MASTER THESIS

Thanyarat Onlamun

2023

Hungarian University of Agriculture and Life Sciences
THE FORMER FACULTY OF HORTICULTURAL SCIENCE

Designing a new set of molecular markers to reveal polymorphism in the coding region of European plum genome

Thanyarat Onlamun
MSc in Agricultural Biotechnology

Made at the department of Plant Biotechnology
Supervisors:
Beti Ivanovska
Attila Hegedűs PhD, DSc

Budapest, 2023

…….
Head of Department

…….
Supervisor

CONTENTS

1. USED ABBREVIATIONS

AFLP – Amplified fragment length polymorphism

BLAST – Basic Local Alignment Search Tool

Bp – Base pair

DNA – Deoxyribonucleic acid

EST-SSRs – Expressed sequence tag-derived simple sequence repeat markers

FAO – Food and Agriculture Organization of the United Nations

FaSt – *FallingStones* transposon

GDR – Genome Database for *Rosaceae*

Ha – Hectare

IRAP – Inter-Retrotransposon Amplified Polymorphism

ISSR – Inter-simple sequence repeat

MITE – Miniature inverted repeat transposable element

NCBI – National Center for Biotechnology Information

PCA – Principal component analysis

PCR – Polymerase chain reaction

PIC – Polymorphism information content

PPV – *Plum pox virus*

RAPD – Random amplified polymorphic DNA

RKN – Minor root-knot nematodes

SCAR – Sequence Characterized Amplified Region

SCoT – Start codon targeted

SSR – Simple Sequence Repeats

T – Ton

Ta – Annealing temperature

TBE – Tris-borate-EDTA

Tm – Melting temperature

UPGMA – Unweighted pair-group average algorithm

V – Volt

2. INTRODUCTION

Fruit tree culture faces many challenges worldwide that make continuous breeding activities inevitable. The biggest limitation in the breeding of tree species is the long vegetative period of trees which allows selection for specific traits only after 4-5 years of growth. Keeping the trees in the orchard and providing the required nutrients and plant protection treatments make the process not just long but expensive as well. Molecular markers can help decrease the costs by eliminating the offspring not carrying the required alleles. Several reliable markers are available for specific favorable agronomic traits (e.g., self-compatibility, early and late fruit maturity dates, weeping growth habit). However, the establishment of such an efficient molecular marker requires the identification of candidate genes and mutations rendering specific phenotypes.

The cost-efficient and high-throughput sequencing platforms provide a huge amount of transcriptomic data that might help identify genomic regions responsible for specific traits. However, the application of such assays is still limited in most countries while breeding programs are suffering from the lack of robust markers. Many available molecular markers are associated with the non-coding part of the genome and hence they are useful in the characterization of genetic diversity and testing the relatedness of accessions but not in following the inheritance of an important phenotypic trait.

The genomic arrangement of *Prunus* species is more and more known due to structural genomics studies. Several transposable elements were identified and shown to have a frequent occurrence in the gene-rich euchromatin segments of chromosomes. Molecular marker systems were also created to amplify fragments of genes but the flanking regions may also contain information (mutations in promoter, untranslated regions etc.). However, a strategy using a combination of PCR primers annealing to the transposable elements located close to genes and those annealing to the conserved sequences of genic regions has not yet been tested to estimate genetic variability in *Prunus*. This approach might be used to detect intraspecific variations and provide a low-cost marker system to follow advantageous phenotypic traits in breeding programs.

3. OBJECTIVES

The ultimate aim of this study was to establish a fully new molecular marker strategy for *Prunus* stone fruit species that is able to detect variations predominantly in the coding part of their genome. The basic steps to establish such a marker system in the hexaploid European plum (*Prunus domestica*) were the followings:

To detect polymorphisms by using combinations of Start Codon Targeted (SCoT) and *FallingStone*s (*FaSt*) primers.
Choosing combinations of *FaSt* and SCoT
Characterizing the information content of *FaSt* and SCoT marker combinations
Identifying variations detected by different combinations of *FaSt* and SCoT primers
Determining the DNA sequences of fragments amplified by *FaSt* and SCoT marker combinations
Characterizing marker efficiency based on sequence analyses

Our studies form an initial step in marker development, to see if the recently identified *Prunus*-specific miniature inverted repeat transposable elements, *FallingStones*, might be used to detect genetic polymorphism in European plum. This transposon accumulates in the heterochromatin regions of the genome, so combination with gene-targeted markers (SCoT) may detect DNA sequence variations in the coding part of the genome.

4. REVIEW OF LITERATURE

4.1 Basic features of European plum

The genus *Prunus* (a member of the *Rosaceae* family) comprises about 400 species of trees and shrubs that produce drupes as fruits, commonly called "stone fruits" (Ben, 2015). There are many diploid (*P. salicina* Lindl., *P. cerasifera* Ehrh. and *P. simonii* Carr.) as well as polyploid (*P. spinosa* L. and *P. domestica* L.) species. European plum (*Prunus domestica*) is an important crop with a predominant use in many European countries. It is a hexaploid (2n = 6x = 48) fruit tree species. The plum has a fleshy fruit that called a drupe and is formed by the thickening of the ovary wall after fertilization. Plum is one of the major commercial stone fruits also including cherries (*P. avium* L. or *P. cerasus* L.), apricots [*P. armeniaca* L. and *P. mume* (Sieb.) Sieb. & Zucc.], peach [*P. persica* (L.) Batsch] or almond (*P. dulcis* L.).

Within the *Prunus* genus five subgenera are differentiated (*Amygdalus*, *Cerasus*, *Prunus*, *Padus* and *Laurocerasus*) according to Rehder, 1949. Three clades of *Prunus* were named after their typical inflorescence structures (Chin et al. 2013): 1) the 'Solitary' clade includes the peaches, almonds apricots and plums (subgenera *Amygdalus* and *Prunus*); 2) the '*Corymbose*' group involves cherries (subgenus *Cerasus*); and 3) the '*Racemose*' clade is comprising the species in *Padus* and the *Laurocerasus* subgenera.

The European plum (*Prunus domestica*) probably originated in eastern Europe or western Asia around the Caucasus and the Caspian Sea. Current evidence suggests that *P. domestica* originated is an interspecific hybrid of a diploid *P. cerasifera* and a tetraploid *P. spinosa* that itself may have been an interspecific hybrid of *P. cerasifera* and an unknown Eurasian plum species (Zhebentyayeva et al. 2019). This complexity of the *P. domestica* genome makes the sequencing and assembly of the genome particularly challenging but a first draft genome was published using Illumina short-read technology (Callahan et al. 2021). The assembled genome was annotated resulting in 130,866 gene models and available on Genome Database for *Rosaceae* (GDR), https://www. rosaceae.org/8.

European plums are consumed as fresh fruit as a healthy snack or a dessert. Plums also have low caloric content and high nutritive value. Carbohydrates, including sucrose, glucose, and fructose, as well as organic acids such as citric and malic acids are contained in these fruits. The fruits are high in health-promoting antioxidants such as flavonoids, polyphenols such as anthocyanins, and carotenoids, such as cryptoxanthin, lutein, and zeaxanthin (Jayasankar et al. 2016). Dried plums or prunes as we normally call are sweet taste and juicy. Plums and prunes are known for their useful effect. This effect has been attributed to various compounds present in the fruits, such as dietary fiber, sorbitol (Jason, 2007). Prunes are known to help regulate in the digestive system. Highly used in people who on a diet. In plum, varieties of plum with dark purple colored skin showed 200% higher total phenolic than others (Rupasinghe et al. 2006). The Japanese plum cvs. Black Beauty and Angeleno were especially rich in phenolics (Tomás et al. 2001). Plum species and cultivars are very diverse in fruit characteristics for example shape,

size, aroma, color, texture, taste, and quality. Plant characteristics are also very diverse, ranging from shrubs to large trees, spreading to upright, thick to thin leaves, and early to late blooming (Ramming and Cociu, 1990).

4.2 Economic importance of European plum

The European plum (*Prunus domestica* L.) is a valuable fruit-bearing plant that belongs to the *Rosaceae* family and is grown predominantly in Europe (Sottile et al. 2022). Unfortunately, the FAO database contains only information for all plums and sloes (merging data for diploid, tetraploid, and hexaploidy crop species). However, the latest official figure for plum production is 12014482 tons, produced in 90 countries (Faostat, 2021). The estimated value of the harvested area is 2602436 ha, and the yield is 4.62 t/ha.

The main producer country is China, but its production covers diploid species. The main producers for the hexaploid European plum are Romania (807'170 t) and Serbia (412'778 t). It takes a significant proportion of production quantity in many European regions, like the Balkan countries (Republic of Moldova, Bosnia and Herzegovina, North Macedonia etc.), or Central (Hungary, Poland, Czech Republic etc.) and Western Europe (France, Germany etc.) (Faostat, 2021). In Hungary, environmental conditions are optimal for growing plums, and hence it is the second *Prunus* species in production quantity with approx. 33,250 t (Faostat, 2021). The plum season lasts in Hungary for almost 3 months.

Plums are consumed in varied ways as fresh fruit or processed into jam, marmalade, juice, prunes (dried fruit), or spirits. Plums are also used for making popular pálinkas. Pálinka is a special distilled product, „Hungaricum" with deep traditions and is protected as a geographical indication of the European Union (Hegedűs et al. unpublished).

4.3 Breeding techniques, aims and results

Hybridization along with selection of clonal variant have been widely used in plum cultivars. This technique is a dominant technique. Seedlings are produced by this technique. However, long reproductive cycle with long juvenile periods, complex reproductive biology and high degree of heterozygosity are the main problems of conventional breeding. Frequently, as a way to obtain a new offspring that meets the desired agronomic and commercial characteristics it is necessary to perform several rounds of introgressive backcrossing (Petri and Scorza, 2008). European plum has the average generation time around 3-7 years and it would take 15-20 years to release the fruits.

There are a great number of breeding objectives to improve European plum cultivars. The main breeding goals include resistance to biotic and/or abiotic stress, chilling requirements, tree size, productivity and fruit quality traits (Callahan, 2008; Neumüller, 2011). However, Most of the European plum are cultivated in regions with severe weather and cold hardiness which make European plum breeding has some challenging. Related to plum-affecting diseases breeding efforts have been focused on: brown rot, caused by the fungus *Monilinia spp*.; bacterial canker, caused by *Pseudomonas syringaevan* Hall; bacterial spot, caused by *Xanthomonas campestris* pv. and plum leaf

scald, caused by the bacteria *Xylella fastidiosa*, the most important disease affecting stone fruits, caused by the *plum pox virus* (PPV) (Petri et al. 2018). The long vegetative period of plum made using recent biotechnology a perspective approach for the improvement of European plums. Although induced mutations and polyploidy were carried out in a limited number of programs, it does not play a role in current breeding programs (Neumüller, 2011). *In vitro* micropropagation techniques have been complicated for plant multiplication and rooting, but the nature of plums makes the interventions very challenging. In most of woody fruit species, transformation and adventitious regeneration are difficult, with low efficiency and often limited to a few genotypes or to seed-derived tissues (Petri and Burgos, 2005) However, the European plum has been the most successful species among Prunus to transform. (Petri et al. 2018). Nowadays, two mains advantage have been used for European plum are genetic transformation and regeneration. First, Transformation may require less time, labor, and field space. Second, Transformation may improve some genetic traits. Genetic engineering was used in *P. domestica* for a single purpose: to induce resistance against the most devastating viral disease of *Prunus* trees, *plum pox virus* (PPV). Two strategies were applied: expressing parts of the PPV coat protein gene or the host translation initiation factor *eIF(iso)4E* gene. Both strategies resulted in transgenic plants with stable resistance to PPV (Scorza et al. 1994; Wang et al. 2013).

4.4 Molecular markers in European plum breeding

Several molecular markers have been used previously for studies of plums, including genetic diversity and cultivar identification, such as Restriction Fragment Length Polymorphism, Random Amplified Polymorphic DNA, Amplified Fragment Length Polymorphism (Aradhya et al. 2004), Simple Sequence Repeat and Inter-Simple Sequence Repeat.

In some plant families, the basic knowledge of chromosome number and ploidy level is essential because hybridization and ploidy levels can cause problem. There are not yet molecular markers for agronomic traits available in *Prunus domestica*, which could be applied in breeding programs, due to the highly polymorphic hexaploid genome of this species (Neumüller, 2011).

4.5 Random amplified polymorphic DNA (RAPD)

Random amplified polymorphic DNA belongs to the first generation of PCR based molecular marker strategies (Williams et al. 1990). Its development does not require preliminary sequence information as it applies randomly chosen, 10-nt oligomer primers to arbitrarily amplify regions of the tested genome. Differences in banding patterns might be detected among species and varieties and used for designing more specific markers.

The first study in *P. domestica* tested 145 primers, the size of the amplification products ranged from 250 to 2600 bp, with an average of 9 bands per primer/cultivar combination; however, only 6 primers detected enough genetic variation to allow complete differentiation of all the cultivars tested (Gregor et al. 1994). RAPD markers were used for cultivar identification in European plums many times (Ortiz et al. 1997; Yu et al. 2013) and also for

characterizing genetic diversity and revealing phylogenetic relationships among related wild species and other germplasm resources (Liu et al. 2006). European and Japanese plums have been distinguished based on random amplified polymorphic DNA analysis or RAPD. (Shimada et al. 1999).

However, the shortcomings of RAPD markers including their dominant expression (the homo- and heterozygotes cannot be differentiated) and especially their poor repeatability and reliability made them outdated and other markers became predominant.

4.6 Sequence Characterized Amplified Region (SCAR)

Sequence characterized amplified regions or SCAR is a PCR based method. Sequencing markers amplified was developed and converted into this technique. The process, longer primer can provide the greater degree of specificity so designing longer primers are better. Accomplished by converting dominant RAPD to co-dominant SCAR can give better assessment of F2 individuals. The researcher suggested that *Prunus* species has resistance to controls all major and minor root-knot nematodes (RKN) species tested especially *Prunus cerasifera.* In *P. cerasifera*, the clones P.2175 and P.2980 are heterozygote\use to test for *Ma* single dominant gene and carry *Ma1* and *Ma3* alleles. The result found that the allele has a high-level resistance to root-knot nematodes. Consequently, these SCARs appear to be powerful tools to screen for RKN resistance conferred by the *Ma* gene (Lecouls et al. 2004)

4.7 Simple Sequence Repeats (SSR) or microsatellites

Microsatellites or simple sequence repeats are tandem repeats of 1-6 bp sequence motifs in plants' genomes (Gupta et al. 1996). Due to replication slippage, the number of motifs is highly variable, which makes intra-specific differentiation possible. The flanking regions of SSRs are conserved and hence PCR primers can be designed to anneal to the flanking regions and amplify the SSRs. An accurate length determination (e.g. sizing the fluorescently labeled amplicons on an automated sequencer) can detect size variants, i.e. alleles. The development of SSR markers requires sequence information but they are co-dominant (differentiate homo- and heterozygotes) and characterized by reliable repeatability and transferability, meaning the markers can be applied in other species, as well (Hormaza, 2002).

SSR analysis was used to distinguish *P. domestica* from *P. cerasifera* and *P. spinosa* as well as to identify different genetic groups, where damson plums were clearly separated from greengage plums (Horvath et al. 2011). SSR markers were used to distinguish cultivars in Hungary (Makovics et al. 2017), Spain (Urrestarazu et al. 2018), and many other countries in Europe (Xuan et al. 2011; Gasi et al. 2020) were also used to genotype embryos to identify pollen donors and the success rate of individual pollinizers (Meland et al. 2020). Kazija et al. 2013 used a set of 11 SSR markers to describe variations within the 'Besztercei' plum germplasm.

EST-SSRs (i.e. SSR markers part of a transcribed DNA sequence) were also identified and used in *P. domestica*. EST-SSRs detected a significantly smaller number of alleles but the genotypic evenness and the observed heterozygosity were comparable to genomic SSRs (Manco et al. 2019). Decroocq et al. 2004 demonstrated the utility of apricot EST-SSR markers for genotype fingerprinting of the hexaploid plum cultivars. Their results indicated three Cacak cultivars ('Cacanska najbolja', 'Cacanska rana' and 'Cacanska lepotica') originated from a cross between 'Stanley' and 'Ruth gerstetter'.

4.8 Amplified fragment length polymorphism (AFLP)

The AFLP technique is based on the amplification of genomic restriction fragments by using PCR. Firstly, Restriction enzyme is used to cut the DNA and ligate the double-stranded adapted to the ends of the DNA fragment to generate template DNA for amplification. Only restriction fragments in which the nucleotides flanking the restriction site match the selective nucleotides will be amplified (Vos et al. 1995). Among PCR-based molecular markers, AFLPs are highly reproducible multi-locus marker systems developed by Vos et al. 1995. The advantages of AFLP are High levels of polymorphism and high degrees of discriminative capacity. This method has been widely used to identify the genetic variability in many species including fruit trees, but the use of this method has been very limited in plums.

One of these studies compared only cherry plum belong to *Prunus cerasifera*. The objective of this study was to characterize 14 cultivars in turkey by using AFLP markers and use to determine whether AFLP markers are appropriate for describing taxonomic relationships. The results suggested that AFLP is a good method to determine genetic relatedness among plum cultivars (Llgin et al. 2009). Another experiment to analyze *P. domestica* genotypes in Iraq used three primer combinations that generated a total of 106 bands and among them 86 were polymorphic (81.1%), while 20 (18.9%) were monomorphic (Ali et al. 2015). Firstly, Plums were observed by using AFLP and using UPGMA to analysis the relatedness. The results suggested that AFLP was a good marker to determine genetic relatedness among plum genotypes in Duhok city, Iraq. Another study, twenty plum samples from Mediterranean region were used to analyze their genetic diversity by using AFLP. The genetic diversity showed the value from 0.829-0.985 which was low similarity. However, the majority of green plum genotypes derived from an ecologically distinct location within the Mediterranean region were more diverse genetically than those from other locations (Ayanoğlu et al. 2007).

4.9 Inter-simple sequence repeat (ISSR) markers

ISSR is DNA fragment and amplified by PCR using primers with 15-25 base pairs and it should consist of repeated DNA motifs. ISSR-PCR is usually conducted with an annealing temperature (Ta) of 45–60°C, depending on the melting temperature (Tm) of the ISSR primer (Reddy et al. 2002).

The ISSR markers were applied to analyze the genetic diversity and relationships among 30 individual plants from 5 types of *P. domestica* germplasm across Xinjiang (Sun et al. 2015). Froom 0.5552 to 0.9968 was the genetic similarity coefficients value which mean they showed a certain level of genetic diversity. 3 groups were divided such as European plum varieties, Jiashi Smoked plum and wild European plums. ISSR was used to assess diversity and relatedness of 23 Moroccan *P. domestica* and *P. salicina* cultivars (Ait et al. 2021). From the result, ISSR markers have proved to be efficient. The percentage of polymorphic bands obtained (83.14%) was comparable to that found in previous reports by Athanasiadis et al. 2013 and Ali et al. 2015 in Greek plum cultivars based on ISSR markers (81.81%) and Iraq plum cultivars using AFLP markers (81.1%). Another study, *ex situ* collection of plum cultivars was conducted in Morocco by using 20 ISSR primers. A genetic variability showed 95.68% as the high polymorphism rate. Overall results showed that Plum cultivars in Morocco was characterized by high genetic diversity. Also, polymorphism information content (PIC) showed that used ISSR markers were informative and relevant for discriminating the cultivars evaluated (Hamdani et al. 2022)

4.10 Start codon targeted (SCoT) markers

SCoT is a relatively novel marker system described by Collard and Mackill 2009. The method is based on the short, conserved region in plant protein encoding genes surrounding the ATG translation start codon that has been well characterized. Genetic fidelity of *in vitro* cultures of plum (*P. salicina*) cvs. Santa Rosa and Frontier multiplied for 5 years (60 passages) through enhanced axillary bud proliferation was tested and compared with 22-year-old *in vitro* raised and mother plants of respective cultivars using SCoT marker. Eight primers out of twenty-six generated reproducible bands, thus suggesting that *in vitro* propagation using axillary buds is the safest mode to produce clonal planting material in plums (Manisha et al., 2021).

A total of 289 polymorphic bands were amplified with 23 SCoT primers, showing a polymorphism percentage of 97.94% and an average is 12.6 polymorphic bands per primer. The SCoT21, SCoT32, and SCoT53 primers amplified up to 17 bands, and the polymorphism percentage was 100%. The minimum number of bands amplified by SCoT3 was 9, and the polymorphism percentage was 90%. Therefore, SCoT molecular markers were shown to be highly polymorphic and suitable for genetic diversity studies of *P. sibirica* in Inner Mongolia. The results indicated that the efficient SCoT molecular marker-based genetic diversity analysis of *P. sibirica* in Inner Mongolia can provide a reference for *P. sibirica* variety breeding and resource development (Buer et al., 2022)

4.11 Inter-Retrotransposon Amplified Polymorphism (IRAP)

Inter-Retrotransposon Amplified Polymorphism or IRAP is retrotransposon based molecular marker. This method is usually used to study phylogenetic relationship and genetic variability. 23 *Prunus domestica* varieties was analyzed with comparing Cassandra IRAP fingerprints and constructing a dendrogram based on the genetic similarity to obtain polymorphic profiles. IRAP-based polymorphism was proved to be a suitable tool for analyzing

the DNA variability of European plums. The study was proved that Cassandra is also suitable for making an IRAP primer. Unique fragments were obtained for the cultivars Švestka domácí, Čačanská ranná and Elena and specific banding profile was obtained for genotypes from Germany, Slovakia and former Yugoslavia (Senková et al. 2017).

4.12 FallingStones (FaSt)

Despite their high copy numbers in genomes, only a limited number of active MITEs have been identified so far in species like rice and maize (Zhang et al. 2001; Kikuchi et al. 2003; Shirasawa et al. 2012). *FallingStones* (*FaSt*) is a new non-autonomous transposon in the *Prunus* genome. It was first found in apricot. It was identified a miniature inverted repeat transposable element or MITE with features (349-bp size, 82-bp terminal inverted repeats and 9-bp target site duplications) that are consistent with this MITE being a putative member of the Mutator-like transposase superfamily (Halász et al. 2014). This non-autonomous element was labeled as *FaSt*. *FaSt* showed a preferential accumulation in the short AT-rich segments of the euchromatin region of the peach genome (Fig. 1). The pattern highlights the frequence of occurrence of *FaSt* in the gene-rich part of peach and other stone genomes.

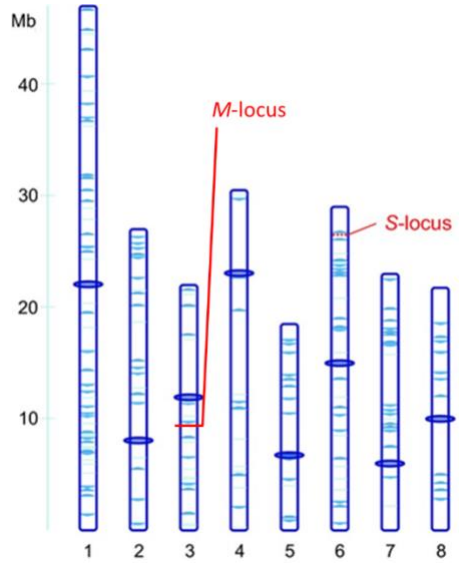

Fig. 1. The position of *FallingStones* (*FaSt)* transposon sequences on *Prunus persica* chromosomes (Halász et al., 2014). Triangles and lines indicate the full length and partial elements, respectively. The position of centromer is shown by an oval in each chromosome. Red lines refer to genes with disrupted function due to *FaSt* insertions.

In apricot, the same transposon was observed to be inserted in two genes located on chromosome 3 (*disulfide bond A-like oxidoreductase*) and 6 (*S-haplotype-specific F-box*). Interestingly, the loss-of-function mutation of both genes resulted in the same phenotypical consequences, breaking the self-incompatibility barrier and rendering the mutant accessions being self-compatible. Those examples prove *FaSt* might be inserted in coding part of certain genes and contribute to the formation of genetic variability in *Prunus.*

.

5. MATERIALS AND METHOD

5.1 Plant material

The study was performed on 13 domestic plum (*Prunus domestica* L.) cultivars: 'Toptaste', 'Haroma', Topend plus', 'Haganta', 'Presenta', 'Tophit', 'Topfive', 'Empress', 'Elena', 'Hanita', 'Jojo', 'Nemtudom P3', and 'Besztercei Bt2'. 11 bud samples were obtained from Soroksári Botanical Garden of the Hungarian University of Agriculture and Life Sciences, Hungary and 2 samples were obtained from tissue culture. Samples were stored in -80^{o}C until used for DNA extraction.

5.2 DNA extraction

The genomic DNA extraction was performed from buds using a Dneasy Plant Mini Kit (Qiagen, Hilden, Germany). Afterwards, The DNA concentration and purification parameters were measured by Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The SCoT primers (Start Codon Targeted markers; Collard et al. 2009) combined with *FaSt* (*FallingStones*, Halász et al. 2014), *FaSt* sequences: FaSt-Rev2 (5'-TCTTAGAAATTACAAAACTACC-3') were used for PCR amplification. In PCR reaction was performed in PCR Thermal Cycler in a total volume of 13 µl reaction mixture by adopting the following PCR program: 5 min 94°C, 40 cycles of 1 min at 94°C, 1 min at 49°C, 2 min 72°C. The success of the PCR reactions was monitored by running 9 µl of PCR reactions on 1.2% agarose in 1x TBE buffer with 80 V applied current.

5.3 Cloning and sequencing

The PCR products were chosen for cloning in case of each having a unique pattern of differently sized bands. For cloning, PCR products were carried out using pTZ57R/T vector (Thermo Fisher Scientific). The ligated plasmid vectors were transformed into JM109 Escherichia coli competent cells (Zymo Research Corp., Irvine, CA, United States). After that the blue-white screening method was used to visualize the successfully transformed cell. Colony PCR was conducted by using M13 primers. PCR products were run on an agarose gel to confirm the presence or absence.

Table 1. SCoT primer sequences.

SCoT	primer Sequence (5'-3')	%GC
1	CAACAATGGCTACCACCA	50
2	CAACAATGGCTACCACCC	56
3	CAACAATGGCTACCACCG	56
4	CAACAATGGCTACCACCT	50
5	CAACAATGGCTACCACGA	50
6	CAACAATGGCTACCACGC	56
7	CAACAATGGCTACCACGG	56
8	CAACAATGGCTACCACGT	50
9	CAACAATGGCTACCAGCA	50
10	CAACAATGGCTACCAGCC	56
11	AAGCAATGGCTACCACCA	50
12	ACGACATGGCGACCAACG	61
13	ACGACATGGCGACCATCG	61
14	ACGACATGGCGACCACGC	67
15	ACGACATGGCGACCGCGA	67
16	ACCATGGCTACCACCGAC	56
17	ACCATGGCTACCACCGAG	61
18	ACCATGGCTACCACCGCC	67
19	ACCATGGCTACCACCGGC	67
20	ACCATGGCTACCACCGCG	67
21	ACGACATGGCGACCCACA	61
22	AACCATGGCTACCACCAC	56
23	CACCATGGCTACCACCAG	61
24	CACCATGGCTACCACCAT	56
25	ACCATGGCTACCACCGGG	67
26	ACCATGGCTACCACCGTC	61
27	ACCATGGCTACCACCGTG	61
28	CCATGGCTACCACCGCCA	67
29	CCATGGCTACCACCGGCC	72
30	CCATGGCTACCACCGGCG	72
31	CCATGGCTACCACCGCCT	67
32	CCATGGCTACCACCGCAC	67
33	CCATGGCTACCACCGCAG	67
34	ACCATGGCTACCACCGCA	61
35	CATGGCTACCACCGGCCC	72
36	GCAACAATGGCTACCACC	56

of the gene of interest. The differently sized plasmid DNA fragments were purified using the EZ-10 Spin Column Plasmid DNA kit (Bio Basic Inc., Markham, Canada) and then sequenced in an automated sequencer ABI PRISM 3100 Genetic Analyzer (Applied Biosystems).

5.4 Sequence analysis

The Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI) and *P. domestica* genome sequences in the Genome Database of *Rosaceae* (https://www.rosaceae.org/blast) were used for homology searches (Altschul et al. 1990). Alignments were created using MEGA7 (Kumar et al. 2016) and were presented with the BioEdit program v.7.2.5 (Hall, 1999).

5.5 Data analysis

PCR bands detected on gels were scored as absent (0) or present (1), only clear reproducible bands were scored. For the diversity and phylogenetic analyses. The unweighted pair-group average algorithm (UPGMA) was used to construct a dendrogram based on Dice similarity coefficients with the software PAST 2.17c (Hammer et al. 2001). Numbers on branches represent bootstrap supports from 1000 replicates. Principal component analysis (PCA) was also carried out using PAST software.

6. RESULTS

6.1. Testing the polymorphism produced by the combination of *FaSt* and different SCoT primers

We used different SCoT primers in combination with a reverse primer annealing to a unique sequence region of the *FallingStones* transposon to detect genetic polymorphism among 13 European plum cultivars. The assayed cultivars include popular commercial cultivars from German and Italian breeding programs and Hungarian landrace selections. A total of 34 SCoT primers were used (SCoT1, SCoT2, SCoT3, SCoT4, SCoT5, SCoT7, SCoT8, SCoT9, SCoT10, SCoT11, SCoT12, SCoT13, SCoT14, SCoT15, SCoT16, SCoT17, SCoT18, SCoT19, SCoT20, SCoT21, SCoT22, SCoT23, SCoT24, SCoT25, SCoT26, SCoT27, SCoT28, SCoT29, SCoT30, SCoT32, SCoT33, SCoT34, SCoT35, and SCoT36.

These primer combinations produced variable results with some of those not amplifying a single band in any cultivars (SCot20, SCoT29, SCoT30, and SCoT35). A subset of primer combinations resulted in the amplification for most of the assayed samples except only some of the assayed cultivars. Those primers were the followings: SCoT10, SCoT12, SCoT16, SCoT17, SCoT21, SCoT22, SCoT23, SCoT24, SCoT25, SCoT26, SCoT27, SCoT33, and SCoT36.

The best combinations provided a great number of amplicon sizes exceeding 100 and resulted in a very complex pattern (Fig. 2). Characteristic alterations were also observed in the banding pattern of different cultivars. An example is the band around 200-bp being present in ‘Haroma’, Topend plus’, ‘Topfive’, ‘Empress’, ‘Elena’ and ‘Hanita’, while missing from all the other samples. Several more differentiating fragments were noted, like the slightly bigger (approx. 230-bp) occurring in samples ‘Topfive’, ‘Empress’, ‘Elena’ and ‘Hanita’. The fragment around 600-bp size was amplified the in the same cultivars and also in ‘Jojo’, ‘Nemtudom P3’, and ‘Besztercei Bt2’. The band around 1000-bp was characteristic for the samples ‘Toptaste’, ‘Haroma’, ‘Presenta’, ‘Topfive’, ‘Empress’, and ‘Jojo’. The fragment of 1500-bp were only amplified in samples of ‘Haroma’, ‘Tophit’, ‘Topfive’, ‘Empress’, ‘Elena’, and ‘Jojo’. The one with 2000-bp size amplified exclusively in cultivar ‘Jojo’.

However, many of the fragments were not polymorphic and the complex pattern made the evaluations of the gels challenging. The intensity of the bands also varied and it was problematic to decipher the presence of faintly appearing bands while in other cases sharp and intense fragments were observable. Since many of the combinations produced enough differentiating bands, only the ones easily identified were used for further analysis.

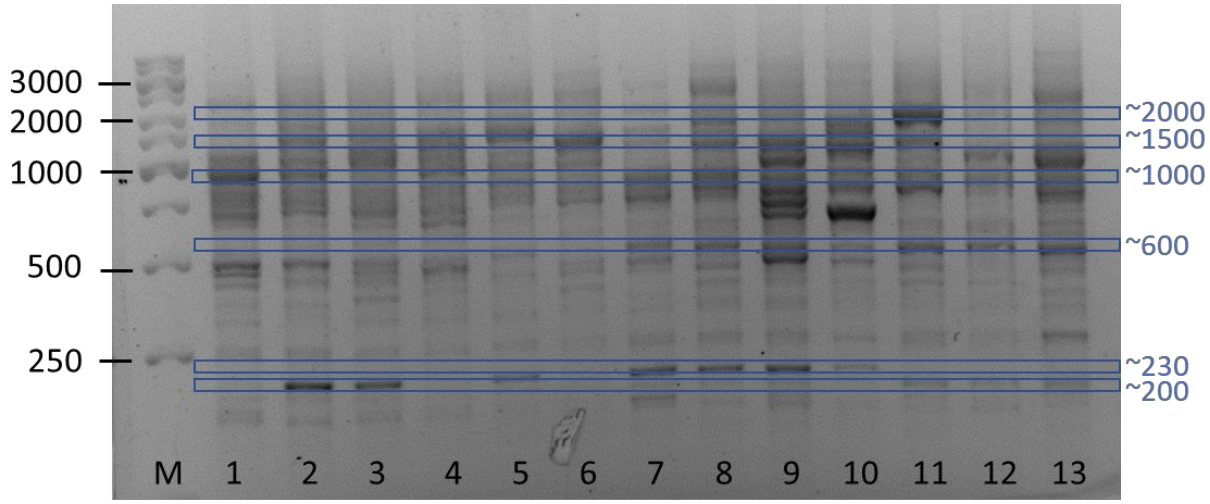

Fig. 2. Patterns of amplified fragments in polymerase chain reaction of European plum cultivars using a combination of forward and reverse primers, SCoT1 and FaSt-Rev2, respectively. Blue squares indicate the analysed differentiating fragments. Labels are the followings: (M) 1-kb DNA ladder, (1) 'Toptaste', (2) 'Haroma', (3) Topend plus', (4) 'Haganta', (5) 'Presenta', (6) 'Tophit', (7) 'Topfive', (8) 'Empress', (9) 'Elena', (10) 'Hanita', (11) 'Jojo', (12) 'Nemtudom P3', and (13) 'Besztercei Bt2'.

The combinations providing the most informative patterns were SCoT1, SCoT2, SCoT4, SCoT7, SCoT9, SCoT18, and SCoT19, some of which are shown in Fig. 3. The gels demonstrate the differences in the number of amplified fragments and polymorphism detected among different cultivars. The primer combination of SCoT1 and FaSt-Rev2 amplified many more fragments than SCoT34 and FaSt-Rev2. The polymorphism percentages value sof European plum cultivars using a combination of several SCoT primers and FaSt-Rev2 were as follows: SCoT1, SCoT4, SCoT7, SCoT9, SCoT15, and SCoT34 were 41.8%, 31.7%, 82.7%, 92.3%, 100% and 60.0% respectively. The most polymorphic combination was the SCoT15 and FaSt-Rev2, while the lowest was the SCoT4-FaSt-Rev2. The less complex patterns allowed a much accurate identification of differentially appearing fragments for most of these combinations, but changes in the patterns of amplified fragments were also observed in some combinations.

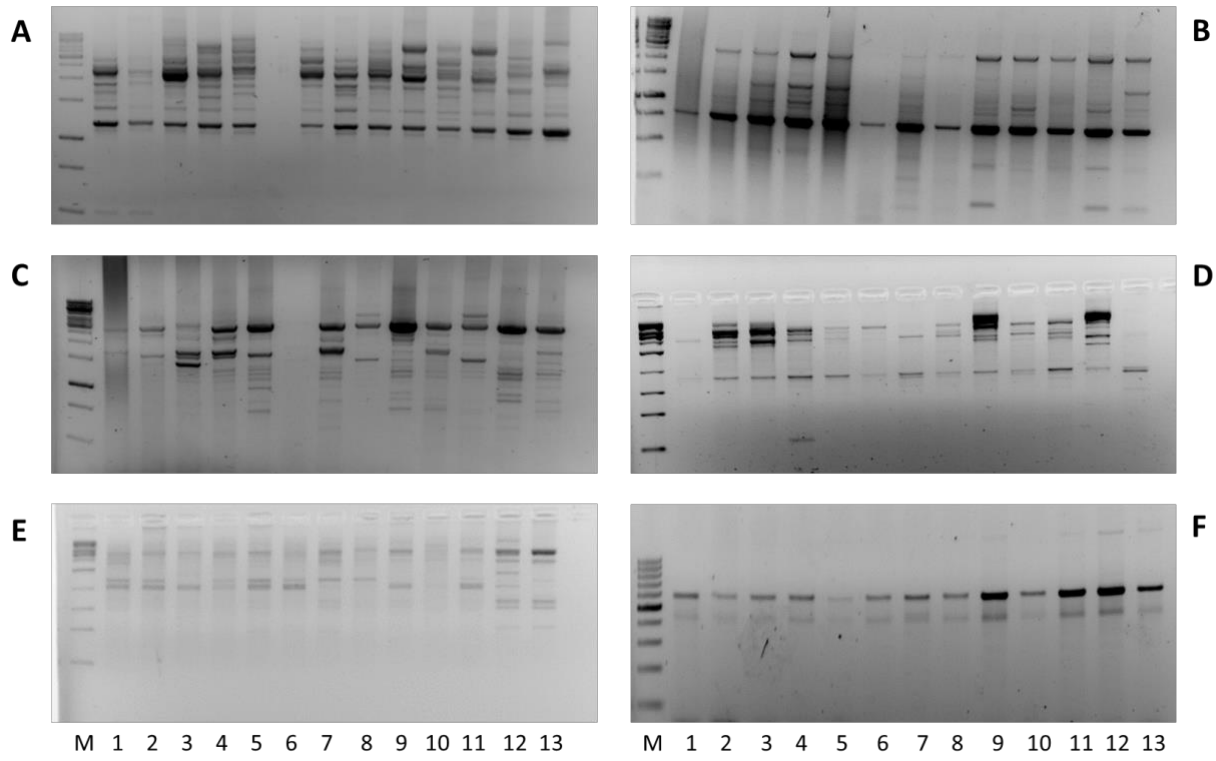

Fig. 3. Patterns of amplified fragments in polymerase chain reaction of European plum cultivars using a combination of several SCoT primers and FaSt-Rev2: (A) SCoT1, (B) SCoT4, (C) SCoT7, (D) SCoT9, (E) SCoT15, and (F) SCoT34. Labels are the followings: (M) 1-kb DNA ladder, (1) 'Toptaste', (2) 'Haroma', (3) Topend plus', (4) 'Haganta', (5) 'Presenta', (6) 'Tophit', (7) 'Topfive', (8) 'Empress', (9) 'Elena', (10) 'Hanita', (11) 'Jojo', (12) 'Nemtudom P3', and (13) 'Besztercei .

We wanted to see how reliable the amplified patterns are, hence we carried out several repeated PCRs with some of the primer combinations. The banding patterns of three independent PCRs are shown in Fig. 4.

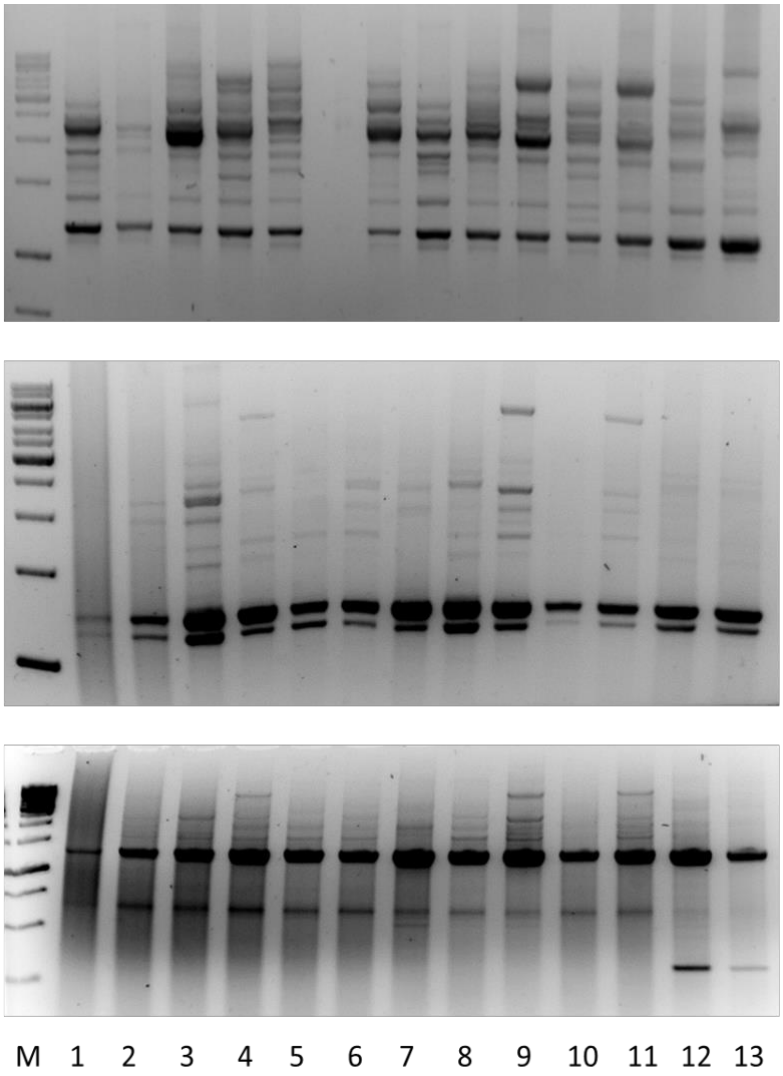

Fig. 4. Banding patterns of polymerase chain reaction amplified fragments of European plum cultivars using a combination of forward and reverse primers, SCoT2 and FaSt-Rev2, respectively. Labels are the followings: (M) 1-kb DNA ladder, (1) 'Toptaste', (2) 'Haroma', (3) Topend plus', (4) 'Haganta', (5) 'Presenta', (6) 'Tophit', (7) 'Topfive', (8) 'Empress', (9) 'Elena', (10) 'Hanita', (11) 'Jojo', (12) 'Nemtudom P3', and (13) 'Besztercei Bt2'.

Although minor differences appear among the gels with samples occasionally not providing any fragments, the characteristic fragments and alterations in banding patterns could be seen in each of the gels. The differences might be attributed to PCR efficiency and electrophoresis time.

For the analysis of marker efficiency, we considered only the polymorphic fragments, fragments that were not present in each and every accession. Such presence/absence data of 31 primer combinations were entered into the PAST 4.03 software for further analyses. A total of 211 amplicons were registered and the number of the amplified

fragments in different cultivars ranged between 43 ('Tophit') to 119 ('Nemtudom P3'). A heatmap indicating the presence of each of the 211 amplicons in the tested cultivars is shown in Fig. 5A, while a dendrogram was depicted to demonstrate genetic relationships among cultivars (Fig. 5B). 'Besztercei Bt.2' and 'Nemtudom P3' cultivars formed a separate clade with 100% bootstrap support.

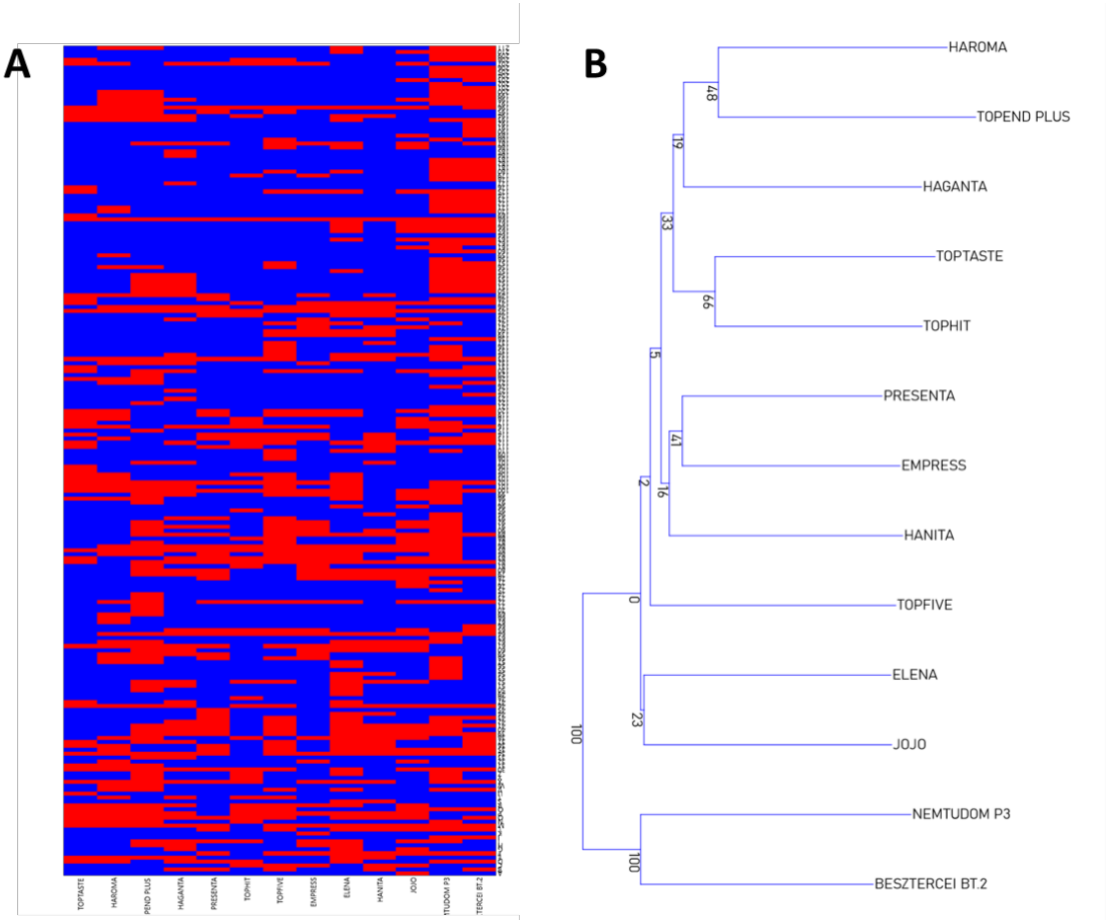

Fig. 5. Genetic variability and phylogenetic analysis of European plum cultivars using 36 SCoT and *FaSt* primer combinations. (A) Heatmap representation of the occurrence of analysed amplicons in cultivars (red: the allele is present, blue: the allele is absent). (B) Phylogenetic analysis indicated genetic relatedness among cultivars based on the Euclidean similarity index. Bootstrap values are given as percentages on or next to the branches.

Other clades receiving strong bootstrap support were the one containing all cultivars originated in a German breeding program and a subclade of 'Toptaste' and 'Tophit'. A PCA analysis was also carried out to verify the genetic relationships reflected by the dendrogram and provide further details. The first two principal axes accounted for 21.7% and 13.5% of the total variation. The PCA analysis confirmed that 'Besztercei Bt.2' and 'Nemtudom P3' were found in a more distant position than the rest of the tested cultivars (Fig. 6). PC1 separated sharply these cultivars from the ones originated in a German breeding program, which were differentiated along PC2.

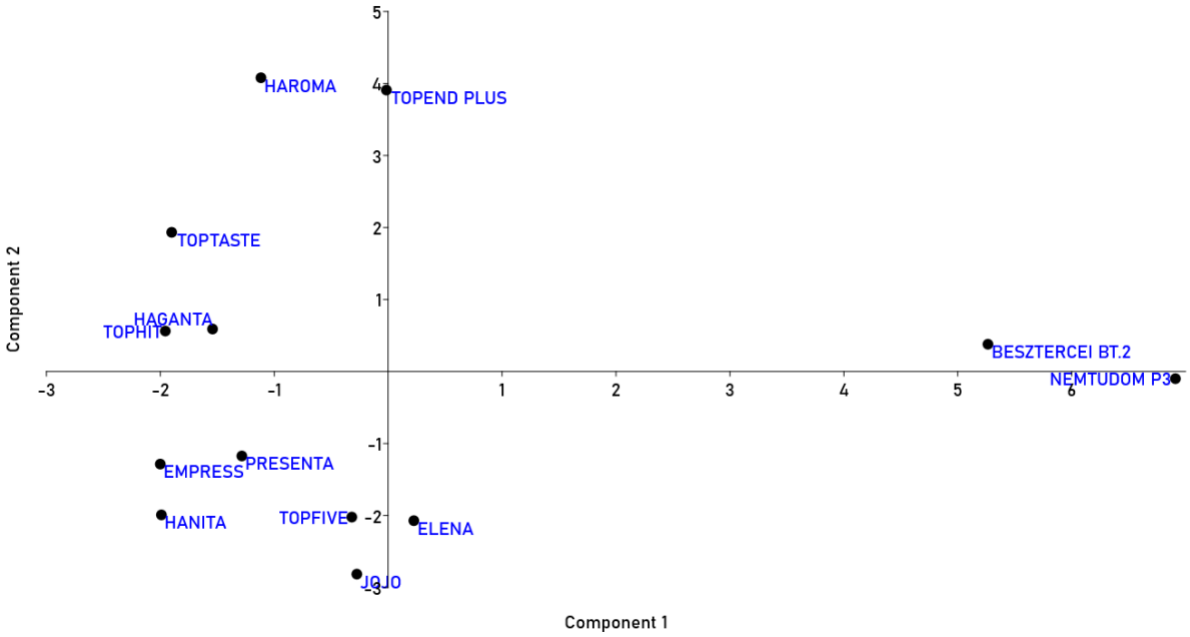

Fig. 6. Distribution of European plum cultivars on the two first principal component analysis axes determined from the amplification pattern of 31 SCoT and *FaSt* primer combinations.

6.2. Sequence analysis of fragments amplified by SCoT and *FaSt* primers

A total of 25 PCR amplified fragments were cloned and sequenced to get information on the genomic regions targeted by different combinations of SCoT and *FaSt* primers. Some of the sequenced amplicons are shown on the characteristic electropherograms (Fig. 7) but sequence analysis identified even many more fragments than could be detected due to the limited resolution power of agarose gels. The obtained sequences, their main characteristics and the results of their homology searches are shown in Table 2. We used a label for the sequences containing the SCoT marker number (1, 2, 4, 18, 22, 27), the abbreviation for the name of the cultivar from which the sequence was determined (BB: 'Besztercei Bt.2', EL: 'Elena', HR: 'Haroma', NT: 'Nemtudom P.3', TF: 'Topfive', TP: 'Topend Plus') as well as the size of sequence. The *FaSt* targeting FaSt-Rev2 primer was combined with the following SCoT primers (in brackets are the number of sequences): SCoT1 (5), SCoT2 (3), SCoT4 (3), SCoT18 (3), SCoT22 (9), and SCoT24 (2). A total of 25 sequences were determined and used for further analyses.

The preliminary check of the raw sequences identified the cloning vector flanking regions, and the insert sequence was kept for further characterization. We tried to identify the used *FaSt* and SCoT primer sequences, which was successful for 16 sequences, while only one of the two primers was found in the remaining 9 sequences (Table 2). In those cases, only the SCoT primers annealed (in some cases like we could find the primer sequence in both forward and reverse orientations at the two ends) and hence those fragments did not contain any identifiable segment of the *FaSt* transposon. Interesting to note that fragments amplified by only the *FaSt* primers and not containing the corresponding SCoT primer sequences were not found.

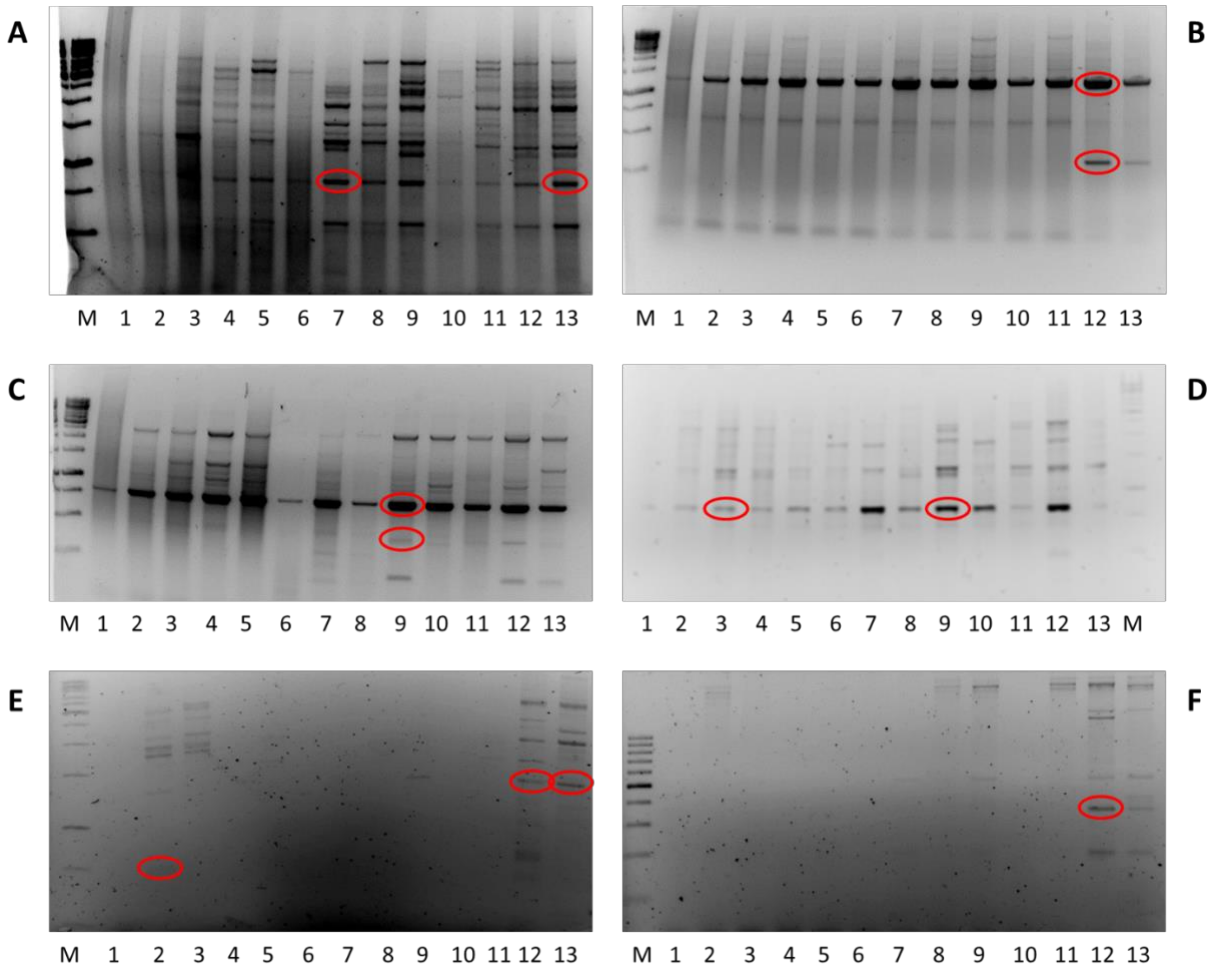

Fig. 7. Some sequenced fragments amplified in polymerase chain reaction of European plum cultivars using a combination of several SCoT primers and FaSt-Rev2: (A) SCoT1, (B) SCoT2, (C) SCoT4, (D) SCoT18, (E) SCoT22, and (F) SCoT27. Labels are the followings: (M) 1-kb DNA ladder, (1) 'Toptaste', (2) 'Haroma', (3) Topend plus', (4) 'Haganta', (5) 'Presenta', (6) 'Tophit', (7) 'Topfive', (8) 'Empress', (9) 'Elena', (10) 'Hanita', (11) 'Jojo', (12) 'Nemtudom P3', and (13) 'Besztercei Bt2'.

From the 16 sequences amplified by the combination of SCoT and *FaSt* primers the blast analysis did not reveal the presence of a recognizable part of the *FallingStones* transposon in the closest peach homolog of the 1-TF-947, 1-BB-948, and 4-EL-806 sequences (Appendix Table 1). When we carried out a blastn analysis on the available *P. domestica* genome, the *FaSt* motif was found in both the 1-TF-947 and 1-BB-948 sequences. It was located upstream of the coding region of the polygalacturonase ADPG2 gene, which was not present in the *P. persica* gene.

Most of the sequences (13 from 16) contained a part of the *FaSt* transposon. Besides a smaller or bigger fragment of the *FaSt*, many of those sequences (1-TF-841, 4-EL-473, 18-TP-834, 18-EL-835, 18-TP-843, 22-NT-374, 22-NT-1239, and 22-BB-1240) contained some intergenic regions without protein coding ability. In the case of 22-HR-295, the only matching part of the sequence was the *FaSt* while other segments did not align with any sequence in the

available *Prunus* and also *P. domestica* genomes, indicating this region was not included in the *P. domestica* genome sequence or occurring only in the tested 'Haroma' cultivar. A similar phenomenon was observed with 22-NT-374 containing the *FaSt* primer annealing site in double copies; however, only the inner core of the sequence could be aligned on the *P. domestica* genome, the region containing a partial *FaSt* element.

Using blastn analysis on the NCBI nucleotide database, only the *FaSt* part of the following sequences, 22-HR-457, 22-BB-486, and 22-BB-460 showed homology to *FaSts* in other species while blasting them on the *P. domestica* genome provided fully matching hits. Approx. half of the nucleotide positions in 22-BB-486, and 22-BB-460 sequences matched the *P. armeniaca FaSt* in the S-haplotype-specific F-box gene, but in *P. domestica* genome the coverage was full with 93% identity (449/481) and 1% gaps (6/481).

In addition, the 18-TP-834, 18-EL-835, 18-TP-843, 22-NT-1239, and 22-BB-1240 sequences contained homologous regions with both *P. dulcis* and *P. domestica,* but *P. domestica* homologs were uninterrupted or contained smaller gaps compared to those in almond genomic sequences. In *P. domestica*, the insertion is of 28 bp, while in *P. dulcis* it is 139 bp and locates close to an alpha/beta-hydrolases superfamily protein.

Table 2. Name, description, size and the results of the blastn analysis of SCoT-*FaSt* sequences.

Name	Description	Size (bp)	*E*-value	Identity (%)	Accession number	Identified regions[1]
SCoT1 and FaSt-Rev2						
1-TF-841	*Prunus dulcis* DNA, pseudomolecule Pd02	841	1e-165	91.0%	AP019298	*FaSt*, IG
1-TF-947	PREDICTED: *Prunus persica* polygalacturonase ADPG2, mRNA	947	0.0	95.9%	XM_020568329	gene
1-BB-948	PREDICTED: *Prunus persica* polygalacturonase ADPG2, mRNA	948	0.0	95.4%	XM_020568329	gene
SCoT4 and FaSt-Rev2						
4-EL-473	*Prunus dulcis* DNA, unplaced-scaffold_262	473	3e-89	88.0%	AP020599	*FaSt*, IG
4-EL-806	*Prunus mume* 7-deoxyloganetin glucosyltransferase-like, mRNA	806	0.0	98.0%	XM_008221234	partial FaSt
SCoT18 and FaSt-Rev2						
18-TP-834	*Prunus dulcis* DNA, pseudomolecule Pd08	834	4e-155	93.4%	AP019304	*FaSt*, IG, gaps
18-EL-835	*Prunus dulcis* DNA, pseudomolecule Pd08	835	4e-155	93.4%	AP019304	*FaSt*, IG, gaps
18-TP-843	*Prunus dulcis* DNA, pseudomolecule Pd08	843	9e-152	92.9%	AP019304	*FaSt*, IG, gaps

Name	Description	Size (bp)	*E*-value	Identity (%)	Accession number	Identified regions[1]
SCoT22 and FaSt-Rev2						
22-HR-295	*Prunus dulcis* DNA, pseudomolecule Pd08	295	1e-100	94.7%	AP019304	*FaSt*
22-NT-374	*Prunus dulcis* DNA, pseudomolecule Pd08	374	2e-75	91.28%	AP019304	nested *FaSt*
22-HR-457	*Prunus armeniaca* clone Armed-3 MITE *Falling Stone*, complete sequence	457	8e-100	96.1%	KF956794	*FaSt*
22-NT-895	*Prunus dulcis* ethylene-responsive transcription factor ERF053-like (LOC117623168), mRNA	895	0.0	97.1%	XM_034354046	gene, *FaSt* (3'UTR)
22-NT-1239	*Prunus dulcis* DNA, unplaced-scaffold_1226	1239	0.0	88.3%	AP021563	*FaSt*, IG, gaps
22-BB-1240	*Prunus dulcis* DNA, unplaced-scaffold_1226	1240	0.0	88.4%	AP021563	*FaSt*, IG, gaps
22-BB-486	*Prunus armeniaca* clone Armed-3 MITE *Falling Stone*, complete sequence	486	1e-98	95.7%	KF956794	*FaSt*
22-BB-460	*Prunus armeniaca* clone Armed-3 MITE *Falling Stone*, complete sequence	460	3e-104	97.0%	KF956794	*FaSt*

[1]*FaSt*, *Falling Stones* transposon, IG, intergenic regions, UTR, untranslated region.

Undoubtedly, the most interesting sequences were the four containing a fragment of identifiable genes, including 1-TF-947, 1-BB-948, 4-EL-806, and 22-NT-895. Two of the four sequences were homologous to *Prunus persica* polygalacturonase ADPG2 gene, while the *FaSt* element in the sequence aligned separately with several *P. persica FaSt* sequences. However, the blast analysis on the *P. domestica* genome showed 98% identity and two gaps on a 936 bp long alignment. The *FaSt* element located upstream of the start codon of the open reading frame (Fig. 8). The same sequence without major alterations were also identified in 'Topfive' and 'Besztercei Bt.2'.

The 4-EL-806 sequence showed homology to the *Prunus mume* 7-deoxyloganetin glucosyltransferase-like (LOC103319662) gene with an *E*-value of 0. The *FaSt* primer annealing site was identified but other regions of the *FaSt* were not found. The 22-NT-895 sequence was found to be closely related to the *Prunus dulcis* ethylene-responsive transcription factor ERF053-like protein encoding gene. The *FaSt* element could be found downstream of the stop codon, indicating its location in the 3' untranslated region.

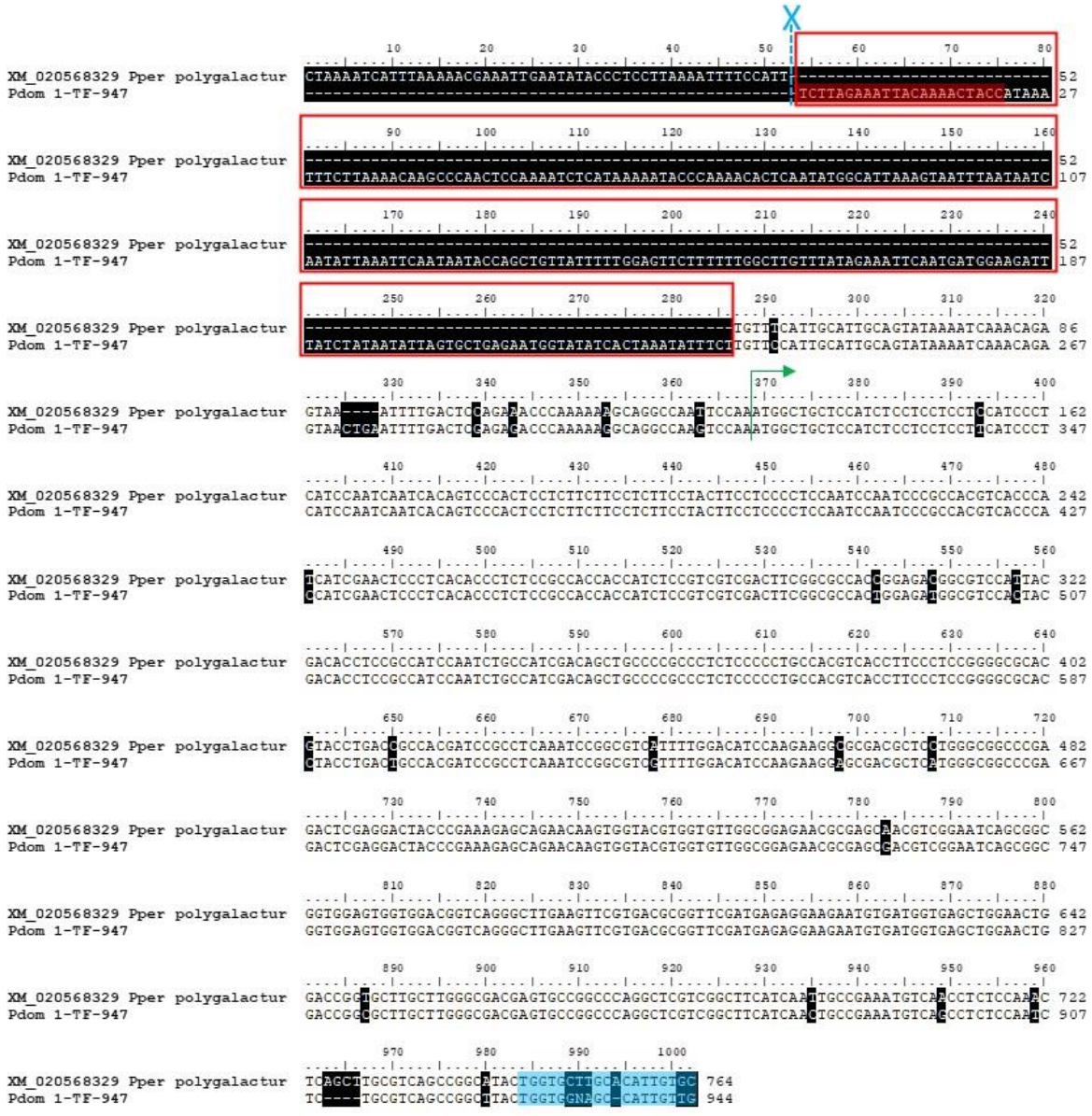

Fig. 8. The alignment of *Prunus domestica* polygalacturonase *ADPG2* and its *P. persica* homolog. The *FallingStones* motif is squared in red, the annealing sites of the SCoT1 and FaSt-Rev2 primers are indicated by red and blue highlighting. The blue cross shows a region where the alignment is interrupted, the green arrow refers to the initiation site and direction of translation. Differing positions are labelled by white characters against black background colour.

7. DISCUSSION

7.1. The evaluation of the SCoT-*FaSt* marker system

The number of amplified fragments and the percentage of polymorphic fragments varied according to the different SCoT-*FaSt* combinations. The ratio of polymorphic bands gives the most information on how much discrimination potential is provided by the marker. This value ranged between 32 and 100%. The most informative of all combinations were the SCoT15, SCoT9 and SCoT7 for the analysis of genetic diversity of European plum cultivars. The least informative primer was SCoT34 with small number of amplified fragments and a small polymorphism ratio. Buer et al. 2022 used SCoT markers to analyse the genetic diversity of *P. sibirica* populations in Mongolia and found the polymorphism percentages ranging between 90 to 100%. Ahmed et al. 2022 determined 89.7% polymorphic percentage for SCoT markers among 5 *Prunus* species including *P. domestica*. The elevated polymorphism detected by SCoT markers relative to other marker systems (RAPD, SSR, ISSR) was also confirmed by other studies (Ait et al. 2021; Thakur et al. 2022; Antanynienė et al. 2023). The SCoT and *FaSt* primer combinations can detect variable levels of polymorphisms and hence it may become an efficiently applied molecular marker in *Prunus* species.

Since RAPD markers are infamous of low repeatability, the original work reporting on the establishment of SCoT marker strategy (Collard and Mackill, 2009), a bit similar one-primer marker, was focused on assessing its repeatability. They concluded that the primer length of 18 nucleotides and annealing temperature of 50°C did not guarantee reproducibility for all tested primers while others were highly reproducible. We have carried out the PCRs in three independent repetitions for some of the primer combinations and could detect some minor alterations in the presence/absence of some fragments but all the major patterns were consistent. It suggests the reproducibility of the combination of SCoT and *FaSt* markers is likely to be not an issue although a careful consideration remains inevitable.

The phylogenetic analysis of European plum cultivars revealed two major clusters, one containing the cultivars from the German and US breeding programs and the other one encompassing the two cultivars of Hungarian origin. A traditional cultivar ‘Besztercei Bt.2’ and ‘Nemtudom P3’ is a clone of an autochthonous landrace which have been cultivated in Hungary for so long and was shown in many former studies to be genetically much differentiated from other European plum accessions (Makovics-Zsohár et al. 2017, Hegedűs et al. unpublished). It was further confirmed by the PCA analysis showing the Hungarian cultivars in a different position relative to the others.

The PCA graph showed an interesting arrangement. The cultivars ‘Empress’, ‘Hanita’, and ‘Presenta’ formed a close group. ‘Empress’ is a variety from the USA (Kovács et al. 2012) but it formed a common group with some cultivars from Germany. However, they are all offspring of ‘President’. The cultivars ‘Elenea’ and ‘Jojo’ were also positioned together on both the phylogenetic tree and PCA scatter plot and both cultivars have ‘Stanley’ among

their parents (Surányi and Erdős, 2006). Interestingly, Makovics-Zsohár et al. 2017 using SSR analysis could not detect the genetic relatedness between these cultivars in the neighbor-joining dendrogram based on Jaccard's indices. However, the Structure analysis from Bayesian clustering placed them in the same genetic group. It indicates our SCoT-*FaSt* marker analysis could provide a reliable result in relation to genetic relatedness among different cultivars. It is further corroborated by the fact that 'Haroma', which has 'Stanley' among its grandparents, was located in a more distant position compared to 'Elenea' and 'Jojo'. However, the grouping of other cultivars was not that clear-cut: 'Topend Plus', 'Haganta' and 'Topfive' are all descendents of 'Cacanska najbolja' but they were scattered along the axis of component 2. It might be explained by the fact that a limited number of cultivars were used in the breeding of the analysed cultivars and hence many of those share common parents. For example, 'Topfive' and 'Hanita' are descendents of 'Auerbacher' and hence 'Topfive' was positioned closer to 'Hanita', or 'Haganta' and 'Tophit' are offspring of 'Cacanska najbolja', which explains their close proximity in the PCA graph. However, another statistically supported subclade was formed within the German-American plum clade, which contained the cultivars 'Toptaste' and 'Tophit'. Both of those are originated in the Geisenheim breeding program, Germany but their reported pedigrees do not show common progenitors as 'Toptaste' is a hybrid of 'Valor' and 'Hauszwetsche', while 'Tophit' was obtained by a cross between 'Cacanska najbolja' and 'President' (Mészáros et al. 2015). Their close relationship within all the other plum cultivars from the same breeding program could not be explained by the results. Unfortunately, these two cultivars have never been included together in a phylogenetic analysis which should be necessary to understand if this subclade and a closer genetic relationship can be confirmed or not.

In conclusion, the combination of SCoT and *FaSt* primers resulted in an unexpectedly high resolution of genetic relationships among the analysed cultivars. The fact that many of such cultivars were originated from the same breeding program and are sharing common progenitors, indicates the reliable performance of such markers in the genetic analysis of European plum.

7.2. Sequence analysis of the SCoT-*FaSt* primer pairs

The theoretical foundation of combining the SCoT and *FaSt* specific primers comes from the fact that both markers were found to be associated with the coding part of plant genome. SCoT primers were designed from the conserved regions flanking the start codon of plant genes (Collard and Mackill, 2009) while *FallingStones* are *Prunus*-specific miniature inverted repeat transposable elements preferentially accumulating in the AT rich segments of the euchromatin regions of chromosomes (Halász et al. 2014). The most important question of the combined application of the two primers was whether they are able to amplify regions close to or within gene sequences.

Nine of the 25 sequences contained only the SCoT primer sequences while DNA sequence of *FaSt* primer was not detected. Since SCoT was described to use a single 18-mer primer, it was not unexpected that such primers

will be able to amplify intergenic regions between two genes in opposite directions (Collard and Mackill, 2009). However, the introduction of a *FaSt*-specific primer and its application in combination with SCoT primers may also amplify those regions where two genes are not that close or are in identical orientation but there is a *FaSt* around or within a gene. We could confirm that many of the combinations could have amplified more fragments than SCoT or *FaSt* primers alone (data not shown). It is also confirmed by the fact that most (64%) of the sequenced fragments were amplified by the combination of SCoT and *FaSt* primers.

The *E*-values of the homology searches on the NCBI nucleotide database were ranged from 2e-75 to 0, indicating that homologous regions were found in other *Prunus* species, mainly *P. dulcis* (10 sequences), *P. armeniaca* (3), *P. persica* (2) and *P. mume* (1). In seven of those sequences, besides *FaSt*, intergenic regions were identified with blastn on the NCBI nucleotide database, while four other sequences contained regions that could not have been identified by NCBI blastn analysis but most likely they contain intergenic regions. It was confirmed by blastn carried out on the *P. domestica* genome sequences. Those sequences were not identified to be related to any described genes, which points to the possibility that the original 18-bp SCoT markers might anneal to other regions of the genome, not exclusively being specific to start codon flanking regions under the applied conditions. We do not know about studies carried out to identify the efficiency of SCoT markers in the amplification of gene-related sequences.

Some sequences (22-HR-457, 22-BB-486, 22-BB-460, 18-TP-834, 18-EL-835, 18-TP-843, 22-NT-1239, and 22-BB-1240) were shown to have partial homologs in other genomes like *P. dulcis* and *P. persica* but full length alignments were only provided by blasting them on the *P. domestica* genome sequence. When insertions/deletions were present in the aligned sequences those were much smaller when aligned with the *P. domestica* sequences than those from *P. dulcis* and *P. persica*. It is consistent with the evolutionary relationships of those species (Badenes and Parfitt, 1995; Shi et al. 2013). Our results gave additional evidence on the major evolutionary changes and genomic rearrangements among different lineages of *Prunus*.

A sequence (22-NT-374) was shown to contain the *FaSt* primer sequence in double copy, which is consistent with this being a nested *FaSt* element. Similar nested *FaSt* copies were seen in *P. armeniaca* (Halász et al. 2014) and *P. domestica* (Gyuris, 2015). This is not unexpected as 10% of MITEs were described to be present in multimers or nested arrangements in rice genome (Jiang and Wessler, 2001). The blastn analysis on the *P. domestica* genome did not reveal this sequence, which might be explained by the not yet complete genome sequence of this hexaploid and heterozygous species or the sequence is uniquely present in the 'Nemtudom P3' cultivar and the related landrace germplasm.

Only four (16%) of the sequenced fragments contained a *FaSt* element in close proximity of a protein encoding gene. Two of those (1-TF-947 and 1-BB-948) contained the *FaSt* upstream of the start codon of a putative polygalacturonase ADPG2 gene, within its promoter region. This arrangement was only found in *P. domestica* while the homologous gene sequences in peach were free of any inserted *FaSt* copies. It has been shown that an almond C-repeat binding factor (CBF) gene contained a *FaSt* element in the promoter region (Ivanovska et al., 2022). Barros

et al. 2012 confirmed that PdCBF1 expressed during late summer while PdCBF2 was induced by autumn chilling temperatures. The differences in the regulation of PdCBF1 might be due to the presence of *FaSt* providing binding sites for a range of transcription factors, similarly to the mMoshan transposons in peach which can increase gene expression significantly (Wang et al., 2016). A functional test on the occasional differences in the *P. domestica* polygalacturonase ADPG2 gene function may worth further investigation.

The 4-EL-806 and 22-NT-895 sequences contained a *FaSt* related sequence in the 3' UTR of a 7-deoxyloganetin glucosyltransferase-like (LOC103319662) gene and a putative ethylene-responsive transcription factor encoding gene, respectively. The insertions do not disrupt the coding region of the genes but length variations like MITE insertions may have consequences as shown in other plant genes (Li et al. 2014; Vignesh et al. 2013). It is also important to consider that besides the *FaSt* primer sequence, other regions of *FaSt* could have not been identified in 4-EL-806. Halász et al. 2014 described 121 full and 60 additional copies having a fragment of more than 50% of the original length in peach genome but many more small deletion derivatives of *FaSt* were also noticed (Hegedűs A. personal communication). It indicates that *FaSt*-specific primers may find even more annealing sites in the *Prunus* genome than the copy number of full length or major fragments of *FaSt*, which can increase significantly the amplification potency of this marker system.

8. CONCLUSIONS

Our results confirmed the possibility of combining the SCoT and *FaSt*-specific primers in order to characterize intergenomic variations among commercial cultivars of the hexaploid *Prunus domestica*. The level of polymorphism detected by different SCoT primers combined with the *FaSt*-annealing primer showed great variations and hence a careful evaluation is required to change the best combinations. The level of polymorphisms detected was eligible to have an accurate demonstration of genetic relationships among cultivar groups sharing a parental genotype. However, deviations were also noted.

Sequence analysis revealed that SCoT markers alone were also able to amplify fragments and the presence of *FaSt* primers could increase the number of amplified fragments compared to the exclusively used SCoT primers. Most of the amplified fragments contained a *FaSt* element but not any regions of protein encoding genes. It indicates that SCoT primers are not strictly specific for annealing to start codon flanking regions in *Prunus* genome. Many small deletion derivatives of *FaSt* are spread over the *Prunus* chromosomes, which increases the amplification potential of the SCoT-*FaSt* marker system. Only a small number of fragments contained both *FaSt* elements and parts of a protein coding genes but those might be crucially important genomic regions when screening for genetic variants inducing phenotypic changes. In breeding programs, such genes might be used for designing specific markers to follow important traits.

Although further experiments are needed to check the repeatability of different primer combinations and understand the factors affecting the performance of this newly described molecular marker system, SCoT and *FaSt* combinations might be useful in future breeding programs of European plum and probably other stone fruit species.

9. SUMMARY

Many studies showed the biggest limitation in the breeding of tree species is the long vegetative period of trees. It can involve 4-5 years of growth to get specific traits manifested in the phenotype. In the meantime, molecular marker technique has been available to help reduce the cost of breeding. Many molecular markers are associated with the non-coding part of the genome which is useful in many ways but not following the inheritance of a significant phenotypic trait. So, our study is focusing to develop a new molecular marker by using a combination of PCR primers. The basic steps to establish this marker system consist of detecting polymorphism by using the *FallingStones* (*FaSt)* and Start Codon Targeted (SCoT) markers in a single PCR, choosing combinations of *FaSt* and SCoT, characterizing the information of *FaSt* and SCoT combinations, and identifying variations detected by different combinations of *FaSt* and SCoT primers.

The initial step was to test the polymorphism produced by the combination of *FaSt* and SCoT primers. During this step, 34 SCoT primers were used. The most polymorphic combination was the SCoT15 and FaSt-Rev2, while the less polymorphic was SCoT4-FaSt-Rev2. The ratio of polymorphic bands gives information on how much discrimination potential is provided by the marker. Only 5 SCoT primers did not amplify a single band in any cultivars while 3 combinations (containing one of the SCoT15, SCoT9, and SCoT7 markers) provided the most informative patterns. The least informative primer was SCoT34 with small number of amplified fragments and a small polymorphism ratio. While SCoT primers alone can amplify fragments, the presence of *FaSt* primers can increase the number of amplified fragments and potentially identify crucially important genomic regions for designing specific markers to follow important traits in a future breeding program.

Also, the reliability of amplified patterns should be considered, so the repeatability was checked. We conducted in three independent repetitions for some of the primer combinations and could detect some minor alterations in the presence/absence of some fragments, but all the major patterns were consistent. It suggests the reproducibility of the combination of SCoT and *FaSt* markers is likely to be not an issue.

For the analysis of marker efficiency, we considered only the polymorphic fragments, Only clear reproducible PCR bands detected on gels were scored as absent (0) or present (1) and entered into the PAST 4.03 software and a PCA analysis was also carried out to verify the genetic relationships reflected by the dendrogram and provide further details. 'Besztercei Bt.2' and 'Nemtudom P3' cultivars formed a separate clade with 100% bootstrap support based on a dendrogram that demonstrated the genetic relationships among cultivars. Further data of the PCA analysis confirmed that 'Besztercei Bt.2' and 'Nemtudom P3' were found in a more distant position

than the rest of the tested cultivars, confirming their well-known differentiation from international commercial cultivars.

Additionally, our results revealed that combinations of SCoT and *FaSt* primers could detected the cultivars 'Elena' and 'Jojo' were also positioned together on both the phylogenetic tree and PCA scatter plot and both cultivars have 'Stanley' among their parents while a formerly published SSR analysis could not detect the genetic relatedness between these cultivars. This marker system can accurately demonstrate genetic relationships among cultivar groups sharing a parental genotype, but evaluation is required for the best combinations of primers.

A total of 25 PCR amplified fragments were cloned and sequenced to see whether they are able to amplify regions close to or within gene sequences. Most of the sequences contained a part of the *FaSt* transposon. Besides a smaller or bigger fragment of the *FaSt*, many of those sequences contained some intergenic regions without protein coding ability. However, four of the sequenced fragments contained a *FaSt* element in close proximity of a protein encoding gene (in its promoter region or 3'UTR). Further analyses are required to check if those have any functional consequences, but the most important outcome of using SCoT and *FaSt* marker combinations might be the identification of such genomic regions.

This study evaluated the efficiency of using SCoT and *FallingStones* (*FaSt)* markers to analyze the genetic diversity of European plum cultivars. The results showed that the SCoT and *FaSt* primer combinations can detect variable levels of polymorphisms and provide an unexpectedly high resolution of genetic relationships among the analyzed cultivars. The combination of SCoT and *FaSt* primers mainly resulted in the amplication of intergenic regions and less frequently regions close to or within genes. The finding suggests that the SCoT and *FaSt* marker may become a promising molecular marker in *Prunus* species, though its efficiency should be further tested and increased.

10. ACKNOWLEDGEMENTS

I would like to express my deep gratitude to Prof. Attila Hegedűs, my research supervisor at the Department of Plant Biotechnology of the Institute of Genetics and Biotechnology (MATE), for his patient guidance, enthusiastic encouragement, and useful critiques of this research work. Without his kind direction this thesis would have been a little success. In every phase of project his super vision and guidance shaped this thesis to be completed perfectly.

I would like to thank Ms. Beti Ivanovska, for her advice, valuable support, and assistance in laboratory. Special thanks should be given to Ms. Melisa Chikhi and Dr. Júlia Halász, for their professional guidance in the laboratory. I would also like to extend my thanks to Plant Biotechnology Department for offering all the instruments and resources during the program. Finally, I also wish to thank Mr. Valerii Kaminskii and my family for their support and encouragement throughout my study.

This work was supported by the ELKH-MATE cooperation “Precision breeding for safe food raw materials” (Eötvös Lóránd Research Network) and NKFI K_128874 project.

11. BIBIOGRAPHY

Ahmed, A. R., Soliman, K. A., Al-Senosy, N. K. and Omar S. A. 2022. Genetic Polymorphisms of Four Egyptian Plant Species using some Molecular Markers Techniques RAPD, ISSR, and ScoT. In: Agricultural Chemistry and Biotchnology. 13(1): p.21-24.

Ait, Y., Bouda, S., Khachtib, Y. and Haddioui, A. 2021. Genetic variability of cultivated plum ('*Prunus domestica* L. and *Prunus salicin* Lindl) in Morocco assessed by ISSR markers. In: Australian Journal of Crop Science. 15(6): p.948-954.

Ali, S. H., Sinjare, D. Y. and Jubrael, J. M. 2015. Analysis of Plum (*Prunus domestica* L.) Genotypes of Duhok City Using AFLP Markers. In: International Journal of Bioinformatics and Biomedical Engineering. 1: p.4-69.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. 1990. Basic local alignment search tool. In: J Mol Biol. 215: p. 403-410.

Aradhya, M. K., Weeks, C. and Simon, C.J. 2004. Molecular characterization of variability and relationships among seven cultivated and selected wild species of *Prunus* L. using amplified fragment length polymorphism. In: Sci. Hort. 103: p.131-144.

Athanasiadis, I., Nikoloudakis, N. and Hagidimitriou, M. 2013. Genetic relatedness among cultivars of the Greek plum Germplasm. In: Not Bot Horti Agrobo. 41(2): p. 491-498.

Ayanoğlu, H., Bayazit, S., İnan, G., Bakır, M., Akpınar, A. E., Kazan, K. and Ergül, A. 2007. AFLP analysis of genetic diversity in Turkish green plum accessions (*Prunus cerasifera* L.) adapted to the Mediterranean region. In: Scientia Horticulturae. 114: p.263-267.

Badenes, M. L. and Parfitt, D. E. 1995. Phylogenetic relationships of cultivated *Prunus* species from analysis of chloroplast DNA. In: Theor. Appl. Genet. 80: p. 129-133.

Barros, P. M., Gonçalves, N., Nelson J. M., and Oliveira, M. M. 2012. Cold acclimation and floral development in almond bud break: insights into the regulatory pathways. In: Journal of Experimental Botany. 63(12): p.4585–4596.

Ben, T. H., Walker, D., Ben, M. S., Abdallah, D., Baraket, G., Salhi, H. A. and Zehdi, A. S. 2015. DNA variation and polymorphism in Tunisian plum species (*Prunus spp.*): contribution of flow cytometry and molecular markers. In: Genet Mol Res. 14(4): p.18034-18046.

Buer, H., Rula, S. and Wang, Z.Y. 2022. Analysis of genetic diversity in *Prunus sibirica* L. in inner Mongolia using SCoT molecular markers. In: Genet Resour Crop. 69: p.1057–1068.

Callahan, A. 2008 Plums. In: Kole C, Hall TC (eds) Compendium of transgenic crop plants: transgenic temperate fruits and nuts. 1st edition. Oxford. Blackwell Publishing. P.93–449.

Callahan, A. M., Zhebentyayeva, T. N., Humann, J. L., Saski, C. A., Galimba, K. D., Georgi, L. L., Scorza, R., Main, D. and Dardick, C. D. 2021. Defining the 'HoneySweet' insertion event utilizing NextGen sequencing and a de novo genome assembly of plum (*Prunus domestica*). In: Hortic Res. 8: p.8.

Chin, S. W., Lutz, S., Wen, J. and Potter, D. 2013. The bitter and the sweet: inference of homology and evolution of leaf glands in *Prunus* (*Rosaceae*) through anatomy, micromorphology, and ancestral–character state reconstruction. Int. J. In: Plant Sci. 174: p.27-46.

Collard, B.C.Y. and Mackill, D.J. 2009. Start Codon Targeted (SCoT) Polymorphism: A Simple, Novel DNA Marker Technique for Generating Gene-Targeted Markers in Plants. In: Plant Mol Biol. 27: p.86-93.

Decroocq, V., Hagen, L. S., Favé, M. G., Eyquard, J. P. and Pierronnet, A. 2004. Microsatellite markers in the hexaploid *Prunus domestica* species and parentage lineage of three European plum cultivars using nuclear and chloroplast simple-sequence repeats. In: Molecular Breeding. 13(2): p.135-142.

FAOSTAT. 2021. Food and Agricultural Organization Database Collection. Rome. Italy. http://faostat.fao.org

Gaši, F., Sehic, J., Grahic, J., Hjeltnes, S. H., Ordidge, M., Benedikova, D. and Nybom, H. 2020. Genetic assessment of the pomological classification of plum *Prunus domestica* L. accessions sampled across Europe. In: Genet Resour Crop. 67(5): p.1137-1161.

Gregor, D., Hartmann, W. and Stösser, R. 1994. Cultivar identification in *Prunus domestica* using randomly amplified polymorphic DNA markers. In: Acta Hortic. 359: p.33-40.

Gupta, P. K., Balyan, H. S., Sharma, P. C. and Ramesh, B. 1996. Microsatellites in plants: A new class of molecular markers. In Current Science. 70(1): p. 45–54.

Gyuris, B. 2015. A csonthéjas gyümölcsfajokra jellemző FaSt transzpozon variabilitása. Budapesti Corvinus Egyetem. manuscript.

Halász, J., Kodad, O. and Hegedűs, A. 2014. Identification of a recently active *Prunus*-specific non-autonomous Mutator element with considerable genome shaping force. In: Plant J. 79(2): p.220-231.

Hall, T.A. 1999. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. In: Nucleic Acids Symposium Series. 41: p.95-98.

Hamdani, A., Bouda, S., Houmanat, K., Outghouliast, H., Razouk, R., Adiba, A. and Charafi, J. 2022. Genetic diversity revealed via molecular analysis of moroccan and foreign plum (*Prunus domestica*; *Prunus salicina*) genotypes from an *ex-situ* collection. In: Vegetos. p.1-9.

Hammer, Ø., Harper, D. A. T. and Ryan, P. D. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. In: Palaeontologia Electronica. 4: p.9.

Hegedűs, A., Eusebio, G. P., Szendy, G., Ivanovska, B., Tóth, E.Gy. and Halász, J. 2023. The molecular characterization of traditional Hungarian cultivars of hexaploid plum (*Prunus domestica*) highlights their breeding value. Manuscript.

Hormaza, J. I. 2002. Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. In: Theoretical and Applied Genetics. 104: p.321-328.

Horvath, A., Balsemin, E., Barbot, J. C., Christmann, H., Manzano, G., Reynet, P. and Mariette, S. 2011. Phenotypic variability and genetic structure in plum (*Prunus domestica* L.), cherry plum (*P. cerasifera* Ehrh.) and sloe (*P. spinosa* L.). In: Scientia Horticulturae. 129(2): p.283-293.

Ivanovska, B., Hegedűs, A. and Halász, J. 2022. *In silico* analysis of *FallingStones* (*FaSt*) transposable element inserted in the promoter region of an almond CBF gene. In: Fodor, M., Bodor-Pesti, P., Deák, T. (eds.) Proceedings of János Lippay – Imre Ormos – Károly Vas (LOV) Scientific Meeting. Bp, Hungary, Magyar Agrár- és Élettudományi Egyetem Buda Campus. 814: p. 365-369.

Jason, Z. 2007. Turning Over a New Leaf Change from 'Prune' to 'Dried Plum' Proving Fruitful. http://web.archive.org/web/20071014213752/http://failuremag.com/arch_business_dried_plums.html.

Jiang, N. and Wessler, S. R.2001. Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements. In: Plant Cell. 13(11): p.2553-2564.

Kazija, D. H., Vujević, P., Jelačić, T., Milinović, B., Drkenda, P., Gaši, F., Kurtovic, M., Pejic, I., Simon, S., Zulj, M. and Nikolić, D. 2013. Genetic Identification of 'Bistrica'and Its Synonyms 'Požegača'and 'Hauszwetsche' (*Prunus domestica* L.) Using SSRs. In: Acta Hortic. 976: p. 285-289.

Kikuchi, K., Terauchi, K., Wada, M. and Hirano, H. Y. 2003. The plant MITE mPing is mobilized in anther culture. In: Nature. 421: p.167–170.

Kovács, S., Molnár, Á., Szenci, G. and Tóth, M. 2012. Commercial varieties of European plums grown in Hungary, a comparison of promising foreign varieties with the widely grown traditional ones. In: International Journal of Horticultural Science. 18(2): p.15–22.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. In: Mol Biol Evol. 33: p.1870-1874.

Lecouls, A. C., Bergougnoux, V. and Rubio M. J. 2004. Marker-assisted selection for the wide-spectrum resistance to root-knot nematodes conferred by the Ma gene from Myrobalan plum (*Prunus cerasifera*) in interspecific *Prunus* material. In: Molecular Breeding. 13: p.113–124.

Li, W., Liu, W., Wei, H., He, Q., Chen, J., Zhang, B. and Zhu, S. 2014. Species-specific expansion and molecular evolution of the 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) gene family in plants. In: PloS one. 9(4): p.e94172.

Liu, W. S., Liu, D. C., Feng, C. J., Zhang, A. M. and Li, S. H. 2006. Genetic diversity and phylogenetic relationships in plum germplasm resources revealed by RAPD markers. In: The Journal of Horticultural Science and Biotechnology.81(2): p.242-250.

Llgin, M., Kafkas, S. and Ercisli, S. 2009. Molecular Characterization of Plum Cultivars by AFLP Markers. In: Biotechnology & Biotechnological Equipment. 23(2): p.1189-1193.

Makovics-Zsohár, N., Tóth, M., Surányi, D., Kovács, S., Hegedűs, A. and Halász, J. 2017. Simple Sequence Repeat Markers Reveal Hungarian Plum (*Prunus domestica* L.) Germplasm as a Valuable Gene Resource. In: HortScience. 52(12): 1655-1660.

Manco, R., Basile, B., Capuozzo, C., Scognamiglio, P., Forlani, M., Rao, R. and Corrado, G. 2019. Molecular and phenotypic diversity of traditional European plum (*Prunus domestica* L.) germplasm of Southern Italy. In: Sustainability. 11(15): p.4112.

Manisha, T., Vishal, S. and Anjali, C., 2021. Genetic fidelity assessment of long term *in vitro* shoot cultures and regenerated plants in Japanese plum cvs Santa Rosa and Frontier through RAPD, ISSR and SCoT markers. In: South African Journal of Botany. 140: p.428-433.

Meland, M., Frøynes, O., Fotiric Akšić, M., Pojskić, N., Kalamujić Stroil, B.K., Lasic, L. and Gasi, F. 2020. Identifying Pollen Donors and Success Rate of Individual Pollinizers in European Plum (*Prunus domestica* L.) Using Microsatellite Markers. In: Agronomy. 10: p.264.

Mészáros, M., Kosina, J., Laňar, L. and Náměstek, J. 2015. Long-term evaluation of growth and yield of Stanley and Cacanska lepotica plum cultivars on selected rootstocks. In: Horticultural Science: Czech Academy of Agricultural Sciences. 42(1): p.22-28.

Neumüller, M. 2011. Fundamental and applied aspects of plum (*Prunus domestica*) breeding. In: Fruit Veg Cereal Sci Biotechnol. 5: p.139–156.

Ortiz, A., Renaud, R., Calzada, I. and Ritter, E. 1997. Analysis of plum cultivars with RAPD markers. In: Journal of Horticultural Science. 72(1): p.1-9.

Petri, C., and Burgos, L. 2005. Transformation of fruit trees. useful breeding tool or continued future prospect. In: Transgenic Res. 14: p.15–26.

Petri, C. and Scorza, R. 2008. Peach. In: Kole C, Hall TC (eds) A compendium of transgenic crop plants: temperate fruits and nuts. Oxford. Blackwell Publishing. p.79–92.

Petri, C., Alburquerque, N. and Faize, M. 2018. Current achievements and future directions in genetic engineering of European plum (*Prunus domestica* L.). In: Transgenic Res. 27: p.225–240.

Ramming, D.W. and Cociu, V., 1990. Plums (*Prunus*) 233 287 Moore JN & Ballington JR. In: Genetic resources of temperate fruit and nut crops. 290: p.1-63.

Reddy, M. P., Sarla, N. and Siddiq, E. A. 2002, Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. In: Euphytica. 128: p.9–17.

Rehder, A. 1949. Manual of Cultivated Trees and Shrubs Hardy in North America Exclusive of the Subtropical and Warmer Temperate Regions, 2nd edition. New York. MacMillan.

Rupasinghe, H. P. V., Jayasankar, S. and Lay, W. 2006. Variation in total phenolics and antioxidant capacity among European plum genotypes. In: Scientia Horticulturae. 108: p.243–246.

Scorza, R., Ravelonandro, M., Callahan, A. M., Cordts, J. M., Fuchs, M., Dunez, J. and Gonsalves, D. 1994. Transgenic plums (*Prunus domestica* L.) express the *plum pox virus* coat protein gene. In: Plant Cell Reports. 14(1): p.18–22.

Senková, S., Žiarovská, J., Fernández, E., Bošeľová, D. and Bežo, M. 2017. IRAP variability in *Prunus domestica* (L.) borkh based on Cassandra retrotransposon polymorphism. In: Genetika. 49(3): p.989-1000.

Shi, S., Li, J., Sun, J. Yu, J. and Zhou, S. 2013. Phylogeny and classification of *Prunus sensu lato* (*Rosaceae*). In: J. Integr. Pl. Biol. 55(11): p.1069-1079.

Shimada, T., Hayama, H. and Haji, T. 1999. Genetic diversity of plums characterized by random amplified polymorphic DNA (RAPD) analysis. In: Euphytica. 109: p.143–147.

Shirasawa, K., Hirakawa, H., Tabata, S., Hasegawa, M., Kiyoshima, H., Suzuki, S., Sasamoto, S., Watanabe, A., Fujishiro, T. and Isobe, S. 2012. Characterization of active miniature inverted-repeat transposable elements in the peanut genome. In: Theor. Appl. Genet. 124: p.1429–1438.

Sottile, F., Caltagirone, C., Giacalone, G., Peano, C. and Barone, E. 2022. Unlocking plum genetic potential: where are we at?. In: Horticulturae. 8: p.128.

Sun, Q., Liao, K., Geng, W., Liu, J., Nasir, M., Liu, H. and Cao, Q. 2015. Analysis of Genetic Relationships in *Prunus domestica* L. in Xinjiang Using ISSR Markers. In: Agricultural Science & Technology. 16(3): p.447.

Surányi, D. and Erdôs, Z. 2006. Fontosabb szilvafajták. In: Surányi, D.: Szilva. Mezôgazda Kiadó. Budapest.

Thakur, M., Luharch, R. and Sharma, V. 2022. Molecular markers based genetic relatedness studies in tissue culture propagated japanese plum cultivars Santa Rosa and Frontier. In: Genetic Resources and Crop Evolution. 69: p. 567-575.

Tomás-Barberán, F. A., Gil, M. I., Cremin, P., Waterhouse, A. L., Hess-Pierce, B. and Kader, A. A. 2001. HPLC-DAD- ESIMS analysis of phenolic compounds in nectarines, peaches, and plums. In: Journal of Agricultural and Food Chemistry. 49: p.4748–4760.

Urrestarazu, J., Errea, P., Miranda, C., Santesteban, L. G. and Pina, A. 2018. Genetic diversity of Spanish *Prunus domestica* L. germplasm reveals a complex genetic structure underlying. In: PloS one. 13(4).

Vignesh, K. S., Landero Figueroa, J. A., Porollo, A., Caruso, J. A. and Deepe, G. S. 2013. Granulocyte Macrophage-Colony Stimulating Factor Induced Zn Sequestration Enhances Macrophage Superoxide and Limits Intracellular Pathogen Survival. In: Immunity. 39(4): p.697-710.

Vos, P., Hogers, L., bleeker, M., Van, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. In: Nucleic Acids Res., 23(21): p.4407-4414.

Wang, X., Kohalmi, S. E., Svircev, A., Wang, A., Sanfaçon, H. and Tian, L. 2013. Silencing of the Host Factor eIF(iso)4E Gene Confers *Plum Pox Virus* Resistance in Plum. In: PLoS ONE. 8(1).

Wang, L., Pan, L., Yao, M., Cai, Y., Dong, Z. and Yao, D. 2016. Expression of oncofetal antigen glypican-3 associates significantly with poor prognosis in HBV-related hepatocellular carcinoma. In: Oncotarget. 7(27): p.42150-42158.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. In: Nucleic acids research. 18(22): p.6531-6535.

Xuan, H., Ding, Y., Spann, D., Möller, O., Büchele, M. and Neumüller, M. 2011. Microsatellite markers (SRR) as a tool to assist in identification of European plum (*Prunus domestica*). In: Acta Hortic. 918: p.689-692.

Yu, M., Chu, J., Ma, R., Shen, Z. and Fang, J. 2013. A novel strategy for the identification of 73 *Prunus domestica* cultivars using random amplified polymorphic DNA (RAPD) markers. In: Afr. J. Agric. Res. 8: p.243-250.

Zhang, X., Feschotte, C., Zhang, Q., Jiang, N., Eggleston, W.B. and Wessler, S.R. 2001. P instability factor: an active maize transposon system associated with the amplification of Tourist-like MITEs and a new superfamily of transposases. In: Proc. Natl Acad. Sci. USA. 98: p.12572–12577.

Zhebentyayeva, T., Shankar, V. and Scorza, R. 2019. Genetic characterization of worldwide *Prunus domestica* (plum) germplasm using sequence-based genotyping. In: Hortic Res. 6: p.12.

12. APPENDIX

Table 1. Name, description, size, and the results of the blastn analysis of SCoT-*FaSt* sequences containing only the SCoT primer sequence without *FaSt*.

Label	Description	Size (bp)	*E*-value	Identity (%)	Accession number
1-BB-883a	PREDICTED: *Prunus mume* aspartyl protease family protein At5g10770-like (LOC103323515), mRNA	883	0.0	95.5%	XM_008225508.2
1-BB-883b	PREDICTED: *Prunus mume* aspartyl protease family protein At5g10770-like (LOC103323515), mRNA	883	0.0	95.5%	XM_008225508.2 protein coding gene and 3' UTR
2-NT-907	*Prunus dulcis* UDP-glycosyltransferase (UGT94AF3) mRNA, complete cds	907	0.0	98.2%	MH969427
2-NT-787	*Prunus dulcis* UDP-glycosyltransferase (UGT94AF3) mRNA, complete cds	787	0.0	98.4%	MH969427
2-NT-337	*Prunus dulcis* DNA, pseudomolecule Pd05	337	4e-67	92.6%	AP019301
4-EL-889	*Prunus dulcis* DNA, unplaced-scaffold_1280	889	0.0	94.3%	AP021617
22-NT-382	PREDICTED: *Prunus avium* uncharacterized LOC110769629 (LOC110769629), mRNA	382	5e-106	86.0%	XM_021973657
27-NT-205	*Prunus dulcis* DNA, pseudomolecule Pd03	205	2e-52	86.9%	AP019299
27-NT-210	*Prunus dulcis* DNA, pseudomolecule Pd02 *Prunus avium* protein ELC-like (LOC110770636), transcript variant X8, mRNA	210	3e-86 5e-49	97.9% 90.3%	AP019298 XM_021974828

DECLARATION

on authenticity and public assess of final master's thesis/portfolio

Student's name: Thanyarat Onlamun

Student's Neptun ID: OBNAUQ

Title of the document: Designing a new set of molecular markers to reveal polymorphism in the coding region of European plum genome

Year of publication: 2023

Department: Plant Biotechnology

I declare that the submitted final master's thesis is my own, original individual creation. Any parts taken from an another author's work are clearly marked, and listed in the table of contents.

If the statements above are not true, I acknowledge that the Final examination board excludes me from participation in the final exam, and I am only allowed to take final exam if I submit another final master's thesis.

Viewing and printing my submitted work in a PDF format is permitted. However, the modification of my submitted work shall not be permitted.

I acknowledge that the rules on Intellectual Property Management of Hungarian University of Agriculture and Life Sciences shall apply to my work as an intellectual property.

I acknowledge that the electric version of my work is uploaded to the repository system of the Hungarian University of Agriculture and Life Sciences.

Place and date: Budapest, _2023_ year _04_month _28_day

Student's signature

STATEMENT ON CONSULTATION PRACTICES

As a supervisor of Thanyaratr Onlamun OBNAUQ, I here declare that the final master's thesis has been reviewed by me, the student was informed about the requirements of literary sources management and its legal and ethical rules.

I recommend/don't recommend the final master's thesis to be defended in a final exam.

The document contains state secrets or professional secrets: yes no

Place and date: Budapest,_2023_year _04_ month_28_day

Internal supervisor