213	
BO04627011	240/274	190/220	107/112	101/105	175/221	216/226	206/209	
BO04627008	240/274	190/220	107/112	101/105	175/221	216/226	206/209	
BO93623033	240/242	190/220	107/112	101/132	189/221	226/240	206/209	
BO04640005	240/274	190/220	107/112	101/105	175/221	216/226	206/209	
BO04639287	240/260	180/220	107/112	101/132	197/221	218/226	206/213	
BO05634091	240/260	190/220	107/112	101/132	175/221	226/234	199/206	
BO05634055	240/260	174/220	107/112	101/103	211/221	226/228	199/206	
BO04639073	240/260	180/220	107/112	101/132	197/221	218/226	206/213	
BO04639389	240/260	180/220	107/112	101/132	197/221	218/226	206/213	
BO93623012	240/242	190/220	107/112	101/132	189/221	226/238	206/209	
BO05636099	240/260	180/220	107/112	101/132	197/221	218/226	206/213	
BO05636034	240/260	190/220	107/112	101/105	197/221	218/226	206/213	resistant[1]
BO96621021	240/260	180/220	107/112	101/132	197/221	218/226	206/213	susceptible[1]
BO06603049	240/242	190/220	107/112	101/132	189/221	226/240	206/209	
BO99612023	240/260	190/220	107/112	101/105	175/221	226/228	206/213	
BO96621030	240/260	174/220	107/112	101/103	211/221	226/228	199/206	resistant[1]
BO03615053	240/260	180/220	107/112	101/132	197/221	218/226	206/213	resistant[1]
BO03615011	240/260	180/220	107/112	101/132	197/221	218/226	206/213	resistant[1]
BO07633139	240/260	174/220	107/112	101/103	211/221	226/228	213/213	
BO02614073	240/260	198/220	107/112	101/105	175/221	0/0	199/206	
BO05634191	240/260	174/220	107/112	101/103	197/221	0/0	197/206	
BO05634114	240/274	190/220	107/112	101/105	175/221	214/224	197/206	
BO02611038	240/260	180/220	107/112	101/132	0/0	224/226	206/206	
BO04616003	240/260	174/220	107/112	101/103	0/0	226/228	206/213	
BO99612020	240/260	174/220	107/112	101/103	0/0	6/228	206/213	
BO04639366	240/260	180/220	107/112	101/132	197/197	218/226	209/213	
BO04639350	236/240	200/220	107/112	101/140	191/225	218/234	199/209	
BO04624031	240/260	174/220	107/112	101/103	197/211	218/228	213/213	resistant[1]
BO92639007	240/240	200/220	107/112	0/0	191/221	218/226	199/206	
BO05636033	240/274	0/0	107/112	101/105	197/221	218/226	206/213	
BO04639027	274/274	0/0	107/112	105/109	175/197	216/224	206/209	
BO92645021	240/240	200/218	107/112	101/103	191/221	218/226	199/213	
BO02629001	240/242	190/218	107/112	101/132	221/221	226/226	206/213	
BO04618001	240/260	174/200	107/112	101/103	211/221	226/228	206/213	
BO04602023	240/260	180/218	107/112	101/132	197/221	218/226	206/213	
BO05637010	240/260	190/218	107/112	101/132	175/221	226/234	199/206	
BO05636072	240/260	180/218	107/112	101/132	197/221	218/226	206/213	
BO06609003	240/260	174/218	107/112	101/103	197/221	224/226	197/206	susceptible[1]
BO06609083	240/260	174/218	107/112	101/103	197/221	224/226	197/206	resistant[1]
BO06609133	240/274	190/218	107/112	101/105	175/221	216/226	197/206	resistant[1]
BO06609037	240/260	174/218	107/112	101/103	197/221	224/226	197/206	resistant[1]

BO06609113	240/260	174/218	107/112	101/103	197/221	224/226	197/206	resistant[1]
BO06609068	240/274	190/218	107/112	101/105	175/221	216/226	0/0	resistant[1]
BO05634251	240/260	190/218	107/112	101/132	175/221	226/234	199/213	
BO05634173	240/260	174/218	107/112	101/103	211/221	226/228	199/213	
BO05633050	240/260	180/218	107/112	101/132	197/221	218/226	213/213	
BO05636125	240/274	190/218	107/112	101/105	197/221	218/226	213/213	
BO05634027	240/260	174/218	107/112	101/103	197/221	224/226	209/213	
BO05636075	240/274	190/218	107/112	101/105	197/221	218/226	213/213	
BO05637017	240/260	174/218	107/112	101/103	211/221	226/228	199/213	
BO06609055	240/260	174/218	107/112	101/103	197/221	224/226	209/213	resistant[1]
BO06609129	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609048	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609087	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609079	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609074	240/274	190/218	107/112	101/105	175/221	216/226	209/213	resistant[1]
BO06609033	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609060	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609099	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609036	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609024	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609136	240/274	190/218	107/112	101/105	175/221	216/226	209/213	resistant[1]
BO06609013	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609039	240/260	174/218	107/112	101/103	197/221	224/226	197/213	resistant[1]
BO06609104	240/274	190/218	107/112	101/105	175/221	216/226	209/213	resistant[1]
BO06603087	242/242	190/190	107/112	101/132	189/221	226/240	206/209	
BO96626089	260/260	196/218	107/112	101/103	189/197	216/218	206/209	
BO04610060	236/240	200/218	107/112	0/0	221/225	226/234	209/213	
BO06609012	240/260	174/218	107/112	0/0	197/221	224/226	197/213	resistant[1]
BO06609045	240/260	174/218	107/112	0/0	197/221	224/226	197/213	resistant[1]
BO92636056	260/260	174/188	107/112	0/0	0/0	224/228	206/213	
BO03605095	0/0	188/200	107/112	105/109	0/0	224/226	206/209	
BO03605036	0/0	188/200	107/112	105/109	191/197	224/226	206/209	
BO04635016	242/242	188/190	107/112	109/132	0/0	224/240	206/209	
BO02611035	260/260	188/196	107/112	103/109	197/197	218/224	206/206	
BO02611054	260/260	188/196	107/112	103/109	197/197	218/224	206/206	
BO02611043	260/260	188/196	107/112	103/109	197/197	218/224	206/206	
BO02611043	260/260	188/196	107/112	103/109	197/197	218/224	206/206	
BO04639402	274/274	188/190	107/112	105/109	175/197	216/224	206/209	
BO04602026	260/260	180/188	107/112	109/132	197/197	224/224	197/206	
BO04639402	274/274	188/190	107/112	105/109	175/197	216/224	206/209	
BO04639227	274/274	188/190	107/112	105/109	175/197	216/224	206/209	
BO04639405	274/274	188/190	107/112	105/109	175/197	214/222	206/209	
BO04635015	242/242	188/190	107/112	109/132	189/197	224/240	206/209	
BO02609006	260/260	180/188	107/112	109/132	197/197	218/224	206/213	
BO14604009	260/260	174/188	107/112	103/109	197/197	224/228	199/206	
BO14604011	264/264	174/188	107/112	103/109	197/211	224/228	199/206	
BO99608003	246/266	190/202	107/112	105/130	175/193	208/218	199/211	
BO03605116	236/260	188/220	112/112	103/140	0/0	228/234	209/213	
BO14604010	0/0	0/0	112/112	0/0	0/0	0/0	199/206	
BO03615019	250/258	0/0	112/112	0/0	173/181	222/234	0/0	
BO06628081	236/260	180/200	112/112	0/0	197/225	218/234	213/213	

BO04640015	240/242	190/200	112/112	103/132	0/0	220/240	209/213	
BO04639267	240/260	180/190	112/112	105/132	175/197	216/218	209/213	
BO95640011	240/260	174/206	112/112	103/103	195/211	228/228	211/213	
BO02602019	240/260	174/174	112/112	103/103	211/211	216/228	199/213	
BO86606014	236/240	190/200	112/112	132/140	175/225	206/234	197/209	
BO04614003	240/260	174/200	112/112	103/103	191/211	218/228	199/213	
BO07628820	240/260	180/200	112/112	103/132	191/197	218/218	199/213	
BO99610006	240/260	180/206	112/112	132/136	195/197	218/228	211/213	
BO06603111	242/242	190/190	112/112	132/132	189/221	226/240	206/209	
BO07608013	260/260	174/180	112/112	103/132	197/221	218/226	206/213	
BO03605044	260/260	180/200	112/112	105/132	0/0	218/226	209/213	
BO01616013	236/262	190/200	112/112	105/140	0/0	0/0	209/213	
BO02615061	242/260	190/198	112/112	105/132	0/0	0/0	199/209	
BO01603008	260/260	174/180	112/112	103/132	0/0	218/228	213/213	
BO95612103	236/274	190/200	112/112	105/140	0/0	216/234	209/209	
BO07628501	260/260	174/180	112/112	103/132	0/0	218/228	213/213	
BO02615090	242/260	190/198	112/112	105/132	175/189	226/226	199/206	
BO04602038	260/260	180/180	112/112	132/132	197/197	218/226	206/206	
BO04637020	260/260	174/174	112/112	103/103	197/197	224/226	197/213	
BO87619001	260/260	174/180	112/112	103/132	197/197	0/0	213/213	
BO86604052	260/260	174/190	112/112	103/132	175/211	0/0	209/213	
BO86617004	260/260	174/190	112/112	103/132	175/175	0/0	199/213	
BO01622110	260/260	180/180	112/112	132/132	185/197	0/0	213/213	
BO06621176	260/260	174/174	112/112	103/103	197/211	0/0	213/213	
BO02611033	260/260	180/196	112/112	103/132	197/197	218/218	206/213	
BO03610002	236/260	174/200	112/112	103/140	211/225	228/234	206/209	
BO01617063	260/262	190/196	112/112	103/105	197/197	218/228	206/213	
BO02611040	260/260	180/196	112/112	103/132	197/197	218/218	206/213	
BO03624001	236/260	196/200	112/112	103/140	197/225	218/234	206/213	
BO02611006	260/260	180/196	112/112	103/132	197/197	218/218	206/213	
BO07642006	274/274	190/190	112/112	105/105	175/197	216/224	206/209	
BO06609053	260/274	174/190	112/112	103/105	175/211	216/228	206/209	resistant[1]
BO04624043	260/274	180/190	112/112	105/132	175/197	216/218	206/209	susceptible[1]
BO04639125	260/260	180/180	112/112	132/132	197/197	218/218	213/213	
BO03628004	260/274	180/190	112/112	105/132	175/197	216/218	209/213	
BO04639261	260/274	180/190	112/112	105/132	175/197	216/218	209/213	
BO95602024	236/260	174/200	112/112	103/141	211/223	228/234	209/213	
BO01607174	262/274	190/190	112/112	105/105	175/197	216/228	209/213	
BO86613038	260/260	180/190	112/112	132/132	175/197	218/234	199/204	
BO86606179	236/236	190/200	112/112	105/140	207/225	206/234	197/209	
BO95651020	260/262	174/180	112/112	103/132	197/197	218/224	209/213	
BO02602009	260/260	174/174	112/112	103/103	197/211	228/228	213/213	
BO02602008	260/274	174/190	112/112	103/105	175/211	216/228	209/213	
BO03627079	260/274	174/190	112/112	103/105	175/211	218/228	209/213	
BO02602012	260/274	174/190	112/112	103/105	175/175	216/228	209/213	
BO02602008	260/274	174/190	112/112	103/105	175/175	216/228	209/213	
BO94668002	260/260	180/180	112/112	107/132	197/197	218/218	213/213	
BO92618086	262/262	190/190	112/112	105/105	197/197	228/228	213/213	
BO86608282	236/260	180/200	112/112	132/140	197/225	218/234	209/213	
BO91620003	260/262	174/180	112/112	103/132	175/197	218/220	199/213	
BO86606183	236/260	180/200	112/112	132/140	197/197	218/218	213/213	

BO86615156	262/274	190/190	112/112	105/105	175/197	216/228	209/213	
BO92657005	236/260	180/200	112/112	132/140	197/197	218/234	209/213	
BO03627040	260/274	174/190	112/112	103/105	175/211	218/228	209/213	
BO04635018	260/260	174/174	112/112	103/103	197/197	224/224	197/213	
BO02632039	260/262	174/190	112/112	103/105	197/211	228/228	209/213	
BO92637005	260/260	174/180	112/112	103/132	197/225	224/234	209/213	
BO04635036	242/260	174/190	112/112	103/132	189/197	224/238	209/213	
BO03627040	236/260	174/200	112/112	103/140	197/225	222/234	197/209	
BO06621228	260/260	174/174	112/112	103/103	197/197	218/224	209/213	
BO06619151	242/260	174/190	112/112	103/132	189/197	224/238	209/213	
BO06628112	260/260	174/180	112/112	103/132	197/197	218/224	213/213	
BO06613160	260/274	180/180	112/112	132/132	197/197	216/216	213/213	
BO06628023	260/274	174/190	112/112	103/105	197/197	218/224	213/213	
BO07641136	260/260	180/180	112/112	132/132	197/197	218/218	213/213	
BO92637005	260/260	174/180	112/112	103/132	197/225	224/234	209/213	
BO04624042	260/274	180/190	103/112	105/132	175/197	216/218	213/213	susceptible[1]
BO04624039	260/274	180/190	112/112	105/132	175/197	216/218	209/213	resistant[1]
BO86615123	236/274	190/200	112/112	105/140	175/225	216/234	209/209	

[1]Babini, Phytosanitary Service, CRPV

Supplementary Table 3. Genotypes and phenotypes reported for the 119 accessions plus 14 well-known varieties.

Accession code	PGS1.10	PGS1.21	ZP002	PGS1.24	AMMPG016	AMMPG021	SSRLG_11m52	Phenotype
SPRINGBLUSH*	240/240	184/220	107/107	101/101	221/221	226/226	206/206	resistant[1]
PRIMARINA	240/240	218/220	107/107	101/101	221/221	226/226	206/213	
JLBUDD	240/240	218/220	107/107	101/101	189/221	216/226	206/209	
LE2927	232/240	178/220	107/107	101/103	193/221	208/226	199/206	resistant[1]
LE3662	240/248	178/220	107/107	101/103	203/221	222/226	206/213	
LE3205	232/240	178/220	107/107	101/103	0/0	208/224	197/199	resistant[1]
LE2904*	240/240	174/220	107/112	101/103	211/221	216/226	199/206	
MEDIABELL	240/240	200/220	107/112	101/103	191/221	218/226	199/206	resistant[1]
GILGAT	240/240	180/220	107/112	101/132	197/221	218/226	206/213	resistant[1]
MOGADOR	240/240	190/220	107/112	101/132	175/221	216/226	206/209	resistant[1]
LE3182	240/240	190/220	107/112	101/132	203/221	220/226	197/213	
BERGEVAL	240/240	174/220	107/112	101/103	0/0	214/226	206/209	
BERGAROUGE	240/240	174/220	107/112	101/103	0/0	216/226	199/206	
LE3225	240/240	190/220	107/112	101/132	197/203	0/0	197/213	
ZEBRA	240/260	180/220	107/112	101/132	197/221	218/226	206/213	
PETRA*	240/242	190/220	107/112	101/132	189/221	226/240	206/209	resistant[1]
GOLDRICH*	240/260	180/220	107/112	101/132	197/221	218/226	206/213	resistant[1]
SEO*	240/260	174/220	107/112	101/103	197/221	224/226	197/206	resistant[2]
LITO*	240/260	174/220	107/112	101/103	211/221	226/228	199/206	resistant[1]
FARMINGDALE*	240/260	174/220	107/112	101/103	211/221	226/228	206/213	resistant[4]
MURCIANA	236/240	200/220	107/112	101/140	221/225	226/234	206/209	
MEDAGA	240/260	174/220	107/112	101/103	211/221	226/228	206/213	
LUNA	240/260	180/220	107/112	101/132	197/221	216/226	206/213	susceptible[1]
FLAVORCOT*	240/260	174/220	107/112	101/103	197/221	224/226	206/209	resistant[1]
GG9871	240/266	194/220	107/112	101/140	197/221	218/226	199/206	
ROSA	240/260	190/220	107/112	101/132	175/221	206/226	206/209	

TSUNAMI	240/260	174/220	107/112	101/103	197/221	224/226	197/206	resistant[1]
FLOPRIA	240/274	190/220	107/112	101/105	175/221	216/226	206/209	resistant[1]
RUBISTA*	240/260	174/220	107/112	101/103	211/221	226/228	206/209	resistant[1]
FARLIS	240/274	190/220	107/112	101/105	175/221	216/226	206/209	resistant[1]
LADYCOT	240/274	190/220	107/112	101/105	175/221	216/226	199/206	resistant[1]
PRICIA	240/260	174/220	107/112	101/103	211/221	226/228	0/0	resistant[1]
LILLYCOT	240/274	190/220	107/112	101/105	175/221	222/234	206/209	resistant[1]
ANEGAT	236/240	200/220	107/112	101/140	0/0	226/234	206/209	
BIGRED	240/260	180/220	107/112	101/132	197/207	218/226	206/213	susceptible[1]
SUNNYCOT	240/260	180/220	107/112	101/127	197/197	218/226	206/213	
MIRLOBLANCO*	240/260	174/220	107/112	101/103	175/191	218/220	199/199	resistant[1]
MIRLOROJO	240/260	174/220	107/112	101/103	191/225	218/234	199/201	
CONGAT	240/274	190/220	107/112	101/105	175/211	216/216	209/213	
GK988	260/260	180/220	107/112	101/132	197/221	226/228	206/213	
ROJOPASION	236/260	200/220	107/112	101/140	191/225	218/234	199/209	
WONDERCOT	240/240	200/220	107/112	0/0	191/221	218/226	199/206	resistant[1]
SELENE	240/260	180/220	107/112	0/0	175/221	216/226	206/209	
MOIXENT	236/240	200/220	107/112	0/0	221/225	226/234	206/209	
HARLAYNE*	240/260	0/0	107/112	101/103	0/0	226/228	206/213	resistant[2]
LADYROSE	236/240	200/218	107/112	101/140	221/221	226/234	209/213	
AURORA*	240/260	174/174	107/112	101/103	211/221	226/228	206/213	resistant[1]
SHERPA	240/260	180/190	107/112	101/132	0/0	216/226	0/0	
HARVAL	260/260	196/218	107/112	101/103	189/197	216/218	206/209	resistant[1]
MORMONSSLDG	260/260	174/218	107/112	101/103	203/211	220/228	199/213	
CLUTHAGOLD	260/260	180/218	107/112	101/132	189/197	216/218	211/213	
BORA*	240/260	174/174	107/112	0/0	221/221	226/226	206/213	resistant[1]
HARCOT*	260/260	174/188	107/112	103/109	197/197	0/0	206/213	tolerant[3]
TONDINADITOS SIGNANO	236/274	190/190	107/112	105/105	175/207	0/0	199/209	
FARBALY	274/274	188/190	107/112	105/109	175/197	216/224	206/209	susceptible[1]
FARIUS	274/274	188/190	107/112	105/109	175/197	216/224	206/209	
FARDAO	274/274	188/190	107/112	105/109	175/197	216/224	206/209	
HAROSTAR	260/260	174/188	107/112	103/109	197/211	224/228	206/213	
NINFA	0/0	174/220	112/112	0/0	0/0	0/0	0/0	susceptible[1]
DULCINEA	236/242	190/200	112/112	101/140	175/175	206/216	201/209	
PRIMARIS	240/260	174/200	112/112	0/0	191/211	218/228	199/213	
HANITA	0/0	198/210	112/112	0/0	175/177	0/0	0/0	
MIRABOLANO	250/258	172/196	112/112	0/0	173/181	218/226	0/0	
MILORD	260/274	180/190	112/112	0/0	175/197	216/218	209/213	
FAVORIT	240/240	174/190	112/112	103/132	0/0	216/234	199/213	
CEGLEDIORIAS	240/240	174/190	112/112	103/132	0/0	216/234	199/213	
ESTRELLA	240/240	190/200	112/112	103/132	175/191	0/0	199/213	
SZEGEDIMAMMUT	240/240	174/190	112/112	103/132	175/175	216/234	199/213	
ULEANOS	240/240	190/190	112/112	132/132	175/175	206/206	199/209	
SILVANE	240/240	174/174	112/112	103/103	197/211	216/216	199/199	
PRIMAYA	240/240	190/200	112/112	103/132	175/191	216/218	199/209	
TONI	240/240	190/200	112/112	103/132	175/191	206/218	199/209	
BANEASA2211	240/260	174/174	112/112	103/103	0/0	216/228	199/213	
MAGYARKAISZI	240/260	174/190	112/112	103/132	0/0	220/228	213/213	
ORANGERUBIS	240/260	174/200	112/112	103/103	0/0	218/228	199/213	susceptible[1]
MARAVILLA	240/260	180/200	112/112	103/132	175/191	0/0	199/201	

FARALIA	240/260	180/200	112/112	103/132	191/197	218/218	199/209	susceptible[1]
PISANA	240/260	174/190	112/112	103/132	203/211	220/228	199/213	resistant[1]
ROYALROUSSILON	240/260	174/190	112/112	103/105	175/175	216/228	209/213	
SWEETRED	240/260	180/200	112/112	103/132	191/197	218/218	199/213	
MAIHUANG	240/266	192/198	112/112	105/129	175/175	218/224	204/213	
OLIMP	240/260	174/174	112/112	103/103	211/211	216/228	199/213	
TRZIIBUCUREST I	240/260	174/190	112/112	103/105	211/211	206/228	199/213	
VENUS	236/240	174/190	112/112	103/105	211/211	206/216	199/199	
CEGLEDIBIBOR	240/260	174/190	112/112	103/132	203/211	220/228	213/213	susceptible[1]
SULMONA	236/240	174/174	112/112	103/103	211/211	216/216	199/199	
NJA23	240/260	174/190	112/112	103/132	197/211	224/228	197/213	
OUARDI	240/260	190/190	112/112	132/132	175/175	206/234	199/209	
DORADA	236/240	174/200	112/112	103/140	211/225	216/234	199/201	
MAGIC COT	240/260	174/200	112/112	103/103	191/197	218/224	199/209	
SUBLIME	240/260	180/200	112/112	103/132	191/197	218/218	199/213	
HG9869	240/260	174/190	112/112	103/132	175/197	206/224	209/213	
MIRLONARANJA	240/260	174/200	112/112	103/103	191/211	218/228	199/213	
POPPY	240/260	180/190	112/112	132/132	175/197	216/218	209/213	
COLORADO	240/260	180/190	112/112	132/132	175/197	216/218	209/213	susceptible[1]
AUTUMNROYAL	0/0	174/190	112/112	103/105	165/175	216/228	209/213	
BOREALE	236/236	180/200	112/112	132/140	0/0	0/0	209/213	
OPAL	260/260	174/180	112/112	103/132	0/0	0/0	213/213	
FARFIA	260/260	174/180	112/112	103/132	0/0	216/228	209/213	
TONDADICOSTIGLIOLE	260/274	174/190	112/112	103/105	0/0	216/228	209/213	
JERSEYCOT	232/260	174/202	112/112	103/103	0/0	224/228	213/213	
NJA38	232/260	174/200	112/112	103/105	197/211	0/0	206/213	
SELEZIONESABBATANI	236/260	180/200	112/112	132/140	197/225	0/0	209/213	
MONO	260/260	180/180	112/112	132/132	197/197	0/0	213/213	
ARDORE	260/260	196/196	112/112	103/103	197/197	218/218	206/206	
FIAMMA	260/260	196/196	112/112	103/103	197/197	218/218	206/206	
FARCLO	260/274	174/190	112/112	103/105	175/211	216/228	206/209	susceptible[1]
KYOTO	260/274	180/190	112/112	105/132	197/197	218/218	213/213	susceptible[1]
PORTICI	260/274	174/190	112/112	103/105	175/211	216/230	209/213	susceptible[1]
REALEGRANDI	236/262	190/200	112/112	105/105	197/225	228/234	209/213	
REALEBALDASSARRI	236/262	190/200	112/112	105/140	197/225	228/234	209/213	
BEBECO	260/262	174/174	112/112	103/103	175/211	220/218	199/213	susceptible[1]
REALECASETTAINFERNO	236/262	190/200	112/112	105/105	197/225	228/234	209/213	
SELEZIONECASSANI	236/260	180/200	112/112	132/140	197/225	218/234	209/213	
RIVAL	260/260	174/174	112/112	103/103	197/211	224/228	209/213	
GUZAPRIKOZU	260/274	196/196	112/112	132/132	195/211	222/230	209/213	
PELESEDIGIOVANNIELLO	242/274	190/190	112/112	105/132	175/189	216/216	209/209	
SANANDREA	236/262	190/200	112/112	105/105	197/197	228/234	209/213	
PELECHIELLA	260/274	180/190	112/112	105/132	175/197	218/220	204/213	
AMABILEVECCHIONI	260/262	180/190	112/112	105/132	197/197	218/228	209/213	
SARRITZU1	236/236	198/200	112/112	130/140	203/225	220/234	209/209	
HELLIN1182	260/260	174/180	112/112	103/132	197/211	218/228	204/213	

DCD11-47-2	260/260	174/174	112/112	103/103	197/211	224/228	213/213	
NJA1	264/264	174/180	112/112	103/103	197/197	216/222	197/213	
TIRYNTHOS	236/260	174/190	112/112	103/105	207/211	206/228	197/199	
SPRINGGIANT	264/264	174/174	112/112	103/103	197/211	224/228	197/213	
YAMAGATA	236/236	180/200	112/112	105/132	175/205	212/230	204/204	
PIEVE	236/260	174/200	112/112	103/140	197/225	224/234	213/213	susceptible[1]
SANCASTRESE	260/260	174/180	112/112	103/132	197/197	218/228	213/213	
REALEIMOLA	236/262	190/200	112/112	105/140	197/225	228/234	209/213	
BELLAIMOLA	236/260	180/200	112/112	132/140	197/225	218/234	209/213	susceptible[1]
M1020040	260/274	180/190	112/112	105/132	175/197	216/218	209/213	
RUBY	260/274	174/190	112/112	103/105	175/211	216/228	209/213	

[1]Babini, Phytosanitary Service, CRPV; [2]Martínez-Gómez et al., 2003; [3]Stylianidis et al., 2005; [4]Trandafirescu et al., 2013; *: accessions used as control.

5.6 References

Atanasoff, D., 1933. Plum pox. A new virus disease. Annuaire de l'Universit´e de Sofia. 11, 49–69.

Badenes, M. L., Llácer, G., 2006. Breeding for resistance: breeding for Plum pox virus resistant apricots (Prunus armeniaca L.) in Spain. EPPO Bull. 36, 323–326.

Bassi, D., 2006. Breeding for resistance: breeding for resistance to Plum pox virus in Italy. EPPO Bull. 36, 327–329.

Bassi, D., Audergon, J. M., 2006. Apricot breeding, update and perspectives. Acta Hort. 701, 279–294.

Baumgartnerova, H., Slovakova, L., Petrusova, N., 1998. Relationship between concentration of plum pox virus and content of pigments in virus-infected symptomatic apricot leaves. Acta Virol. 42, 216–18.

Cambra, M., Capote, N., Myrta, A., Llacer, G., 2006. Plum pox virus and the estimated costs associated with Sharka disease. EPPO Bull. 36, 202–204.

Capote, N., Cambra, M., Llacer, G., Petter, F., Platts, L. G., Roy, A. S., Smith, I. M., 2006. Current status of Plum pox virus and Sharka disease worldwide. EPPO Bull. 36, 205–218.

Decroocq, S, Chague, A., Lambert, P., 2014. Selecting with markers linked to the PPVres major QTL is not sufficient to predict resistance to Plum Pox Virus (PPV) in apricot. Tree Genetics & Genomes. 10, 1161-1170.

Dondini, L., Lain, O., Geuna, F., Banfi, R., Gaiotti, F., Tartarini, S., Bassi, D., Testolin, R., 2007. Development of a new SSR based linkage map in apricot and analysis of synteny with existing Prunus maps. Tree Genet Genomes. 3, 239–249.

Dondini, L., Lain, O., Vendramin, V., 2011. Identification of QTL for resistance to plum pox virus strains M and D in 'Lito' and 'Harcot' apricot cultivars. Molecular Breeding. 27, 289-299.

Gildow, F., Damsteegt, V., Stone, A., Schneider, W., Luster, D., Levy, L., 2004. Plum pox in North America: identification of aphid vectors and a potential role for fruit in virus spread. Phytopathology. 94, 868–74.

Hammer, Ø., Harper, D. A. T., Ryan, P. D., 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica. 4, 9 pp.

Karayiannis, I., 2006. Breeding for resistance: conventional breeding for Plum pox virus resistant apricots (Prunus armeniaca L.) in Greece. EPPO Bull. 36, 319–322.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. *28*, 1647-1649.

Labonne, G., Lauriaut, F., Yvon, M., Quiot, J. B., 1994. Diss´emination du plum pox potyvirus par les pucerons: analyse des vecteurs potentiels du virus dans un verger d'abricotiers. EPPO Bull. 24, 681–90.

Lalli, D., A., Abbott, A. G., Zhebentyayeva, T., N., Badenes, M. L., Damsteegt, V., Polak, J., Krska, B, Salava, J., 2008. A genetic linkage map for an apricot (Prunus armeniaca L.) BC1 population mapping Plum pox virus resistance. Tree Genet Genomes. 4, 481–493.

Lambert, P., Dicenta, F., Rubio, M., Audergon, J. M., 2007. QTL analysis of resistance to sharka disease in the apricot (Prunus armeniaca L.) 'Polonais' × 'Stark Early Orange' F1 progeny. Tree Genetics & Genomes. 6, 291-304.

Lommel, S. A., McCain, A, H., Morris, T. J., 1982. Evaluation of indirect linked immunosorbent assay for the detection of plant viruses. Phytopathology. 72, 1018–1022.

Marandel, G., Salava, J., Abbott, A., Candresse, T., Decroocq, V., 2009. Quantitative trait loci meta-analysis of Plum pox virus resistance in apricot (Prunus armeniaca L.): new insights on the organization and the identification of genomic resistance factors. Mol. Plant Pathol. 10, 347–360.

Mariette, S., Wong Jun Tai, F., Roch, G., Barre, A., Chague, A., Decroocq, S., Groppi, A., Laizet, Y., Lambert, P., Tricon, D., Nikolski, M., Audergon, J. M., Abbott, A. G., Decroocq, V., 2015. Genome wide association links specific genes to resistance to Plum Pox Virus in apricot (Prunus armeniaca). The New Phytologist. 209, 447-886.

Martinez-Gomez, P., Dicenta, F., 2000. Evaluation of resistance of apricot cultivars to a Spanish isolate of plum pox potyvirus (PPV). Plant Breed. 119, 179–181.

Martínez-Gómez, P., Dicenta, F., Audergon, J. M., 2000. Behaviour of apricot (Prunus armeniaca L.) cultivars in the presence of sharka (Plum pox potyvirus): a review. Agronomie. 20, 407–422.

Martìnez-Gòmez, P., Rubio, M., Dicenta, F., 2003. Evaluation of resistance to Plum Pox Virus of North American and European apricot cultivar. Hort. Science. 38. 568-569.

Nagy, S., Poczai, P., Cernák, I., Gorji, A. M., Hegedűs, G., Taller, J., 2012. PICcalc: an online program to calculate polymorphic information content for molecular genetic studies. Biochemical Genetics. 50, 670-672.

Németh, 1986. Virus, Mycoplasma and Rickettsia Diseases of Fruit Trees. Budapest: Akadémiai Kiadó. p. 463.

Pilarova, P., Marandel, G., Decroocq, V., Salava, J., Krska, B., Abbott, A., G., 2010. Quantitative trait analysis of resistance to plum pox virus in the apricot F1 progeny 'Harlayne' x 'Vestar. Tree Genet. Genomes. 6, 467–475.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., Sham, P. C., 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics. 81, 559–575.

Rubio, M., Martinez-Gomez, P., Dicenta, F., Audergon, J. M., 2004. Testing of genetic control hypothesis for Plum pox resistance in apricot. Acta Hort. 663, 265–267.

Scholthof, K. B. G., Adkins, S., Czosnek,H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C., Foster, G., 2011. Top 10 plant viruses in molecular plant pathology. Mol. Plant Pathol. 12, 938–954.

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nature America Inc. 18, 1-2.

Soriano, J., Domingo, M., Zuriaga, E., 2012. Identification of simple sequence repeat markers tightly linked to plum pox virus resistance in apricot. Molecular Breeding. 30, 1017-1026.

Soriano, J. M., Vera-Ruiz, E. M., Vilanova, S., Martınez-Calvo, J., Llacer, G., Badenes, M. L., Romero, C., 2008. Identification and mapping of a locus conferring Plum pox virus resistance in two apricot improved linkage maps. Tree Genet. Genomes. 4, 391–402.

Stylianidis, C. D., Sotiropoulos, E. T., Syrgiannidis, G., Therios, N. I., Mainou, A., Karayiannis, I., A., Isaakidis, 2005. Effect of Sharka (Plum pox virus) disease on the nutrient status of the apricot (Prunus armeniaca L.) cvs ‘Harcot’ and ‘Bebecou’ and expression of symptoms of the physiological disorder Tip Burn on fruits of cv ‘Bebecou’. European Journal of Horticultural Science. 70, 121-124.

Sutic, D., 1971. Etat des recherches sur le virus de la sharka. Annales de Phytopathologie Supplement. 161–170.

Trandafirescu, M., Dumitru, L. M., Trandafirescu, I., 2013. Evaluating the resistance to the Plum Pox Virus of some apricot tree cultivars and hybrids in south-eastern Romania. Proceedings of the Latvian Academy of Sciences. 67, 203-206.

Vera Ruiz, E., Soriano, J. M., Romero, C., 2011. Narrowing down the apricot Plum pox virus resistance locus and comparative analysis with the peach genome syntenic region. Molecular Plant Pathology. 12, 535-547.

Wetzel, T., Candresse, T., Ravelonandro, M., Dunez, J., 1991. A polymerase chain reaction assay adapted to plum pox potyvirus detection. J. Virol. Methods. 33, 355–365.

Zhebentyayeva T. N., Reighard, G. L., Lalli, D. A., 2008. Origin of resistance to plum pox virus in Apricot: what new AFLP and targeted SSR data analyses tell. Tree Genetics and Genomes 4, 403-417.

Zuriaga, E., Soriano, J. M., Zhebentyayeva, T., 2013. Genomic analysis reveals MATH gene(s) as candidate(s) for Plum pox virus (PPV) resistance in apricot (Prunus armeniaca L.). Molecular Plant Pathology. 14, 663-677.

6 SHORT COMMUNICATION: DEVELOPMENT OF A HIGH-RESOLUTION MELTING APPROACH FOR RELIABLE AND COST-EFFECTIVE GENOTYPING OF *PPVRES* LOCUS IN APRICOT (*P. ARMENIACA*)

6.1 Introduction

The MAS approach represents one of the most promising strategies to boost the development of novel apricot (*Prunus armeniaca* L.*)* breeding lines resistant to Sharka disease, caused by the Plum Pox Virus (PPV) potyvirus. Currently, most of apricot European cultivars are susceptible to PPV infection (Dosba et al., 1991), whose symptoms and deterioration at fruit level make them unsuitable for consumption or industrial processing (Kegler et al., 1998). A set of short simple repeats (SSR) and/or single sequence length polymorphism (SSLP) markers associated to PPV-resistance trait have been identified on chromosome 1 in several linkage and/or association-based QTL mapping studies, allowing their use for a quick screening of promising breeding selections (Lambert et al., 2007; Soriano et al., 2008; Lalli et al., 2008; Marandel et al.; 2009, Pilarova et al., 2010; Dondini et al., 2011; Vera Ruiz et al., 2011; Soriano et al., 2012; Zuriaga et al., 2013). One of these markers (ZP002) identified a 5 base pair deletion allele on the main candidate gene (ppb022195m) belonging to the MATH (TRAF-like) family proteins (Zuriaga et al., 2013) and possibly conferring resistance to Sharka. The mutation results in the expression of a truncated protein that may interfere with the pathway of virus translocation throughout the plant, although the exact mechanism of action is still unknown. Other two SSR markers, PGS1.21 and PGS1.24 were identified in the flanking regions of ppb022195m gene, covering together with ZP002, the main locus for Sharka resistance, known as *PPVres* (Soriano et al., 2012).

Although this marker set is necessary but not sufficient to unambiguously assign PPV resistant cultivars (Decroocq et al., 2014), it still represents a highly valuable tool for breeders and nurserimen to perform a preliminary screening on seedlings or novel released cultivars. Moreover, considering the highly expensive evaluation procedures for a reliable evaluation of Sharka disease susceptibility (at least 4 - 5 years of trials in adequate facilities for the containment of a quarantine pathogen), an effective reduction of individuals to be tested is highly desirable. From a practical point of view, the application of SSR markers for PPV-MAS shows its own limitation when the number of samples becomes larger, requiring the use of expensive genotyping analytical platforms.

A set of CAPS (Cleaved Amplified Polymorphic Sequences) markers was also developed for a gel-based (agarose) genotyping (Decroocq et al., 2014; Mariette et al., 2015). However, CAPS markers are laborious and time-consuming, due to the use of restriction enzymes and gel electrophoresis screening, and, thus, difficult to adopt for mass selection programs.

High-resolution melting (HRM) analysis is a closed-tube method for the analysis of genetic variation within PCR amplicons (Reed and Wittwer, 2003). This method relies on the use of a saturating intercalating dye to monitor fluorescence variation during the thermal denaturation of DNA (i.e. dsDNA to ssDNA transition). Genetic variants with differences in base composition are discriminated by their characteristic melting profiles and/or melting temperature (T_m) (Liew et al., 2004). HRM analysis has proven to be a highly sensitive, rapid and cost-effective method, particularly for SNPs genotyping, and successfully applied in many plant species (Simko, 2016).

In this research, we developed a quick and cost-effective protocol for genotyping PPV-resistance in apricot by developing and validating HRM assays on target markers at the *PPVRes* locus.

6.2 Results and Discussion

As previously described, genotyping for PPV resistance in apricot is commonly performed using two SSRs and one SSLP marker, that jointly cover the *PPVRes* locus: PGS1.21, PGS1.24 and ZP002 (Soriano et al., 2012; Decroocq et al., 2014). Except for ZP002, with a unique allelic variant (a 5-bp deletion), PGS1.21 and PGS1.24 markers are characterized by one resistant allele and several SSR variants. Despite many works successfully adapted SSR marker to HRM technology (Ganopoulos et. al., 2011; Mader et. al., 2008), such allelic variability makes them unsuitable for a simple and reliable use in HRM. Therefore, two novel assays (PGS1.21_SNP and PGS1.24_SNP) were designed by selecting two A/G polymorphisms in linkage with resistance (allelic variants with A nucleotide) (S. Decroocq, personal communication) and located in the proximity of PGS1.21 and PGS1.24 markers (Figure 1). In order to increase the sensitivity of HRM assay and with the purpose of using standard DNA melting instruments (not equipped with high-resolution systems), amplicon size of ZP002 marker was reduced from 111-116 to 86-91 bp, designing novel primers around the 5 bp deletion of the ppb022195m gene (ZP002_DEL) (Figure 1). In addition, an alternative assay was also designed for the ZP002 (ZP002_SNP) based on the A/T polymorphism located in first intron of ppb022195m gene and in linkage with the deletion (Figure 1).

A

Name	Type	Size	Position (bp)	Forward primer (5' - 3')	Reverse primer (5' - 3')	Variant
PGS1.21[1]	SSR	172-220	8527745	CCCTGGTGTTCTGCTCTCTC	CATCCACAAATGGGAAGCAT	
PGS1.24[1]	SSR	101-141	8668808	GTAAATGAGTGCCTGCGTGT	TGCGAGAGTTGTGATTGATG	
ZP002[2]	SSLP	107-112	8607078	TGTATCCTCCAGCTTCAAAGTC	AACATTTTCTGATTCAATGCCA	
PGS1.21_SNP	SNP	118	8491690	ACCCGGTGAAAGAAAAGTGA	TGGATCGCTTCTACATGTCAAG	A/G
PGS1.24_SNP	SNP	70	8774594	ACTAATAAACAGGACCTGCAATAT	CTATTCAATTTCGGTTTATGC	A/G
ZP002_SNP	SNP	64	8606821	GCCAGTTTCCTGTAGCAAAACCAC	ATTCTTATTTCCAAGCTGCATTA	A/T
ZP002_DEL	SSLP	86-91	8606972	CAGCTTCAAAGTCTTCCGATTCA	ATGCCAACTCATTACACGTTCAA	GTTTG

B

PGS1.21_SNP
ACCCGGTGAAAGAAAAGTGAAATATCTAGCAAAACCAGCCAACTAAAATATCACAGTAAGAAGGCAATCTGACCGAT[G/A]AGCATGCGA
ATTCCAATTCTTGACATGTAGAAGCGATCCA

PGS1.24_SNP
ACTAATAAACAGGACCTGCAATAT[G/A]CTTAAGTGTTACCTCTTAGGAAGTGCATAAACCGAAATTGAATAG

ZP002_SNP
GGATAR AAAGCCAGTTTCCTGTAGCAAAACCAC[A/T]AATTGAGAGTACTGTAATGCAS CTTGGAAATAAGAATAAAAT

ZP002_DEL
CAGCTTCAAAGTCTTCCGATTCATATCTCTCCAGTGAAT[GTTTG]GACATCAATGAAAACGACTGTATTTTGAACGTGTAATGAGTTGGCAT

C

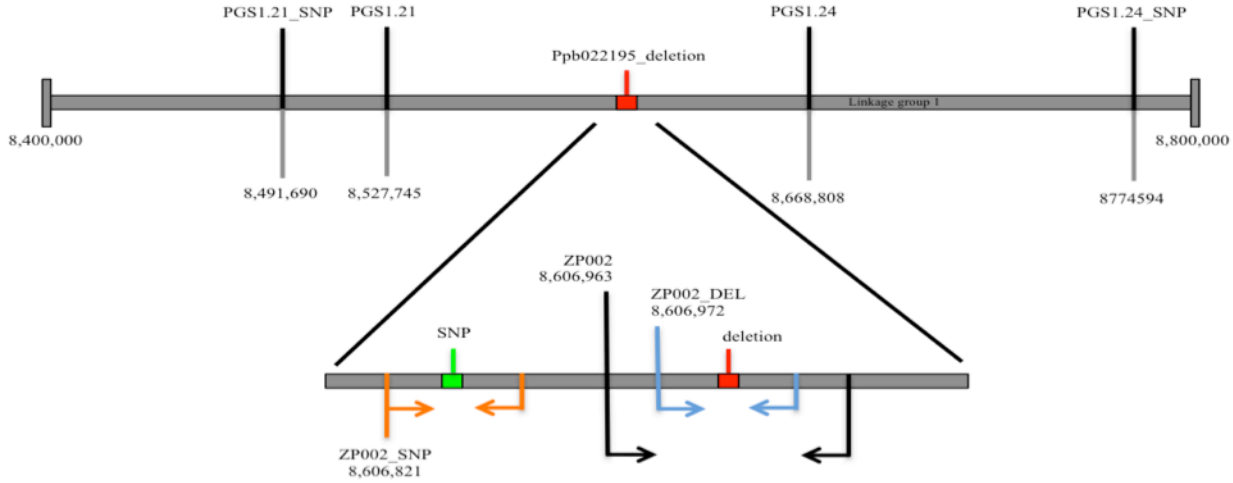


Figure 1. Schematic representation of the PPVres region and flanking markers.

(**A**) Oligonucleotides used in the HRM analysis, [1]Soriano et al., 2012, [2]Decroocq et al., 2014. (**B**) Amplicons sequences of the assessed markers (primers in grey and SNP/SSLP in black). (**C**) Schematic representation of the PPVRes locus on linkage group 1 of apricot.

SNP-flanking regions were retrieved from a BAC library assembly of 'Lito' cultivar. HRM primers were designed by using Primer3 software (http://primer3.ut.ee/) and pre-validated *in silico* by using uMELT-HETS and uMELT-BATCH melting prediction tools (Dwight et al., 2011), available at uMELT website (www.dna.utah.edu/umelt/umelt.html) (Supplemental Figure 1). MELTSIM thermodynamic set and default values of monovalent cations and magnesium concentrations were used for prediction.

Young leaves of 51 accessions and breeding selections already phenotyped for PPV resistance (list is shown in Supplemental Table 1) were sampled from the CRPV (Centro Ricerche

Produzioni Vegetali) experimental station (Cesena, Italy) using a leaf-disc sampler. About 50 mg of fresh tissue for each plant were collected in a 1 ml tubes, lyophilized and then stored at room temperature. In order to reduce DNA extraction time and cost, a simplified CTAB protocol (Doyle and Doyle, 1990) was implemented. Lyophilized samples were ground by a TissueLyser (Qiagen, Germany) through the addition of carborundum. A volume of 200 µl of CTAB extraction buffer and 10 µl RNase A (Sigma, Cat# R4875) were added to the powdered tissue, followed by incubation at 65° C for 10 minutes. The composition of buffer was the following: 4% CTAB, 4 M NaCl, 0.5 M EDTA (pH 8.0), 1M Tris-Cl (pH 8.0), 3% PVP (MW 40 kDa) and (just prior to use) 2% β-mercaptoethanol. After spinning for 1 min at 5000 g to separate cell debris and carborundum, the mixture was transferred to a clean 96-well plate, adding 0.5 volume of chlorophorm:isoamyl alcohol 24:1 (v/v). Then, the solution was mixed and centrifuged for 5 min at 11,000 g. The upper aqueous phase was collected and transferred in a clean 96-well plate, adding 0.6 volume of isopropanol and 0.1 volume of 3M sodium acetate. After precipitating DNA for 20 min at 4° C, surnatant was discarded and the pellet washed 2 times with 80% ice-cold ethanol, dried under a fume-hood and dissolved in 30 µl DEPC water. Samples were directly quantified using the QuBit dsDNA HS Assay kit and the QuBit 3.0 fluorometer following the manufacturer's instructions (ThermoFisher, Italy). DNA concentration varied from a minimum of 10 to a maximum of 30 ng/µl (data not shown).

The four HRM assays were tested on three DNA templates from cultivars 'Lito', 'Portici' and the selection 'BO03615049', previously scored by using SSR/SSLP markers and carrying different combinations of the resistant alleles (homozygote, heterozygote or absent) (Table 1). HRM analyses were carried out using an Eco Real-Time PCR System (Illumina, USA). PCR mixture consists of 5 ng/µl template DNA, 0.25 mM forward and reverse primers, 1X EVAGREEN Precision Melt Supermix (Bio-Rad, USA) and H_2O DEPC in a final volume of 12 µl. The reactions were carried out with the following program: 2 min at 94° C, 35 cycles of 30 s at 94° C, 30 s annealing at 58° C and 30 s at 72° C, followed by a final melting step over a 65-95°C gradient with 0.1° C/s ramp rate. Data were analyzed using EcoStudy software (Illumina, USA). Melting data were normalized according to the operator's manual and visualized as fluorescence/temperature melting curves profile and derivative plot (-dF/dT).

Table 1. Comparison of the resistance alleles calling with the three SSR markers by the HRM assays, on a subset of fourteen accessions/breeding selections of apricot

Accession	PGS1.21	PGS1.21_SNP	ZP002	ZP002_DEL	PGS1.24	PGS1.24_SNP
AURORA	174 – 174	T / T	**107** – 112	**-** / GTTTG	**101** – 103	**A** / T
BELLA DI IMOLA	180 – 200	T / T	112 – 112	GTTTG / GTTTG	132 – 140	T / T

BO03615049	**220** – **220**	**A** / **A**	**107** – **107**	**-** / **-**	**101** – **101**	**A** / **A**
BO96621002	**220** – **220**	**A** / **A**	**107** – **107**	**-** / **-**	**101** – **101**	**A** / **A**
FARBALY	188 – 190	T / T	**107** – 112	**-** / GTTTG	105 – 109	T / T
FLAVORCOT	174 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 103	**A** / T
GOLDRICH	180 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 132	**A** / T
KYOTO	180 – 190	T / T	112 – 112	GTTTG / GTTTG	105 – 132	T / T
LADYCOT	190 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 105	**A** / T
LITO	174 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 103	**A** / T
MOGADOR	190 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 132	**A** / T
NINFA	174 – **220**	**A** / T	112 – 112	GTTTG / GTTTG	103 – 103	T / T
PETRA	190 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 132	**A** / T
PORTICI	174 – 190	T / T	112 – 112	GTTTG / GTTTG	103 – 105	T / T
SEO	174 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 103	**A** / T
TSUNAMI	174 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 103	**A** / T
WONDERCOT	200 – **220**	**A** / T	**107** – 112	**-** / GTTTG	**101** – 103	**A** / T

Resistance alleles are shown in bold.

Considering the ZP002_DEL marker, the resistance allele (with a 5 bp deletion) is clearly distinguishable from the susceptible one in both homozygote and heterozygote combinations (Figure 2A, B). Derivative melting plot shows two clear peaks with a T_m of about 76.8° and 77.8°C for the resistant and susceptible alleles, respectively and a different melting shape for the heterozygote, characterized by two broadened peaks with a T_m of about 73.1°C and 77.4°C, respectively (Figure 2A). The marker PGS1.24_SNP shows a T_m values of about 73.7° and 74.6°C for the homozygous resistant and susceptible alleles, consistent with the expected combination (AA and GG, respectively) (Figure 2C, D). The heterozygous individuals can be easily identified by the different melting curve shape, determined by A/G heteroduplexes formation. A similar result was also obtained for PGS1.21_SNP, (homozygotes for A and G alleles show a T_m difference of about 0.4°C) although the broadened peak of heterozygote A/G genotype can be more easily recognized through the melting curve profile rather than the derivative plot (Figure 2E, F). The ZP002_SNP marker does not provide the expected results, since the resistant genotype (homozygote for A allele) show a higher T_m compared to the susceptible one (homozygote for T allele), although the assay is apparently able to discriminate both (Figure 2G, H). A BLAST search against the SRA dataset of 66 apricot accessions (Bioproject PRJNA292050) confirmed the presence of a G/C mutation within the reverse primer not in linkage with the A/T, resulting in different allelic combinations and consequently, melting profiles. Moreover, the selected SNP is located in a highly polymorphic region, which does not allow excluding other unlinked polymorphisms from the amplicon (data not shown). For such reasons, the assay was excluded from the analysis.

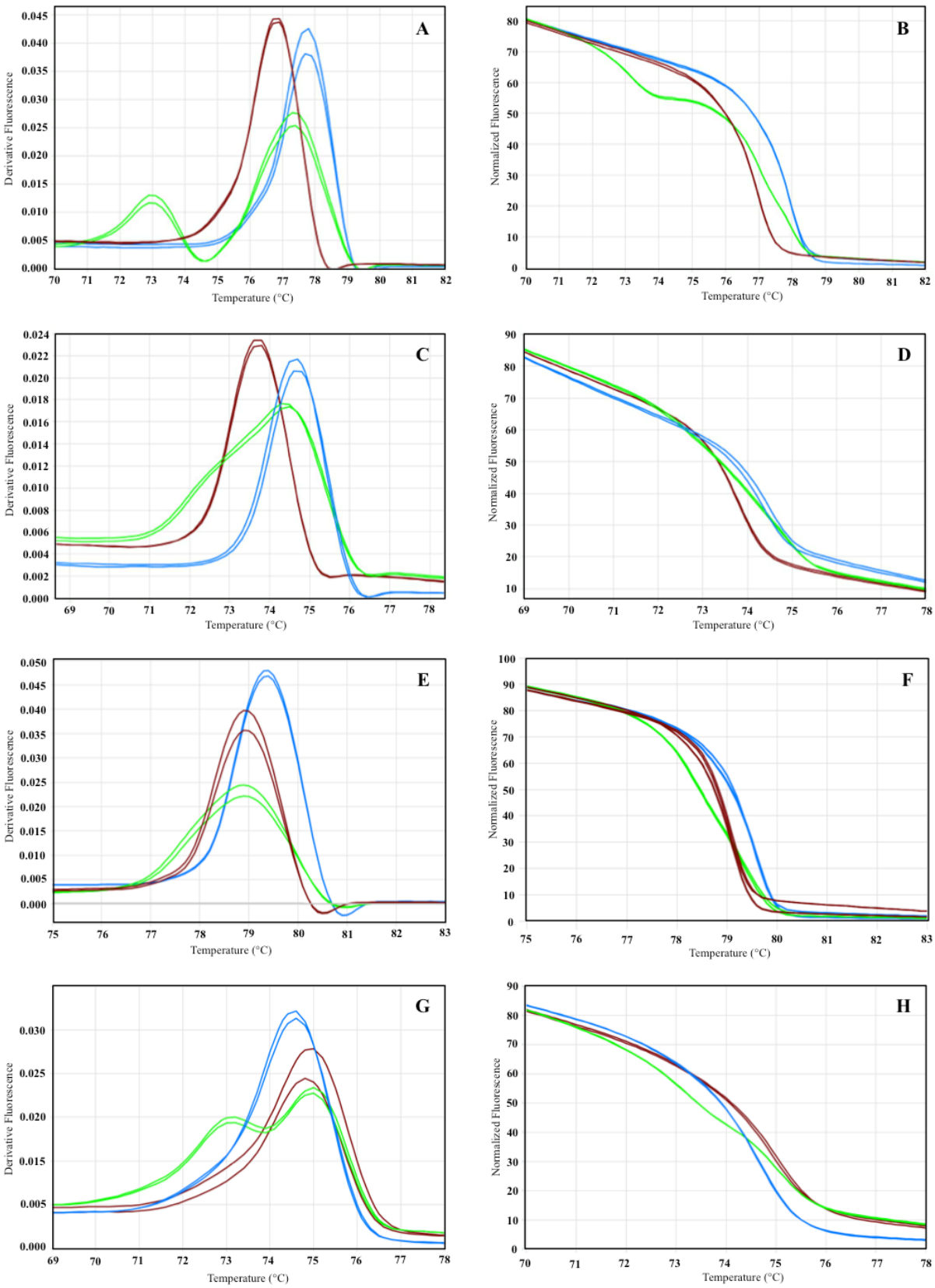


Figure 2. Derivative fluorescence and normalized fluorescence profiles of the HRM analysis. The following markers are shown: ZP002_DEL (**A**, **B**); PGS1.24_SNP (**C**, **D**); PGS1.21_SNP (**E**, **F**); ZP002_SNP (**G**, **H**) for two biological replicates of samples 'BO03615049' (brown), 'Lito' (green) and 'Portici' (blue).

Only for marker ZP002_DEL, melting curve analyses were also repeated on an ABI7300 Real Time PCR instrument (Thermo Scientific, USA) not equipped with high-resolution melting systems (i.e. ramp rate higher than 1° C/s), using the same PCR reaction protocol above described except for PCR mix: 1X SYBR Premix Ex Taq (Takara, Shiga, Japan) and final volume of 20 µl. Interestingly, the assay is able to discriminate genotypes, although with a slightly lower resolution (Supplementary Figure 2).

After the HRM assay, the expected amplicon sizes were confirmed by 3% agarose gel electrophoresis (Supplementary Figure 3). A set of 48 samples was analyzed to assess the reproducibility of HRM (Figure 3), confirming their high sensitivity also in multi-sample conditions, allowing to unambiguously genotype all individuals, particularly using normalized fluorescence melting profile (Figure 3 B, D, F) (Supplementary Table 1). Data were also cross-validated by genotyping a subset of fourteen accessions/breeding selections with the three SSR markers PGS1.21, ZP002 and PGS1.24. Forward primers were labelled on the 5' position with the PET, FAM and NED fluorophores, respectively. The PCR reaction consisted of 5 ng/µl template DNA, 0.10 mM dNTPs, 0.25 mM of each primer, 1X PCR buffer, 2.5 mM $MgCl_2$ and 0.025 U/µl Taq DNA polymerase (Eurx LTD, Poland) in a final volume of 20 µl. Reactions were carried out with the following program: 5 min at 94°C, 2 cycles of 30 s at 94°C, 1.30 min at 55°C and 1 min at 72°C, 35 cycles of 30 s at 94°C, 30 s at 55°C and 30 s at 72°C with a final extension of 10 min at 72°C. The size of PCR products was scored through electrophoresis on a capillary sequencer AB3730 (Thermo Scientific, USA). As reported in Table 1, SSR data support the resistance alleles assignment by HRM assays.

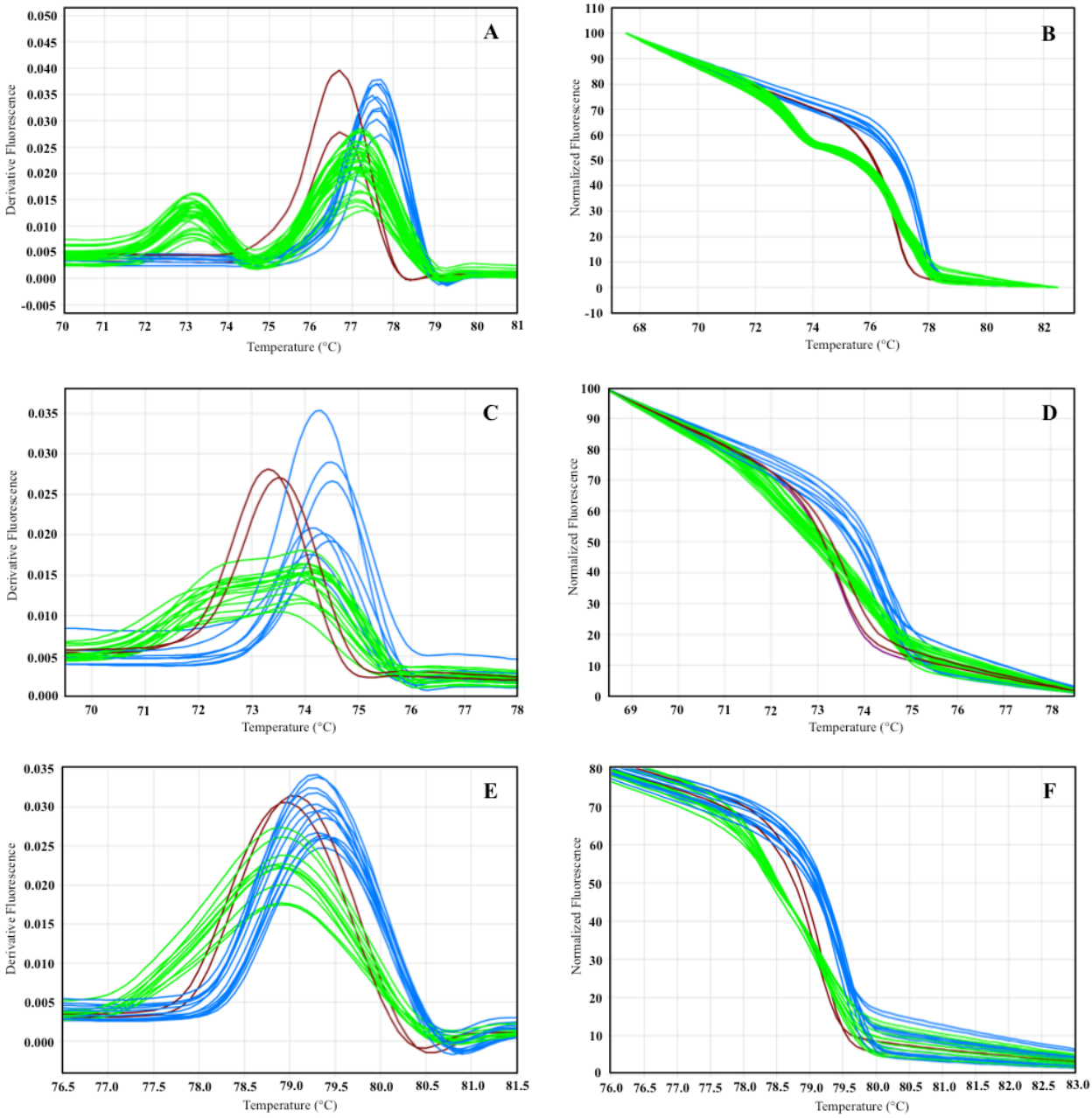


Figure 3. Derivative fluorescence and normalized fluorescence profiles of the HRM analysis. for the markers ZP002_DEL (**A**, **B**), PGS1.24_SNP (**C**, **D**) and PGS1.21_SNP (**E**, **F**) on a randomised set of 48 samples of apricot accessions/breeding selections carry on different combinations of the resistant alleles: homozygote (brown), heterozygote (green) or absent (blue).

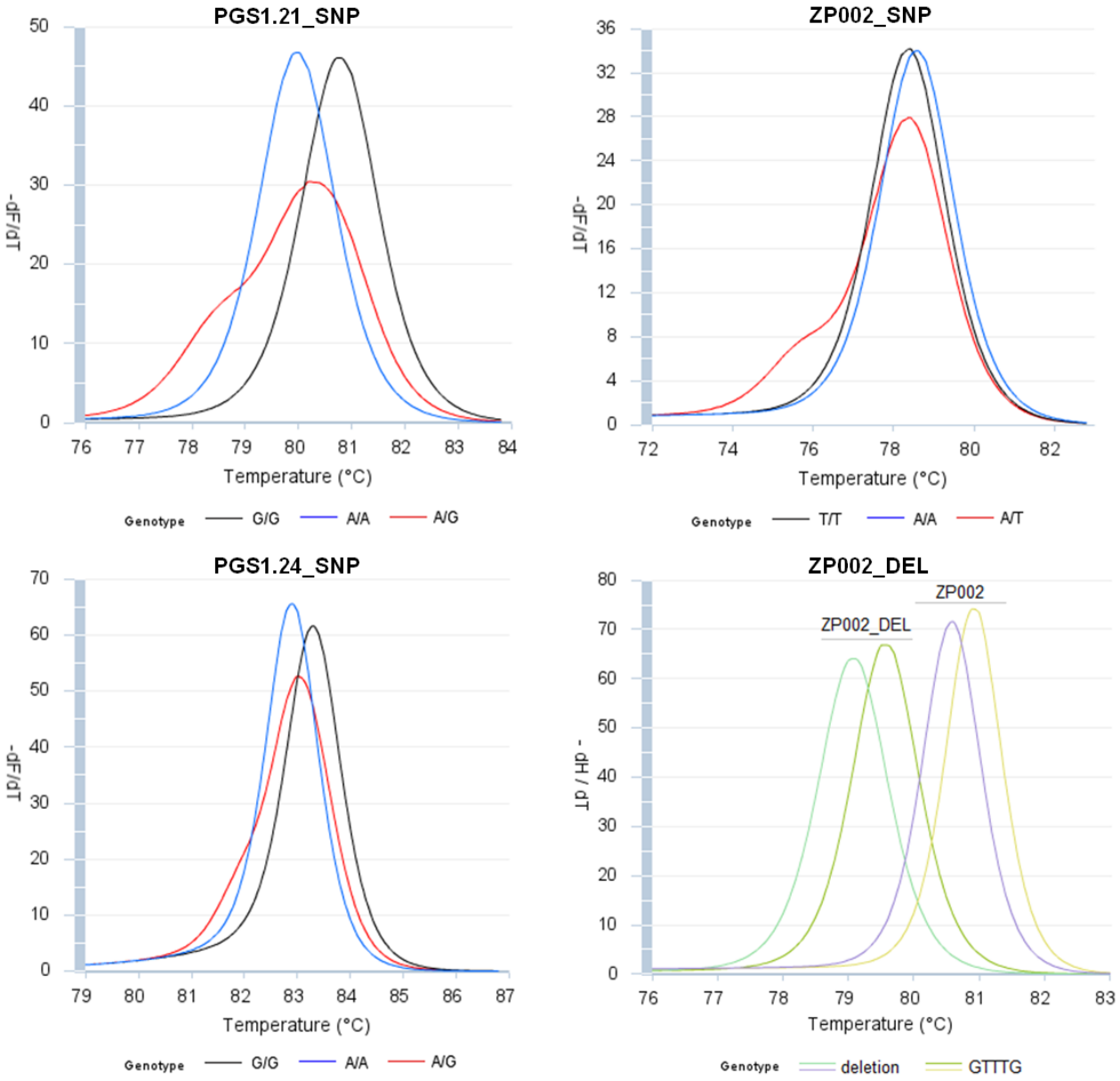
A cost-survey was also conducted to compare the available molecular tools for MAS of PPV-resistance (Supplemental Table 2). SSR markers resulted the most expensive, especially when an automatic sequencer is not available. In this case the cost of MAS reach the value of 7.56 euros per sample. CAPS technology reduces the overall cost to 6.39 euros per sample, avoiding fragment analysis. However, the needs for restriction enzymes led to saving just 15% compared with SSR.

KASP technology (Chunlin et. al., 2014), provide high-throughput genotyping system using SNP markers, with the possibility to include DNA extraction. KASP services at LGC Genomics (Hoddesdon, UK) costs 4.8 euros per sample, configuring itself as one of the best options particularly when a molecular biology facility is not available. The HRM genotyping method proposed here is the less expensive (3.96 euros per sample), and also the fastest because of the only two-step requirements to reach the final data, although it requires a Real-Time system. By replacing the commercial kit for DNA extraction with our extraction method the price drops significantly for all the assays excluding LGC Genomics services.

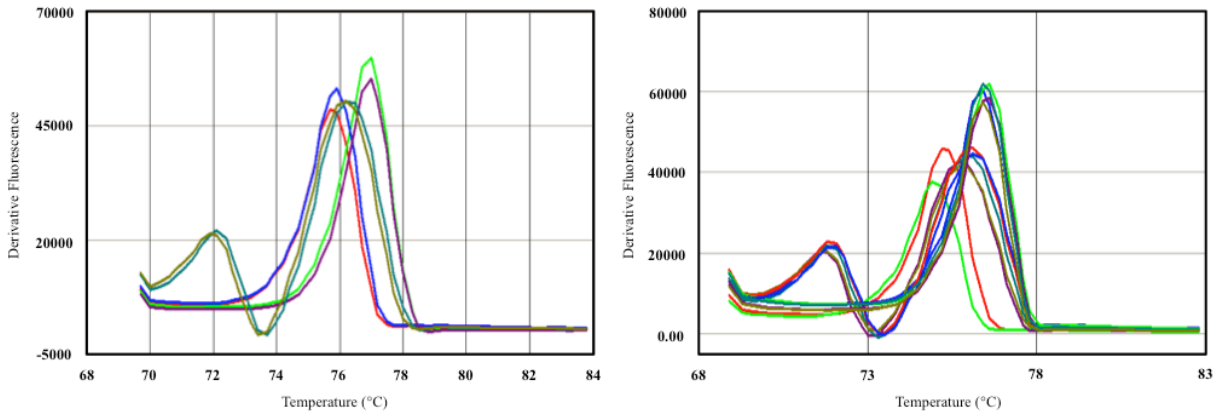
6.3 Conclusions

In conclusion, we developed a reliable and user-friendly HRM-based method for the genotyping of PPV resistance locus in apricot. This approach is able to consistently reduce costs and time of laboratory analyses, thus giving an important contribution to the adoption of the MAS strategy in conventional breeding programs, boosting the rapid development of novel Sharka resistant lines.

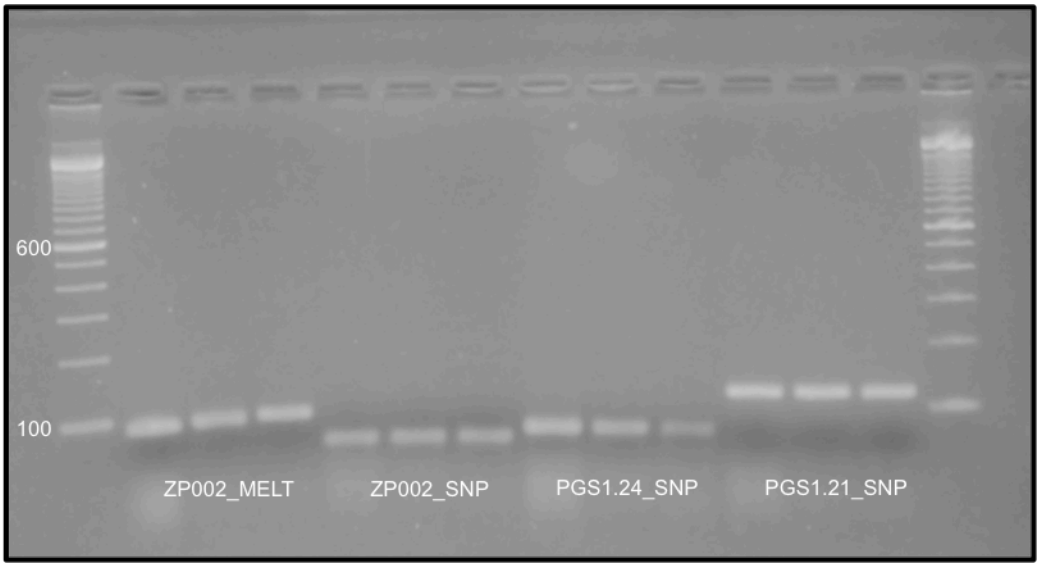
6.4 Supplementary Materials



Supplementary Figure 1. *In silico* pre-validation of the experimental HRM assays. The output profiles are shown of uMELT-HETS and uMELT-BATCH melting prediction programs for each of the four novel SNP/SSLP markers.



Supplementary Figure 2. Derivative fluorescence melting curve analyses for marker ZP002_DEL on a 7300 Real Time PCR instrument. Two biological replicates are depicted of accessions 'BO03615049' (blu, red), 'Lito'(green, dark green) and 'Portici' (light green, violet) (left panel) and of a randomised set of 48 samples of apricot accessions and breeding selections carried out on different combinations of the resistant alleles (right panel). Only a subset of the samples are depicted on the right panel in order to allow a better visualization of the melting curves.



Supplementary Figure 3. Size confirmation of the expected amplicons on a 3% agarose gel electrophoresis for the four assessed markers. A 100 bp Ladder is used for reference.

Supplementary Table 1. Phenotype and HRM assays results for 51 accessions and breeding selections.

Accession	Genotype			Phenotype
	PGS1.21_SNP	ZP002_DEL	PGS1.24_SNP	
BO03615049	A / A	- / -	A / A	resistant [1]
LITO	A / T	- / GTTTG	A / T	resistant [1]
PORTICI	T / T	GTTTG / GTTTG	T / T	susceptible [1]
AURORA	T / T	GTTTG / GTTTG	A / T	resistant [1]
BEBECO	T / T	GTTTG / GTTTG	T / T	resistant [1]
BELLA DI IMOLA	T / T	GTTTG / GTTTG	T / T	susceptible [1]
BIGRED	A / T	- / GTTTG	A / T	susceptible [1]
BO03615034	A / A	- / -	A / A	resistant [1]
BO03615053	A / T	- / GTTTG	A / T	resistant [1]
BO04624031	A / T	- / GTTTG	A / T	resistant [1]
BO04624042	T / T	GTTTG / GTTTG	T / T	susceptible [1]
BO04624043	T / T	GTTTG / GTTTG	T / T	susceptible [1]
BO05636034	A / T	- / GTTTG	A / T	resistant [1]
BO06609048	T / T	- / GTTTG	A / T	resistant [1]
BO06609053	T / T	GTTTG / GTTTG	T / T	resistant [1]
BO06609055	T / T	- / GTTTG	A / T	resistant [1]
BO96621002	A / A	- / -	A / A	resistant [1]
BO96621021	A / T	- / GTTTG	A / T	susceptible [1]
BO96621030	A / T	- / GTTTG	A / T	resistant [1]
BORA	T / T	- / GTTTG	A / T	resistant [1]
CEGLEDIBIBOR	T / T	GTTTG / GTTTG	T / T	susceptible [1]
COLORADO	T / T	GTTTG / GTTTG	T / T	susceptible [1]
FARALIA	T / T	GTTTG / GTTTG	T / T	susceptible [1]
FARBALY	T / T	- / GTTTG	T / T	susceptible [1]
FARCLO	T / T	GTTTG / GTTTG	T / T	susceptible [1]
FARLIS	A / T	- / GTTTG	A / T	resistant [1]
FARMINGDALE	A / T	- / GTTTG	A / T	resistant [4]
FLAVORCOT	A / T	- / GTTTG	A / T	resistant [1]
FLOPRIA	A / T	- / GTTTG	A / T	resistant [1]
GILGAT	A / T	- / GTTTG	A / T	resistant [1]
GOLDRICH	A / T	- / GTTTG	A / T	resistant [1]
HARCOT	T / T	- / GTTTG	T / T	tolerant [3]
HARLAYNE	T / T	- / GTTTG	A / T	resistant [2]
HARVAL	T / T	- / GTTTG	A / T	resistant [1]
KYOTO	T / T	GTTTG / GTTTG	T / T	resistant [1]
LADYCOT	A / T	- / GTTTG	A / T	resistant [1]
LILLYCOT	A / T	- / GTTTG	A / T	resistant [1]
LUNA	A / T	- / GTTTG	A / T	susceptible [1]
MEDIABELL	A / T	- / GTTTG	A / T	resistant [1]
MIRLOBLANCO	A / T	- / GTTTG	A / T	resistant [1]
MOGADOR	A / T	- / GTTTG	A / T	resistant [1]
NINFA	A / T	GTTTG / GTTTG	T / T	susceptible [1]
ORANGERUBIS	T / T	GTTTG / GTTTG	T / T	susceptible [1]
PETRA	A / T	- / GTTTG	A / T	resistant [1]
PIEVE	T / T	GTTTG / GTTTG	T / T	resistant [1]
PISANA	T / T	GTTTG / GTTTG	T / T	resistant [1]
PRICIA	A / T	- / GTTTG	A / T	resistant [1]
RUBISTA	A / T	- / GTTTG	A / T	resistant [1]
SEO	A / T	- / GTTTG	A / T	resistant [2]
TSUNAMI	A / T	- / GTTTG	A / T	resistant [1]
WONDERCOT	A / T	- / GTTTG	A / T	resistant [1]

[1]Babini, Phytosanitary Service, CRPV, [2]Martínez-Gómez et al., 2003; [3]Stylianidis et al., 2005; [4]Trandafirescu et al., 2013.

Supplementary Table 2. Survey of costs of the common methods for high-throughput screening of the *PPVres* locus in apricot for Sharka resistance.

Step	SSR (€)	CAPS (€)	LGC (€)	HRM (€)
DNA extraction	2.488	2.488		2.488
Primers	0.069	0.014		0.014
PCR	0.775	0.775		1.455
Agarose gel	0.259[1]	0.519[2]		
DNA ladder	0.268	0.268		
Enzyme		2.327		
Run service	3.700			
All inclusive service			4.800	
Total	7.560	6.392	4.800	3.957

Product	Price (€)	Reactions (n°)
Anza™ 51 BspT #IIVGN0516 (AFLII, PGS1.21) (Thermo Fisher Scientific)	57	300
Anza™ 25 PaeI #IVGN0254 (SPHI, PGS1.24) (Thermo Fisher Scientific)	57	40
Anza™ 38 ScaI #IVGN0386 (SCAI, ZP002) (Thermo Fisher Scientific)	57	80
Capillary electrophoresis (external service)	3.7	1
Certified™ Molecular Biology Agarose #1613102 (Biorad)	692	8000[1]
Certified™ Molecular Biology Agarose #1613102 (Biorad)	692	4000[2]
Fluorerscent Primer (20bp, FAM) (Biorad)	62.2	3000
KASP Assay "all inclusive" (LGC Genomics)	4.8	1
Master Mix For PCR #1665009 (Biorad)	62	240
NucleoSpin® Plant II #740770.250 (Macherey-Nagel)	622	250
Precision Melt Supermix #1725110 (Biorad)	194	400
Standard Primer (20bp) (Biorad)	7.2	3000
100 bp PCR Molecular Ruler #1708206 (Biorad)	143	1600

Prices and estimated number of reactions for the products considered in the cost-survey (lower panel). [1]2% agarose gel. [2]4% agarose gel.

The article has been submitted to Molecular Breeding and is currently under review

6.5 References

Barkley, A. N., Roose, M. L., Krueger, R. R., Federici, C. T., 2006. Assessing genetic diversity and population structure in a *Citrus* germplasm collection utilizing simple sequence repeat markers (SSRs). Theor Appl Genet. 112, 1519–1531.

Chunlin, H., Holme, J., Anthony, J., 2014. SNP Genotyping: The KASP Assay. Crop Breeding: Methods in Molecular Biology. 1145, 75-86.

Collard B. C., Mackill D. J. *2008.* Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil. Trans. R. Soc. *B. 363, 557–572.*

Decroocq, S., Chague, A., Lambert, P., Roch, G., Audergon, J. M., Geuna, F., Chiozzotto, R., Bassi, D., Dondini, L., Tartarini, S., 2014. Selecting with markers linked to the PPVres major QTL is not sufficient to predict resistance to Plum Pox Virus (PPV) in apricot. Tree Genetics & Genomes. 10, 1161–1170.

Dondini, L., Lain, O., Vendramin, V., Rizzo, M., Vivoli, D., Adami, M., Guidarelli, M., Gaiotti, F., Palmisano, F., Bazzoni, A., 2011. Identification of QTL for resistance to plum pox virus strains M and D in Lito and Harcot apricot cultivars. Mol. Breed. 27, 289–299.

Dosba, F., Denise, F., Maison, P., Massonié, G., Audergon, J. M., 1991. Plum pox virus resistance of Accepted Article apricot. Acta Horticulturae. 293, 569-579.

Doyle, J. J., Doyle, J. L., 1990. Isolation of plant DNA from fresh tissue. Focus. 12, 13-15.

Dwight, Z., Palais, R., Wittwer, C. T., 2011. uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. Bioinformatics. 27, 1019 – 1020.

Ganopoulos, I., Argiriou, A., Tsaftaris, A., 2011. Microsatellite high resolution melting (SSR-HRM) analysis for authenticity testing of protected designation of origin (PDO) sweet cherry products. Food Control. 22, 532–541.

Kegler, H., Hartmann, W., 1998. Present status of controlling conventional strains of Plum pox virus. In: Hadidi, A., Khetarpal, R. K., Koganezawa, H. eds. Plant Virus Disease Control. 616–28. APS Press, St. Paul, MN.

Kumar, P., Gupta, V. K., Misra, A. K., Modi, D. R., Pandey, B. K., 2009. Molecular markers in plant biotechnology. Plant Omics Journal. 2, 141–162.

Lalli, D. A., Abbott, A. G., Zhebentyayeva, T. N., Badenes, M. L., Damsteegt, V., Polak, J., Krska, B., Salava, J., 2008. A genetic linkage map for an apricot (*Prunus armeniaca* L.) BC1 population mapping plum pox virus resistance. Tree Genetics Genomes. 4, 481–493.

Lambert, P., Dicenta, F., Rubio, M., Audergon, J. M., 2007. QTL analysis of resistance to sharka disease in the apricot (*Prunus armeniaca* L.) 'Polonais' x 'Stark Early Orange' F1 progeny. Tree Genetics Genomes 3, 299–309.

Liew, M., Pryor, R., Palais, R., Meadows, C., Erali, M., Lyon, E., Wittwer C., 2004. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clin. Chem. 50, 1156–1164.

Mader, E., Lukas, B., Novak, J., 2008. A strategy to setup codominant microsatellite analysis for high-resolution-melting-curve-analysis (HRM). BMC Genetics. 9, 69.

Marandel, G., Salava, J., Abbott, A., Candresse, T., Decroocq, V., 2009. Quantitative trait loci meta-analysis of Plum pox virus resistance in apricot (*Prunus armeniaca* L.): new insights on the organization and the identification of genomic resistance factors. Mol. Plant Pathol. 10, 347–360.

Mariette, S., Wong Jun Tai, F., Roch, G., Barre, A., Chague, A., Decroocq, S., 2015. Genome-wide association links candidate genes to resistance to Plum pox virus in apricot (*Prunus armeniaca*). New Phytol. 209, 773–784.

Martínez-Gómez, P., Rubio, M., Dicenta, F., 2003. Evaluation of resistance to Plum pox virus of North American and European apricot cultivars. Europ. J. Hort. Sci. 38, 568-569.

Moose, S. P., Mumm, R. H., 2008. Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiology. 147, 969-977.

Pilarova, P., Marandel, G., Decroocq, V., Salava, J., Krka, B., Abbott, A. G., 2010. Quantitative trait analysis of resistance to plum pox virus in the apricot F1 progeny 'Harlayne' x 'Vestar'. Tree Genetics Genomes 6, 467–475.

Sánchez-Pérez, R., Ballester, J., Dicenta, F., Arús, P., Martínez-Gómez, P., 2006. Comparison of SSR polymorphisms using automated capillary sequencers, and polyacrylamide and agarose gel electrophoresis: Implications for the assessment of genetic diversity and relatedness in almond. Scientia Horticulturae. 108, 310-316.

Simko, I., 2016. High-resolution DNA melting analysis in plant research. Trends Plant Sci. 21, 528-537.

Soriano, J. M., Vera-Ruiz, E. M., Vilanova, S., Martinez-Calvo, J., Llacer, G., Badenes, M. L., Romero, C., 2008. Identification and mapping of a locus conferring plum pox virus resistance in two apricot improved linkage maps. Tree Genetics Genomes. 4, 391–40.

Soriano, J. M., Domingo, M. L., Zuriaga, E., Romero, C., Zhebentyayeva, T., Abbott, A. G., Badenes, M. L., 2012. Identification of simple sequence repeat markers tightly linked to plum pox virus resistance in apricot. Mol. Breeding. 30, 1017–1026.

Stylianidis, D. C., Sotiropoulos, T. E., Syrgiannidis, G., Therios, I. N., Mainou, A., Karagianni, I., Isaakidis, A., 2005. Effect of the Sharka (Plum pox virus) Disease on the Nutrient Status of the Apricot (Prunus armeniaca L.) Cultivars 'Harcot' and 'Bebeco' and Expression of Symptoms of the Physiological Disorder Tip Burn on Fruits of the cv. 'Bebeco'. Europ. J. Hort. Sci. 70, 121–124.

Trandafirescu, M., Dumitru, L. M., Trandafirescu, I., 2013. Evaluating the resistance to the plum plox virus of some apricot tree cultivars and hybrids in south-eastern Romania. Proceedings of the Latvian Academy of Sciences. 67, 203-206

Vera Ruiz, E. M., Soriano, J. M., Romero, C., Zhebentyayeva, T., Terol, J., Zuriaga, E., Llacer, G., Abbott, A. G., Badenes, M. L., 2011. Narrowing down the apricot Plum pox virus resistance locus and comparative analysis with the peach genome syntenic region. Mol. Plant Pathol. 12, 535–547.

Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G., Pryor, R. J., 2003. High-resolution genotyping by amplicon melting analysis using LC Green. Clinical Chemistry 49, 853–860.

Wu, S. B., Wirthensohn, M. G., Hunt, P., Gibson, J. P., Sedgley, M., 2008. High resolution melting analysis of almond SNPs derived from ESTs. Theoretical and Applied Genetics. 118, 1–14.

Zuriaga, E., Soriano, J. M., Zhebentyayeva, T., Romero, C., Dardick, C., Canizares, J. G., Badenes, M. L., 2013. Genomic analysis reveals MATH gene(s) as candidate(s) for Plum pox virus (PPV) resistance in apricot (*Prunus armeniaca* L.). Molecular Plant Pathology. 14, 663–677.

7 CONCLUSIONS

The tremendous advances in DNA genotyping have been made possible by the development of new high-troughput technologies that are revolutionizing horticulture research and its applications to modern breeding. However the cost of technologies such as NGS or SNP-array chips not always make them the most rational approach, particularly when working on minor crop species or with large number of samples. Minor crops indeed do not enjoy the same money investment reserved to other major species and, on the others hand, working with numbers tipically required by in-field routine analyses necessitates a dramatic increase of the overall costs in terms of both time and resources.

In this thesis different kinds of molecular markers were applied in order to reach significative results in a cost-effective manner and with practical applications in the horticultural sector.

Two AFLP markers have been identified with the ability to discriminate Chinotto from 'Chinotto di Savona', an uninvestigated traditional *Citrus* species cultivated in the Liguria region (Italy) that is gaining increasing interest for the production of high-quality niche food and beverages. These markers may be used to preserve and increase the agro-economic heritage of the variety, defining its peculiarity and protecting its identity by introducing the certification of the nursery propagation material.

The development of resistant cultivars against Sharka disease is one of the main priorities in the frame of the breeding programmes for the safeguarding of apricot production. However the relevant costs , in terms of time and money, associated to phenotyping require the adoption of cost-effective molecular methods to support the common breeding practices. Until now, SSR markers represented a valuable choice for assisting the screening of resistant samples between populations, selections or accessions of apricot, allowing a reduction of resources spent into the application of phenotyping protocols. A further reduction of costs and times was proposed in this work by reducing the number of markers needed to obtain a reliable prediction of resistance and by replacing the SSR markers with new SNP markers to process using HRM technology.

We demonstrate here that the wise choice of the instruments to use, together with the original implementation of new techniques, could easily overcame the limitations connected with the study of minor, less studied crops or the practical applications of technologies, still getting reliable results.

8 BIBLIOGRAPHY

Akagi, H., Yokozeki, Y., Inagaki, A., Nakamura A., Fujimura, T., 1996. A co-dominant DNA marker closely linked to the rice nuclear restorer gene, Rf-1, identified with inter SSR fingerprinting. Genome. 39, 1205–1209.

Armour, J. A. L., Alegre, S. A., Miles, S., Williams, L. J., Badge, R. M., 1999. Minisatellites and mutation processes in tandemly repetitive DNA. In: Goldstein, D. B., Schlotterer, C., (ed) Microsatellites: evolution and applications. Oxford University Press, Oxford, 24–33.

Baird, V., Abbott, A., Ballard, R., Sosinski, B., Rajapakse, S., 1997. DNA Diagnostics in Horticulture, 111–130. In: Gresshoff, P., (Ed.), Current Topics in Plant Molecular Biology: Technology Transfer of Plant Biotechnology. CRC Press, Boca Raton.

Bartolozzi, F., Warburton, M. L., Arulsekar, S., Gradziel, T. M., 1998. Genetic characterization and relatedness among California almond cultivars and breeding lines detected by randomly amplified polymorphic DNA (RAPD) analysis. J. Am. Soc. Hort. Sci. 123, 381-87.

Batley, J., Barker, G., O'Sullivan, H., Edwards, K. J., Edwards, D., 2003a. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol. 132, 84–91.

Batley, J., Mogg, R., Edwards, D., O'Sullivan, H., Edwards, K. J., 2003b. A high-throughput SNP assay for genotyping SNPs in the flanking regions of Zea mays sequence tagged simple sequence repeats. Mol. Breed. 11, 111–120

Beckman, J., S., Soller, M., 1986. Restriction fragment length polymorphism and genetic improvement of agricultural species. Euphytica. 3, 111-124.

Blake, T. K., Kadyrzhanova, D., Shepherd, K. W., Islam, A. K. M. R., Langridge, P. L., McDonald, C. L., Erpelding, J., Larson, S., Blake, N. K., Talbert, L. E., 1996. STS-PCR markers appropriate for wheat-barley introgression. Theor. Appl. Genet. 82, 715-721.

Borevitz, J. O., Liang, D., Plouffe, D., Chang, H. S., Zhu, T., Weigel, D., Berry, C. C., Winzeler, E., Chory, J., 2003. Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res. 13, 513-523. 16.

Botstein, D., White, R. L., Skolnick, M., Davis, R. W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32, 314–331.

Bruford, M.W., Cheesman, D. J., Coote, T., Green, H. A. A., Haines, S. A., O'Ryan, C., William, T. R., 1996. Microsatellites and their application to conservation genetics. In: Smith, T. B,. Wayne, R. K., (ed) Molecular genetic approaches in conservation. Oxford University Press, New York, 278–297.

Chambers, G. K., MacAvoy, E. S., 2000. Microsatellites: consensus and controversy. Comparative Biochem. Physiol. (Part B). 126, 455–476.

Cordeiro, G. M., Casu, R., McIntyre, C. L., Manners, J. M., Henry, R. J., 2001. Microsatellite markers from sugarcane (Sacharum spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci. 160, 115–1123.

Corley Smith, G. E., Lim, C. J., Kalmar, G. B,, Brandhorst, B. P., 1997. Efficient detection of DNA polymorphisms by fluorescent RAPD analysis. Bio Techniques. 22, 690-692.

Cui, X. P., Xu, J., Asghar, R., Condamine, P., Svensson, J. T., Wanamaker, S., Stein, N., Roose, M., Close, T. J., 2005. Detecting single- feature polymorphisms using oligonucleotide arrays and robustified projection pursuit. Bioinfomatics. 21, 3852- 3858.

Das, S., Bhat, P. R., Sudhakar, C., Ehlers, J. D., Wanamaker, S., Roberts, P. A., Cui, X., Close, T. J., 2008. Detection and validation of single feature polymorphisms in cowpea (Vigna unguiculata L. Walp) using a soybean genome array. BMC Genomics. 9, 107.

Davis, G. L., Mcmullen, M.D., Baysdorfer, C., Musket, T., Grant, D., Staebell, M., Xu, G., Polacco, M., Koster, L., Melia-Hancock, S., Houchins, K., Chao, S., Choe, E. H. Jr., 1999. A maize map standard with sequenced core markers, grass genome reference points, and 932 expressed sequence tagged sites (ESTs) in a 1736 locus map. Genetics. 152, 1137–1172.

Decroocq, V., Fave, M. G., Hagen, L., Bordenave, L., Decroocq, S., 2003. Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor. Appl. Genet. 106, 912–922.

Devos, K. M., Gale, M. D., 1992. The use of random amplified polymorphic DNA markers in wheat. Theor. Appl. Genet. 84, 567-572.

Diwan, N., Cregan, P. B., 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor. Appl. Genet. 95, 723–733.

Drenkard, E., Richter, B. G., Rozen, S., Stutius, L. M., Angell, N. A., Mindrinos, M., Cho, R. J., Oefner, P. J., Davis, R. W., Ausubel, F. M., 2000. A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. Plant Physiol. 124, 1483–1492.

Eisen, J. A., 1999. Mechanistic basis for microsatellilte instability. In: Goldstein, D. B., Schlotterer, C., (ed) Microsatellites: evolution and applications. Oxford University Press, Oxford, 34–48.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., Mitchell, S. E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6, e19379.

Farooq, S., Azam, F., 2002. Molecular markers in plantbreeding: concepts and characterization. Pakistan journal of biologicalsciences. 5, 1135-1140.

Ganal, M., Altmann, T., Roder, M. S., 2009. SNP identification in crop plants. Current Opinion in Plant Biology. 12, 211–217.

Gao, L. F., Jing, L. R., Huo, N. X., Li, Y., Li, X. P., Zhou, R. H., Chang, X. P., Tang, J. F., Ma, Z. Y., Jia, J. Z., 2004. One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. Theor. Appl. Genet. 108, 1392–1400.

Garg, K.,, Green, P., Nickerson, D. A., 1999. Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. Genome. Res. 9, 1087–1092.

Geuna, F., Toschi, M., Bassi, D., 2003. The use of AFLP markers for cultivar identification in apricot. Plant Breed 122: 526–531.

Goldstein, D. B., Pollock, D. D., 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. J. Hered. 88, 335–342.

Grodzicker, T., Williams, J., Sharp, P., Sambrook, J., 1975. Physical mapping of temperature sensitive mutants of adenovirus. Cold Spring Harbor Symp. Quant. Biol. 39, 439-446.

Gupta, M., Chyi, Y. S., Romero-Severson, J., Owen, J., 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple sequence repeats. Theor. Appl. Genet. 89, 998–1006.

Gupta, P. K., Roy, J. K., Prasad, M., 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr. Sci. 80, 524–535.

Gupta, P., Varshney, R., Sharma P., Ramesh, B., 1999. Molecular markers and their applications in wheat breeding. Plant Breed. 118, 369–390.

Hammond, R. L., Saccheri, I. J., Ciofi, C., Coote, T., Funk, S. M., McMillan, W.O., Bayes, M.K., Taylor, E., Bruford, M. W., 1998. Isolation of microsatellite markers in animals. In: Karp, A., Issac, P. G., Ingram, D. S., (ed) Molecular tools for screening biodiversity. Chapman and Hall, London, 279– 287.

Hancock, J. M., 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Goldstein, D. B., Schlotterer, C., (ed) Microsatellites: evolution and applications. Oxford University Press, Oxford, 1–9.

Harushima, Y., Yano, M., Shomur,a A., Sato, M., Shimano, T., Kuboki, Y., Yamamoto, T., Yang Lin, S., Baltazar, A. A., Parco, A., Kajiya, H., Huang, N., Yamamoto, K., Nagamura, Y., Kurata, N., Kush, G. S., Sasai, T., 1998. A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics. 148, 479–494.

Hearne, C. M., Ghosh, S., Todd, J. A., 1992. Microsatellites for linkage analysis of genetic traits. Trends Genet. 8, 288–294.

Helentjaris, T., Slocum, M., Wright, S., Schaefer, A., Nienhuis, J., 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. 61, 650-658.

Henry, R., 1997. Molecular markers in plant improvement. In: Practical Applications of Plant Molecular Biology, 99–132. Chapman and Hall, London.

Hempel, K., Peakall, R., 2003. Cross-species amplification from crop soybean Glycine max provides informative microsatellite markers for the study of inbreeding wild relatives. Genome. 46, 382–393.

Hollingsworth, W. O., Christie, C. B., Nichols, M.A., Neilson, H. F., 1998. Detect ion of variation among and within asparagus hybrids using random amplified DNA (RAPD) markers. New Zealand J. Crop Hortic. Sci. 26, 1-9.

Holtke, H. J., Ankenbauer, W., Muhlegger, K., Rein, R., Sagner, G., Seibl, R., Walter, T. 1995. The digoxigenin (DIG) system for non radioactive labelling and detection of nucleic acids an overview. Cell. Mol. Biol. 41, 883-905.

Huff, D. R., Peakall, R., Smouse, P. E., 1993. RAPD variation within and among natural populations of outcrossing buffalo grass [(Buchloe dactyloides (Nutt.) Engelm]. Theor. Appl. Genet. 86, 927-934.

Jarne, P., Lagoda, P., 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11, 424–429.

Jarvis, P., Lister, C., Szabo, V., Dean, C., 1994. Integration of CAPS markers into the RFLP map generated using recombinant inbred lines of Arabidopsis thaliana. Plant Mol. Biol. 24, 685–687.

Jahufer, M., Barret, B., Griffiths, A., Woodfield, D., 2003. DNA fingerprinting and genetic relationships among white clover cultivars. In: Morton J., (Ed.), Proceedings of the New Zealand Grassland Association. Taieri Print Limited, Dunedin. 65, 163–169.

Jones, C. J., Edwards, K. J., Castaglione, S., Winfield, M. O., Sala, F., van de Wiel C, et al. 1997. Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories. Mol. Breed. 3, 381-390.

Jones, N., Ougham H., Thomas, H., 1997. Markers and mapping: We are all geneticists now. New Phytol. 137, 165–177.

Jongeneel, C. V., 2000. Searching the expressed sequence tag (EST) databases: Panning for genes. Briefings in Bioinformatics. 1, 76-92.

Joshi, S., Ranjekar, P., Gupta, V., 1999. Molecular markers in plant genome analysis. Curr Sci. 77, 230–240.

Joshi, S. P., Prabhakar, K., Ranjekar, P. K., Gupta, V. S., 2011. Molecular markers in plant genome analysis. http/www.ias.ac.in/currsci/jul25/articles15.htm. 1-19.

Kaiser, M. W., Lyamicheva, N., Ma, W., Miller, C., Neri, B., Fors, L., Lyamichev, V. I., 1999. A comparison of eubacterial and archaeal structure-specific 5-exonucleases. J. Biol. Chem. 274 , 21387–21394.

Kantety, R. V., Rota, M. L., Matthews, D. E., Sorrells, M. E., 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum, and wheat. Plant Mol. Biol. 48, 501–510.

Krist, M., Caldo, R., Casati, P., Tanimoto, G., Walbot, V., Wise, R. P., Buckler, E. S., 2006. Genetic diversity contribution to errors in short oligonucleotide microarray analysis. Plant Biotechnol. 4, 489-498.

Konieczny, A., Ausubel, F. M., 1993. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J. 4, 403–410.

Kumar, R., Qiu, J., Joshi, T., Valliyodan, B., Xu, D., Nguyen, H. T., 2007. Single feature polymorphism discovery in rice. PLoS ONE. 2, e284.

Landry, B. S., 1994. DNA mapping in plants. In: Glick, B. R., Thompson, J. E., (ed) Methods in plant molecular biology and biotechnology. CRC Press, Boca Raton-Ann Arbor-London-Tokyo, 269-285.

Levinson, G., Gutman G. A., 1987. Slipped strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol. 4, 203–221.

Lin, J.J., Kuo, J., Ma, J., Saunders, J. A., Beard, H. S., Macdonald, M. H., Kenworthy, W., Ude, G. N., Mathews, B. F., 1996. Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Mol. Biol. Rep. 14, 156-169.

Liu, C. J., Witcombe, J. R., Pittawy, T. S., Nash, M., Hash, C. T., Brusso, C. S., Gale, M. D., 1994. An RFLP-based genetic map of pearl millet (Pennissetum glaucum). Theoretical and applied genet. 8, 481-487.

Lu, Z. X., Reighard, G. L., Baird, W. V., Abbott, A. G., Rajapakse, S., 1996. Identification of peach rootstock cultivars by RAPD markers. Hort. Sci. 31, 127-129.

Lyamichev, V., Mast, A. L., Hall, J. G., Prudent, J. R., Kaiser, M.W., Takova, T., Kwiatkowski, R. W., Sander, T. J., de Arruda, M., Arco, D. A., Neri, B. P., Brow, M. A. D., 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. Nat. Biotechnol. 17, 292–296.

Maeda, M., Uryu, N., Murayama, N., Ishii, H., Ota, M., Tsuji, K., Inoko, H., 1990. A simple and rapid method for HLADP genotying by digestion of PCR-amplified DNA with allele specific restriction endonucleases. Hum. Immunol. 27, 111–121.

Mansfield, E. S., Worley, J. M., McKenzie, S. E., Surrey, S., Rappaport, E., Fortina, P., 1995. Nucleic acid detection using non radioactive labeling methods. Mol. Cell. Probes. 9, 145-156.

Mardis, E. R., 2008. Next-generation DNA sequencing methods. Annu. Rev. Genomics Hum.Genet. 9, 387–402.

McDermott, J. M., Brandle, U., Dutly, F., Haemmerli, U. A., Keller, S., Muller, K. E., Wolf, M.S., 1994. Genetic variation in powdery mildew of barley: Development of RAPD, SCAR and VNTR markers. Phytopathology. 84, 1316-1321.

McDonald, D. B., Potts, W.K., 1997. DNA microsatellites as genetic markers for several scales. In: Mindell, D. P, (ed.) Avian molecular evolution and systematics. Academic Press, San Diego, 29–49.

Meyer, W., Mitchell, T.G., Freedman, E. Z., Vilgays, R., 1993. Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of Cryptococcus neoformans. J. Clin. Microbiol. 31, 2274–2280.

Metzker, M. L., 2010. Sequencing technologies the next generation. Nat. Rev. Genet. 11, 31–46.

Michaels, S. D., Amasino, R. M., 1998. A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant J. 14, 381–385.

Milbourne, D., Meyer, R., Bradshaw, J. E., Baird, E., Bonar, N., Provan, J., Powell, W., Waugh, R., 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Mol. Breed. 3, 127-136.

Mueller, U. G., Lipari, S. E., Milgroom, M. G., 1996. Amplified fragment length polymorphism (AFLP) fingerprinting of symbiotic fungi cultured by the fungus-growing ant Cyphomyurmex minutus. Mol. Ecol. 5, 119-122.

Mullis, K., Fallona, S., Schraf, S., Saiki, R., Horn G., Erlich, H., 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symp. Quant Biol. 51, 263–273.

Nasu, S., Suzuki, J., Ohta, R., Hasegawa, K., Yui, R., Kitazawa, N., Monna, L., Minobe, Y., 2002. Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (Oryza satva, Oryza rufipogon) and establishment of SNP markers. DNA Res. 9, 163-171.

Neff, M. M., Neff, J. D., Chory, J., Pepper, A. E., 1998. dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in Arabidopsis thaliana genetics. Plant J. 14, 387–392.

Pammi, S., Schertz, K., Xu, G., Hart, G., Mullet, J. E., 1994. Random amplified polymorphic DNA markers in sorghum. Theor. Appl. Genet. 89, 80-88.

Paran, I., Michelmore, R., 1993. Development of reliable PCR based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet. 85, 985–993.

Pareek, C. S., Smoczynski, R., Tretyn, A., 2011. Sequencing technologies and genome sequencing. J. Appl. Genet. 52, 413–435.

Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L., Syvanen, A. Minisequencing C., 1997. A specific tool for DNA analysis and diagnostics on oligonucleotide arrays. Genome Res. 7, 606-614.

Paterson, A. H., 1996. Making genetic maps. In: Paterson, A. H., (Ed.), Genome Mapping in Plants, 23–39. Landes, R. G. Company, San Diego, California; Academic Press, Austin, Texas.

Pavy, N., Parsons, L. S., Paule, C., MacKay, J., Bousquet, J., 2006. Automated SNP detection from a large collection of white spruce expressed sequences: contributing factors and approaches for the categorization of SNPs. BMC Genomics. 7:174.

Pease, A. C., Solas, D., Sullivan, E. J., Cronin, M. T., Holmes, C. P., Fodor, S. P., 1994. Light generated oligonucleotide arrays for rapid DNA sequence analysis. Proc. Natl. Acad. Sci. USA 91, 5022-5026.

Perez de la Vega, M. P., 1993. Biochemical characterization of populations. In: Hayward, M. D., Bosemark, N. O., Romagosa, I., (ed) Plant breeding: Principles and prospects. Chapman and Hall, London, 182-201.

Pillay, M., Kenny, S. T., 1995. Anomalies in direct pairwise comparisons of RAPD fragments for genetic analysis. Biotechniques 19, 694-698.

Potter, R. H., Jones, M. G. K., 1991. Molecular analysis of genetic stability. In: Dodds, J. H., (ed) In vitro methods for conservation of plant genetic resources. Chapman and Hall, London, 71-91.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., Rafalski, A., 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2, 225- 238.

Prins, R., Groenewald, J. Z., Marais, G.F., Snap, J. W., Koebner, R. M. D., 2001. AFLP and STS tagging of Lr19, a gene conferring resistance to leaf rust in wheat. Theor. Appl. Genet. 103, 618–624.

Queller, D. C., Strassman, J. E., Hughes, C. R., 1993. Microsatellites and Kinship. Trends Ecol. Evol. 8, 285-288.

Rafalski, A., 2002. Applications of single nucleotide polymorphisms in crop genetics. Curr. Opin. Plant Biol. 5, 94–100.

Rafalski, J., Tingey, S., 1993. Genetic diagnostics in Plant Breed: RAPDs, microsatellites and machines. Trends Genet. 9, 275–280.

Rajapakse, S. L. E., Belthoff, G., He, A. E., Estager, R., Scorza, I., Verde, R. E., Ballard, W. V., Baird, A., Callahan, R., Monet, Abbott A. G., 1995. Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theor. Appl. Genet. 90, 503-10.

Ratnaparkhe, M. B., Tekeoglu, M., Muehlbauer, F. J., 1998. Inter simple sequence repeat (ISSR) polymorphisms are useful for finding markers associated with disease resistance gene clusters. Theor. Appl. Genet. 97, 515–519.

Reamon-Buttner, S. M., Jung, C., 2000. AFLP-derived STS markers for the identification of sex in Asparagus officinalis L. Theor. Appl. Genet. 100, 432–438.

Ribaut, J. M., Hoisington, D., 1998. Marker assisted selection: New tools and strategies. Trends Plant Sci. 3, 236–239.

Roy, A., Frascaria, N., Mackay, J., Bonsquet, J., 1992. Segregating random amplified polymorphic DNA (RAPDs) in Betula alleghaniensis. Theor. Appl. Genet. 85, 173-180.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., Nyren, P., 1996. Real-time DNA sequencing using detection of pyrophosphate release. Anal. Biochem. 242, 84–89.

Rostoks, N., Borevitz, J. O., Hedley, P. E., Russell, J., Mudies, S., Morris, J., Cardie, L., Marshall, D. F., Waugh, R., 2005. Single-feature polymorphism discovery in the barley transcriptome. Genome Biol. 6, R54.

Russell, J. R., Fuller, J. D., Macaulay, M., Hatz, B. G., Jahoor, A., Powell, W., Waugh, R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95, 714-722.

Sanger. F., Nicklen, S., Coulson, A. R., 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74, 5463-5467.

Sankar, A. A., Moore, G. A., 2001. Evaluation of inter simple sequence repeat analysis for mapping in Citrus and extension of genetic linkage map. Theor. Appl. Genet. 102, 206–214.

Schlotterer, C., 1998. Microsatellites. In: Hoelzel, A. R., (ed) Molecular genetic analysis of populations: A practical approach. IRL Press, Oxford, 237–261.

Schmid, K. J., Sorensen, T. R,, Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T., Weisshaar, B., 2003. Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in Arabidopsis thaliana. Genome Res. 13, 1250-1257.

Semagn, K., Bjornstad, A., Ndjiondjop, M. N., 2006. An overview of molecular markermethods for plants. Afr. J. Biotech. 5, 2540–2568.

Shan, X., Blake, T. K., Talbert, L.E., 1999. Conversion of AFLPs to sequence-tagged-site PCR markers in barley and wheat. Theor. Appl. Genet. 98, 1072-1078.

Shendure, J., Ji, H., 2008. Next-generation DNA sequencing. Nat. Biotechnol. 26, 1135–1145.

Singer, T., Fan, Y., Chang, H. S., Zhu, T., Hazen, S. P., Briggs, S. P., 2006. A high- resolution map of Arabidopsis recombinant inbred lines by whole-genome exon array hybridization. PLoS Genet., 2, e144.

Sobrino, B., Briona, M., Carracedoa, A., 2005. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci. International. 154, 181–194

Southern, E. M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503–517.

Staub, J. E., Serquen, F. C., 1996. Genetic markers, map construction, and their application in plant breeding. Hort. Science. 31, 729-740.

Tautz, D., Trick, M., Dover, G. A., 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature. 322, 652-656.

Terachi, T., 1993. The progress of DNA analyzing techniques and its impact on plant molecular systematics. J. Plant Res. 106, 75-79.

Thormann, C. E., Ferreira, M. E., Camargo, L. E. A., Tivang, J. G., Osborn, T. C., 1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88, 973-980.

Tsumura, Y., Ohba K., Strauss, S. H., 1996. Diversity and inheritance of inter simple sequence repeat polymorphisms in Douglasfir (Pseudotsuga menziesii) and sugi (Cryptomeria japonica). Theor. Appl. Genet. 92, 40–45.

Vejl, P., 1997. Identification of genotypes in hop (Humulus lupulus L.) by RAPD analsysis using program Gel Manager for Windows. Rostlinna Vyroba. 43, 325-331.

Vilanova, S., Romero, C., Abbott, A. G., Llacer, G., Badenes, M. L., 2003. An apricot (Prunus armeniaca L.) F2 progeny linkage map based on SSR and AFLP markers mapping plum pox virus resistance and self-incompatibility traits. Theor. Appl. Genet. 107, 239–247.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee T., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, m., Zabeau, M., 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23, 4407–4414.

Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T., Itakura, K., 1979. Hybridization of synthetic oligodeoxyribonucleotides to phi 174 DNA: the effect of single base pair mismatch. Nucleic Acids Res. 6, 3543–3557.

Wang, D., Karle R., Iezzoni, A. F., 2000. QTL analysis of flower and fruit traits in sour cherry. Theor. and Appl. Genet. 100, 535-44.

Wang, G., Mahalingan, R., Knap, H. T., 1998. (C-A) and (G- A) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, Glycine max (L.) Merr. Theor. Appl. Genet. 96, 1086–1096.

Wang, Z.,.Weber, J. L., Zhong, G., Tanksely, S. D., 1994. Survey of plant short tandem DNA repeats. Theor. Appl. Genet. 88, 1–6.

Warburton, M. L., Bliss, F. A., 1996. Genetic diversity in peach (Prunus persica L. Batch) revealed by randomly amplified polymorphic DNA (RAPD) markers and compared to unbreeding coefficients. J. Am. Soc. Hort. Sci. 121, 1012-19.

Weber, D., Helentjaris, T., 1989. Mapping RFLP loci in maize using B-A translocations. Genetics. 121, 583-590.

Weising, K., Nybom, H., Wolff, K., Meyer, W., 1995. Applications of DNA Fingerprinting in Plants and Fungi. CRC Press, Boca Raton.

Weller, J. W., Reddy, A., 1997. Fluorescent detection and analysis of RAPD amplicons using the ABI PRISM DNA sequencers. In: Micheli, M. R., Bova, R., (ed) Fingerprinting methods based on arbitrarily primed PCR, Springer Lab Manual, 81-92.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., Tingey, S. V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535.

Winter, P., Kahl, G.,1995. Molecular marker technologies for plant improvement. World Journal of Microbiology & Biotechnology. 11, 438–448.

Wu, K., Jones, R., Dannaeberger L., Scolnik, P. A., 1994. Detection of microsatellite polymorphisms without cloning. Nucleic. Acids. Res. 22, 3257-3258.

Young, N. D,, Menancio-Hautea, D., Fatokun, C. A., Danesh, D., 1992. RFLP technology, crop improvement, and international agriculture. In: Thottappilly, G., Monti, L. M., Moham, D. R., Moore, A. W., (ed) Biotechnology: Enhancing research on tropical crops in Africa. Technical Center for Agriculture and Rural Cooperation, International Institute of Tropical Agriculture, 221-230.

Zietkiewicz, E., Rafalski, A., Labuda, D., 1994. Genome finger- printing by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. Genomics. 20, 176–183.