ECPGR Activity Grant Scheme – Fifth Call, 2017

Activity Report

Testing, Use and Alignment of genetic data to distinguish unique and characterized accessions in Prunus (Prunus Alignment)

1 January 2018 - 30 April 2020

Matthew Ordidge, Gunārs Lācis, Marine Delmas, Monika Höfer, Pavlina Drogoudi, Stein Harald Hjeltnes, Hedi Kaldmäe, Torben Toldam-Andersen, Zsuzsanna Békefi, Felicidad Fernández Fernández, Gabriella Sonnante, Hilde Nybom, Dunja Kazija, Daniela Giovannini

Participants in the "Prunus Alignment" meeting in Athens © P. Drogoudi

January 2021

Testing, Use and Alignment of genetic data to distinguish unique and characterized accessions in Prunus (Prunus Alignment)

Activity Report

CONTENTS

Testing, Use and Alignment of genetic data to distinguish unique and characterized accessions in Prunus (Prunus Alignment)

Activity Report

INTRODUCTION

Previous ECPGR projects (PRUNDOC and EU.CHERRY) supported the genotyping of sets of plum (38) and sweet cherry (212) accessions to identify unique material for inclusion within AEGIS. The accessions were also characterised using common and harmonized first priority descriptors (FPD). Data were submitted to the Prunus DB manager and prepared for inclusion in EURISCO. A complementary project within the Malus/Pyrus WG (Pomefruit – C&E) has recently built upon equivalent background to utilise existing data, networks, and expertise in bringing together genotypic (SSR) data from multiple national sources to expand on the ECPGR European data base.

PRUNDOC identified, and characterized (with FPD) 57 further accessions which were unable to be genotyped (due to funding limitations) and classified these as, 'in principle' most appropriate accessions (MAAs), subject to a confirmation of uniqueness by subsequent genotyping; 39 of these were held by partners in Prunus Alignment. The EU.CHERRY project widened the SSR data coverage through the inclusion of a series of 112 accessions submitted (and co-funded) through a COST proposal, thus creating a useful base of aligned cherry data from numerous collections. Furthermore, a recent study published by PRUNDOC member Pavlina Drogoudi and colleagues highlighted a potential technique to speed up and improve the efficiency of genotyping in plums.

Together, these previous efforts provided a base on which to further develop our understanding of ECPGR Prunus germplasm and further improve the representation of Prunus within AEGIS and data within EURISCO.

The project aimed to firstly confirm, by conventional SSR genotyping, a series of plum accessions suitable for inclusion in AEGIS (the 'in principle' MAAs) for the majority of which, FPD and second priority descriptor (SPD) data would be available. It also aimed to identify a further set of 40-50 complementary plum accessions with FPD for subsequent inclusion in AEGIS from partner countries that were not included in PRUNDOC and a subset of approximately 20 of these were to be checked for uniqueness by conventional SSR genotyping. The project subsequently aimed to make recommendations on future plum genotyping techniques through a consideration of the use of HRM and the reproducibility of the technique between labs. Lastly, the project aimed to amalgamate a further expanded cherry SSR dataset (based on an alignment of national datasets to the EU.CHERRY dataset) to allow a wider comparison of material for the identification of further unique accessions for inclusion in AEGIS in the future.

Partners 1-6 were able to supply samples from plum accessions identified, but not previously genotyped, in PRUNDOC for analysis (the 'in principle' MAAs) and, with the exception of partner 5 (who was unable due to ongoing regeneration activity) each partner had committed to supplying SPD data for these accessions. Partners 7-10 were able to offer novel plum germplasm to expand upon the PRUNDOC set and were able to commit to supplying FPD data and leaf samples for genotyping in the project.

HAO Demeter (partner 5) was additionally involved in the previous development and testing of a high resolution melt (HRM) based technique for genotyping plums (Merkouropoulos et al., 2017). The main role of partner 5 was therefore to lead in the HRM analysis of plum accessions, in addition to supplying samples for the conventional SSR analysis of accessions outstanding from PRUNDOC. Partners 1,2,8 and 12 were able to offer testing capability for the analysis of the HRM technique across multiple labs.

The expertise and experience of SLU (partner 13) was integral to the SSR genotyping of plums in PRUNDOC and their inclusion as a partner was deemed to be crucial to align with this previous effort.

The main role of partner 13 was therefore to carry out the SSR based genotyping of plum samples; partner 13 additionally supplied cherry SSR data and samples for alignment.

NIAB-EMR (partner 11) provided the SSR genotyping service and expertise for genotyping of cherry samples in EU.CHERRY and their inclusion in this project, to analyse a series of overlapping accessions that would allow alignment of the national datasets, was essential to allow the best alignment of available datasets with the EU.CHERRY data.

Partners 1, 4 and 8 (in addition to partner 13) were able to supply existing cherry SSR data screened with recommended ECPGR (or EU.CHERRY) markers from national datasets for alignment and inclusion in the ECPGR Prunus DB. Each partner was also able to supply leaf samples to partner 11 for replicated analysis to aid with alignment.

Additionally, collaborative agreements were made with Stephanie Mariette (INRA, France) and Markus Kellerhals (Agroscope, Switzerland) in an attempt to align further SSR data for cherry, again produced under national projects within ECPGR member countries.

The project was joined by additional self-funded partners from Croatia, Belgium and the Czech Republic who had either committed to collecting FPD data on self-nominated plum accessions (Croatia), or were interested to act as observers within the project consortium.

MATERIALS AND METHODS

Conventional SSR analysis and associated characterization of plum accessions

Leaf samples from the accessions identified, but not previously genotyped, in PRUNDOC (the 'in principle' MAAs) were supplied by partners 1-6 for genetic analysis by partner 13; in addition, further plum samples were supplied for SSR analysis to expand on the PRUNDOC set by partners 7-10. In total 64 samples were analysed (41 from former PRUNDOC partners [1-6] and 23 from the 'new' partners [7-10]). All plum samples analysed by conventional SSR within the frame of the project are listed in Annex I.

SSR analysis was carried out by partner 13 as described in Gasi et al. (2020).

Data for both FPD and SPD for all of the newly genotyped accessions, in addition to the accessions genotyped previously during PRUNDOC, were supplied to partner 13 and these were used for the analysis presented in Gasi et al. (2020). All characterization data are available in the supplementary material of the publication. Additional accessions, beyond those that it was possible to genotype, were also characterized using the same descriptors.

Analysis of plum accessions by HRM and comparison between labs

Following the analysis of the samples tested by conventional SSR, a subset of 56 samples was selected in order to provide a representative range of the diversity to be reproduced using the HRM technique by partner 5. This subset was selected by the project co-ordinator, with the aid of partner 13 in relation to maintaining alignment with the results of the conventional SSR analysis. Samples were also selected in an attempt to minimise the number of individual partners who would have to collect and send material (for reasons of practicality). A sub-subset of 36 accessions was identified with the intention that these samples would also be supplied to each of partners 1, 2, 8 and 12 with the aim of assessing the

reproducibility of the technique across a set of different laboratories. All samples were collected, lyophilised and sent to each of the respective partners. A full list of samples is included in Annex I.

HRM analysis was carried out according to a standard protocol supplied by partner 5 and based on Merkouropoulos et al. (2017). Isolation of DNA from leaves was performed using the 'NucleoSpin Plant II' kit (Macherey-Nagel, Germany), according to the manufacturer's instructions. DNA concentration was measured using a NanoDrop-1000 (Thermo Scientific, Wilmington, DE, USA) spectrophotometer at 260 and 280 nm ultraviolet lengths, whereas the integrity of the DNAs was estimated by gel electrophoresis in a 0.8% agarose gel. Samples were then diluted to 20 ng/μl working concentration.

For microsatellite analysis, PCR amplification, DNA melting and end point fluorescence level acquisitioning were performed in a total volume of 20 μl in a 72-well carousel of the Rotor-Gene 6000 real-time 5P HRM PCR Thermocycler (Corbett Research, Sydney, Australia) according to Ganopoulos et al. (2011). PCR reaction mixture consisted of 20 ng genomic DNA, 1× PCR buffer, 2.5 mM MgCl2, 0.2 mM dNTP, 300 nM forward and reverse primers, 1.5 mM Syto®9 green fluorescent nucleic acid stain and 1 U KapaTaq DNA polymerase (Kapa Biosystems, USA). The PCR protocol used was as follows: an initial denaturing step at 95°C for 3 min followed by 35 cycles at 95°C for 20 s, 52°C for 30 s, and 72°C for 20 s, then a final extension step at 72°C for 2 min. The fluorescent data were acquired at the end of each extension step during PCR cycles. In order to perform the HRM analysis, the products were initially denatured at 95°C for 5 s, and then annealed at 50°C for 30 s to randomly form DNA duplexes.

HRM protocol steps included: pre-melt at the first appropriate temperature for 90 s, and melt at a ramp of 10°C in an appropriate temperature range at 0.1°C increments every 2 s. The fluorescent data were acquired at the end of each increment step. End point fluorescence level was acquired following the melting process by holding at 60°C for 5 min. The sequences of the microsatellite primers used in the study were chosen from the available published literature (Cipriani et al., 1999; Downey and Iezzoni, 2000; Dirlewanger et al., 2002; Struss et al., 2003; Messina et al., 2004), whereas HRM was performed as described previously (Ganopoulos et al., 2011; Xanthopoulou et al., 2017). For microsatellite genotyping by HRM analysis, the genotype of each DNA sample was determined based on the shape of curves depicted by temperature-shifted melting curves or difference plots, and was scored for the binary data matrix (e.g. '1' denoted the presence of a melting curve, whereas '0' denoted the absence of a melting curve). Briefly for each marker we assign 1 if the HRM curve exists and 0 if it does not by comparing horizontally one genotype with all the others. Thus, the HRM curves become markers, allowing the performance of genotyping via the HRM analysis. The matrices were then analysed by FreeTree v. 0.9.1.50 software (Hampl et al., 2001). Similarity of qualitative data was calculated using the Nei and Li/Dice similarity index (Nei and Li, 1979), and similarity estimates were analysed using NJ (Neighbor Joining). The matrices of mutual coefficients of similarity calculated by FreeTree were converted to MEGA 5 software (Tamura et al., 2011) and the resulting clusters were expressed as dendrograms.

Modifications within each laboratory included the use of differing DNA extraction kits and differing real-time PCR machines. These variables provided a realistic representation of the range of equipment and consumables that would be likely to be available to working group members.

SSR analysis of cherry accessions and alignment of national datasets with EU.Cherry

National datasets of SSR data for cherry were supplied by partners 1, 4 and 13 to partner 8. Datasets were also supplied by the collaborating partners in France and Switzerland. Following an initial attempt to align entries in the national datasets against data generated in EU.Cherry (based on accessions that were analysed in both studies) and an initial analysis of the allele coverage within EU.Cherry samples,

a set of accessions was identified and leaf samples were supplied from each of these by the respective partners to partner 11 for genotyping. The set of accessions was selected in an attempt to represent as many of the alleles as possible from the national datasets (and specifically those which had not already been reproduced in EU.Cherry). Some adjustment was required due to the distribution of samples within the partner country. A number of additional samples, predominantly a series of samples from partner 4 were included where initial attempts at alignment against EU.Cherry data appeared to reveal potential errors in data at either end. In total, 113 samples were identified for collection, although not all of them were able to be collected.

A total of 99 leaf samples were received by partner 11 and these were analysed alongside the standard reference samples (107 samples analysed in total). Leaf discs were preserved on silica gel. The full list of samples can be found in 'Annex 1 – Sample Register' of the partner 11 project report.

DNA extractions were performed in a 96-well plate format as described for EUCHERRY, using the protocol described by Edge-Garza et al 2014 with the following modifications: polyvinylpirrolidone (PVP) was substituted for polyvinylpolypirrolidone (PVPP); 5M sodium chloride was substituted for 6M ammonium acetate; precipitation of DNA at -20°C was performed overnight. DNA pellets were re-suspended in 10mM Tris-HCl pH 8.0 and diluted to 5 ng/µl for use in the polymerase chain reaction (PCR) genotyping.

Initial PCR reactions were carried out using the same dye-labelled primer combinations and multiplexes as reported in the EUCHERRY project (Table 1). PCR reactions were performed in 13 µl reaction volumes using the Type-it™ Microsatellite PCR kit (Qiagen) according to manufacturer's protocol, with 2 µM of each primer. Thermal cycling was carried out in a Veriti 96 well thermal cycler (Applied Biosystems) as follows: initial 5 min denaturation at 95°C; 10 touchdown cycles comprising a 30 s denaturation step at 95°C, followed by 90 s of annealing starting at 55°C in the first cycle and decreasing 0.5°C per cycle, and 30 s of extension at 72°C; 20 cycles of 30 s at 95°C, 90 s at 50°C and 30 s at 72°C; and a final extension step of 30 minutes at 60°C.

Table 1. SSR markers used for genotyping; organisation into multiplex (MP), fluorescent label (Dye), linkage group (LG), and observed allele ranges are shown.

MP	Dye	Marker name	LG	Alelle range	Reference
A	FAM	EMPa002	1	105-133	Clark and Tobutt 2009
A	FAM	CPSCT038[1]	2	184-217	Rosyara et al 2013
A	FAM	CPPCT022	7	227-283	Clark and Tobutt 2009
A	NED	BPPCT034[1]	2	206-287	Rosyara et al 2013
A	HEX	CPPCT006	8	173-205	Clark and Tobutt 2009
A	HEX	PAV-Rf-SSR[2]	3	304-363	Sandefur et al 2016
A	PET	EMPaS02	3	132-192	Clark and Tobutt 2009
B	FAM	BPPCT037	5	121-170	Clark and Tobutt 2009
B	FAM	EMPaS06	4	200-235	Clark and Tobutt 2009
B	PET	EMPa004	6	158-212	Clark and Tobutt 2009
B	PET	EMPa017	2	221-250	Clark and Tobutt 2009
B	PET	EMPa018	8	83-130	Clark and Tobutt 2009
B	HEX	EMPaS12	3	102-157	Clark and Tobutt 2009
B	HEX	EMPaS14	5	168-213	Clark and Tobutt 2009

[1] Markers linked to fruit size
[2] Marker linked to flesh colour

In order to harmonize allele scores and allow comparisons of datasets from project partners, the ECPGR reference genotypes P. avium F12/1, P. avium 'Goodnestone Black', P. avium 'Napoleon', P. avium 'Noble', P. avium 'Noir de Meched', P. incisa E621, P. mahaleb SL64, and P. nipponica F1292 (Clarke and Tobutt 2009) were included in each plate.

Following PCR amplification, fluorescently-labelled products were diluted (1:10) and 1.3 µl used for electrophoresis in the ABI 3130xl Genetic Analyser (Applied Biosystems). Data was collected and allele sizes determined using GENESCAN and GENOTYPER software applications (Applied Biosystems). Estimated allele sizes were rounded and harmonized across plates according to the profiles obtained for the ECPGR standards which had been included in the EUCHERRY analysis ('Napoleon', 'Noble', and/or F12/1), using the VLOOKOUT function on Excel (MS Office). Harmonised allele profiles were compared using GENALEX software (Peakall and Smouse 2006), and identical samples verified manually.

Alignment of the data with national datasets was carried out by partner 8. Initially, this was based upon the creation of a general rule which was formulated based on the direct comparison of all samples replicated between each national dataset and either EU.Cherry or Prunus Alignment centralised data. An independent 'rule' was generated for each marker on a country by country basis. Subsequent comparison of scores for the ECPGR reference samples, where they were available within national datasets, and of alleles in the replicated samples where the 'rule' appeared not to apply, allowed a finer tuning of the alignment factor on an allele by allele, per marker, per country basis. Any profiles that stood out as clearly being in disagreement with the established alignment factor were marked as potential errors in either the national data or EU.Cherry/Prunus Alignment data as appropriate.

Following this initial alignment, a search for identical samples was carried out using Cervus software (Kalinowski et al. 2010). Differential thresholds were applied to the comparison of data from different national datasets, due to a differing level of overlap with the ECPGR and updated EU.Cherry marker set (as an example of the range, the Swedish data contained 5 ECPGR/EU.Cherry markers whilst the British data contained all 14).

RESULTS

Conventional SSR analysis and associated characterization of plum accessions

A full dataset of SSR scores was produced for the selected samples. Along with the data previously generated in the PRUNDOC project, these formed the basis of analysis for a scientific publication that was produced, with the collaborative input of Fuad Gasi (University of Sarajevo and ECPGR Malus/Pyrus working group member, Bosnia and Herzegovina) and colleagues (Gasi et al., 2020). The publication was led by Hilde Nybom (partner 13) and investigated the genetic structure within the germplasm, including the identification of potentially replicate samples within the dataset. All unique genotypes would be expected to be potential candidates for AEGIS.

The generation of data within the project, and the analysis carried out in line with the associated publication (along with the prior analysis carried out within PRUNDOC) was deemed to form a valuable basis on which to make a recommendation within ECPGR of a standard set of markers and reference accessions for the genotyping of plums by conventional SSR. This was discussed during the project meeting and a publication including these recommendations was subsequently submitted to, and has now been published in, the first issue of the new 'Genetic Resources' journal (Nybom et al., 2020). Again, the publication was led by Hilde Nybom (partner 13).

FPD (and SPD) data for the newly nominated accessions that were not planned to be genotyped was generated by partner 8 and self-funded partner 1; an additional 22 genotypes were described with both FPD and SPD and these should be considered as 'in principle' MAAs as per the non-genotyped accessions in PRUNDOC. In addition to this, outstanding characterization data were circulated for 7 Belgian accessions where data were not reported in PRUNDOC. Accession names of these 'in principle MAAs are listed in Annex II.

Analysis of plum accessions by HRM and comparison between labs

HRM data was generated for the subsets of accessions by partners 1, 2 and 12. An initial round of data production was carried out by partner 8 but, unfortunately, a critical failure in the real-time PCR machine at the University of Reading made it impossible to complete a full dataset for comparison with the other partners. Nonetheless, three replicate datasets were produced.

A full set of HRM data were produced for the master set of 58 (including additions from Greece) accessions by partner 5. Using these data, a cluster (UPGMA) tree was generated, along with a detailed analysis of genetic parameters such as allele frequency etc. Although it has not yet been possible to fully compare and discuss the implications of the findings between partners, this is expected to form the basis of a scientific publication.

The normalized HRM melting curves of nine representative plum genotypes (using the microsatellite marker UCD-CH17) are depicted in Figure 1: where only the unique HRM genotypes are shown. The shape of the melting curves could reveal the differences between the cultivars under investigation and show that all cultivars used could be easily distinguished visually by their melting curves, as for example in cultivars 'ITA 8' and 'UK 1'. The results with the rest of the microsatellite markers were similar, showing a clear discrimination of most of the genotypes used.

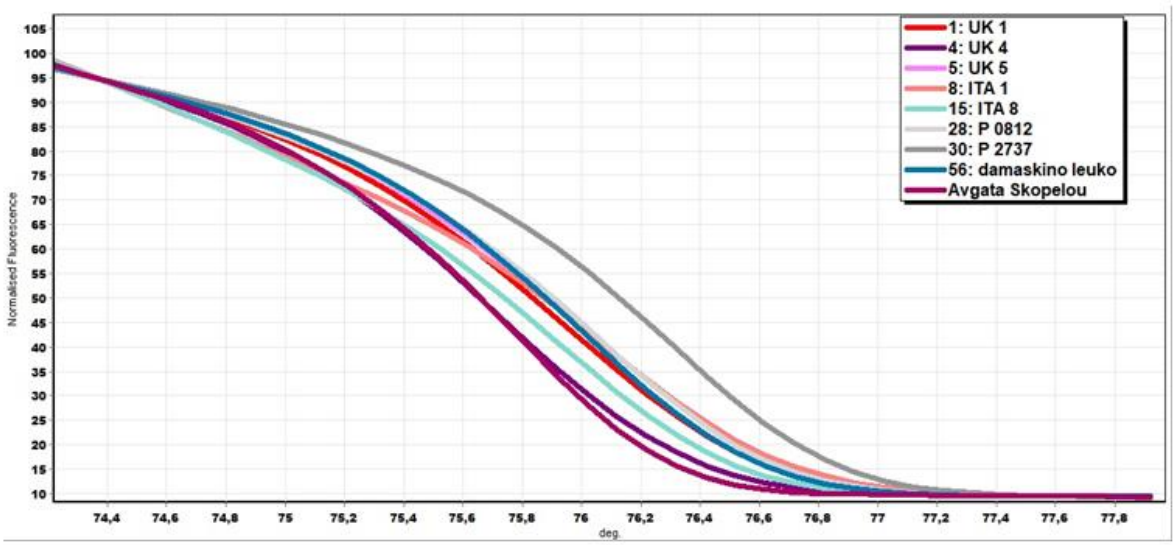
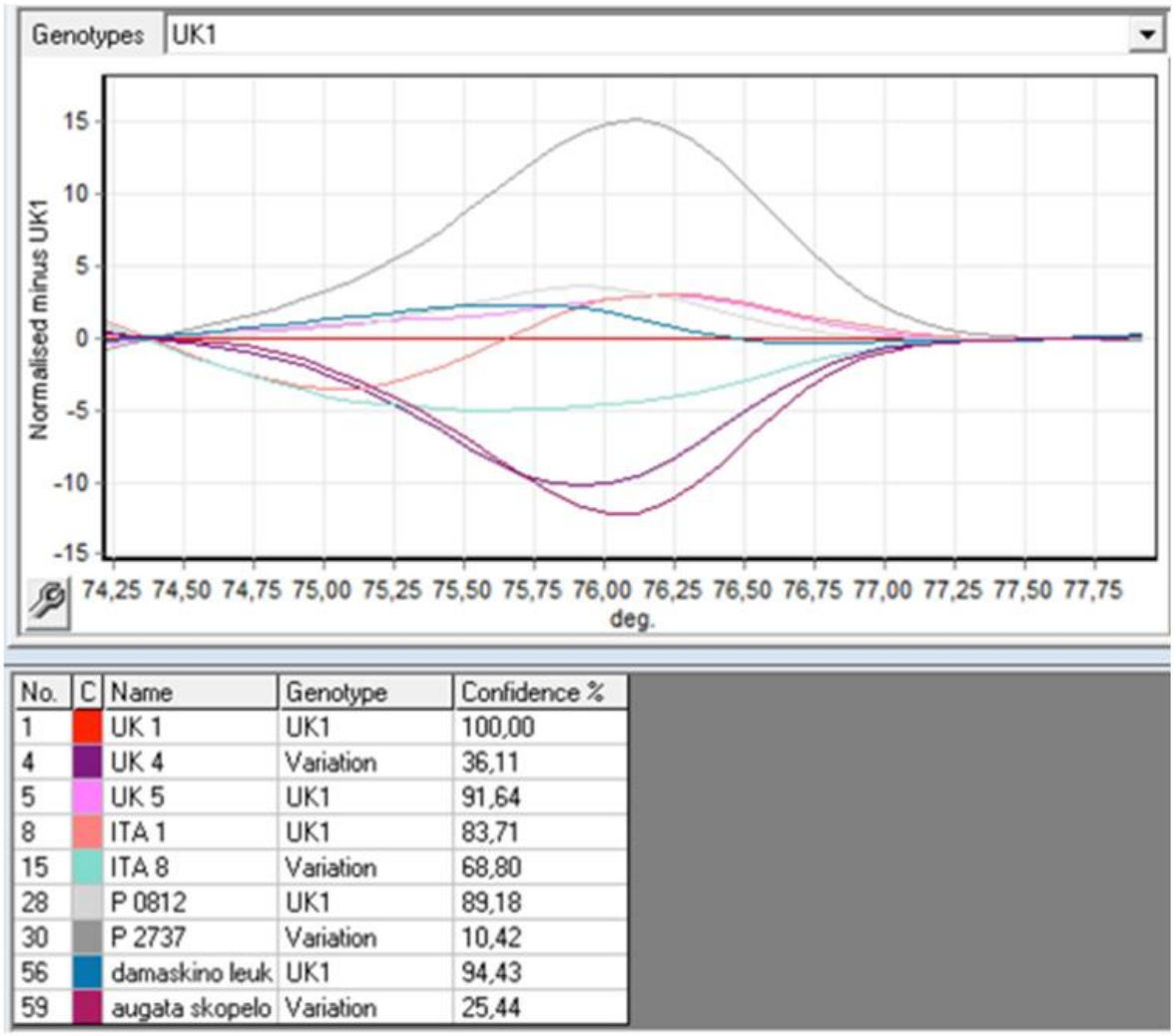


Figure 1. HRM analysis of representative plum genotypes with microsatellite marker UCD-CH17. Normalized HRM melting curves for nine unique plum genotypes that are using HRM analysis.

Figure 2 depicts the difference graph produced by the UCD-CH17 SSR marker on a representative set of 9 unique plum cultivars as compared with cultivar 'UK 1', which was used as the baseline. The confidence value of similarity between 'Black Star' and the rest cultivars was estimated and showed that UCD-CH17 was a sufficient molecular marker to differentiate most of the 60 plum cultivars. UCD-CH17 marker combined with HRM analysis represents a polymorphic microsatellite marker, which identified nine different HRM profiles while all other genotypes followed one of these nine distinct curves.



No.	C	Name	Genotype	Confidence %
1		UK 1	UK1	100,00
4		UK 4	Variation	36,11
5		UK 5	UK1	91,64
8		ITA 1	UK1	83,71
15		ITA 8	Variation	68,80
28		P 0812	UK1	89,18
30		P 2737	Variation	10,42
56		damaskino leuk	UK1	94,43
59		augata skopelo	Variation	25,44

Figure 2**.** Representative profiles of the melting curves (difference plot curves) of UCD-CH17 amplicons for plum genotypes. Difference graph of nine unique genotypes using the cultivar ‘UK 1’ as reference genotype.

The detected polymorphic HRM curves originating from the seven microsatellite markers were used to construct three similarity dendrograms using the NJ cluster algorithm (Figure 3).

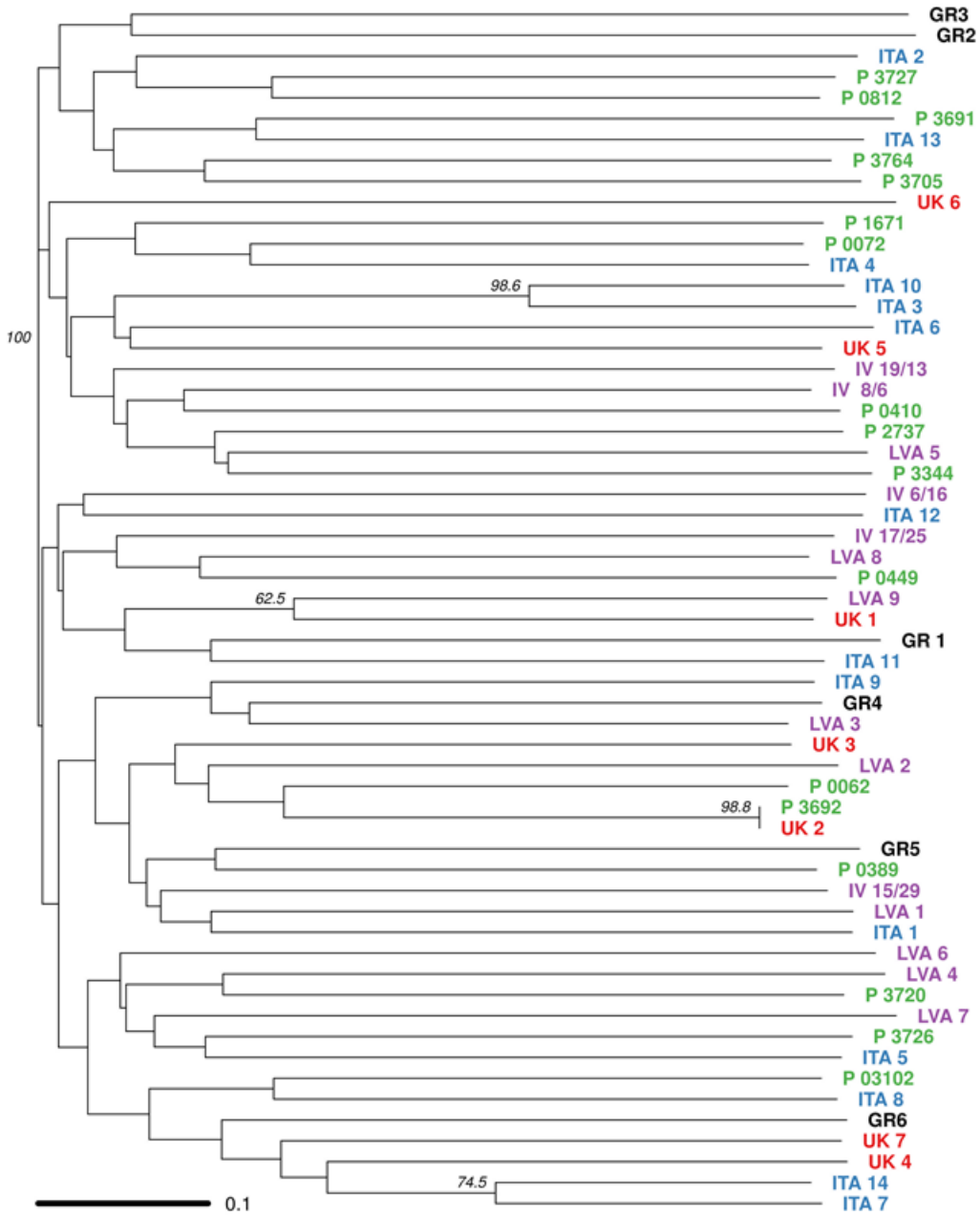


Figure 3. NJ dendrogram showing the genetic relationships among 60 plum cultivars as revealed by the microsatellite genotyping with HRM analysis. Numbers above the lines indicate bootstrap values (percentage of 1000 replicates). Bootstrap values greater than 50% are shown.

To get a clearer picture of the genetic relationships between the analysed groups of cultivars, principal coordinate analysis (PCoA) was performed on the molecular data. In the PCoA scatter plot (Figure 4), a slight separation is present between the traditional Greek cultivars and the foreign plum cultivars supporting the dendrogram (Figure 3) results in a robust way, considering the high percentage of the total genetic diversity in the first two axes (16.17%).

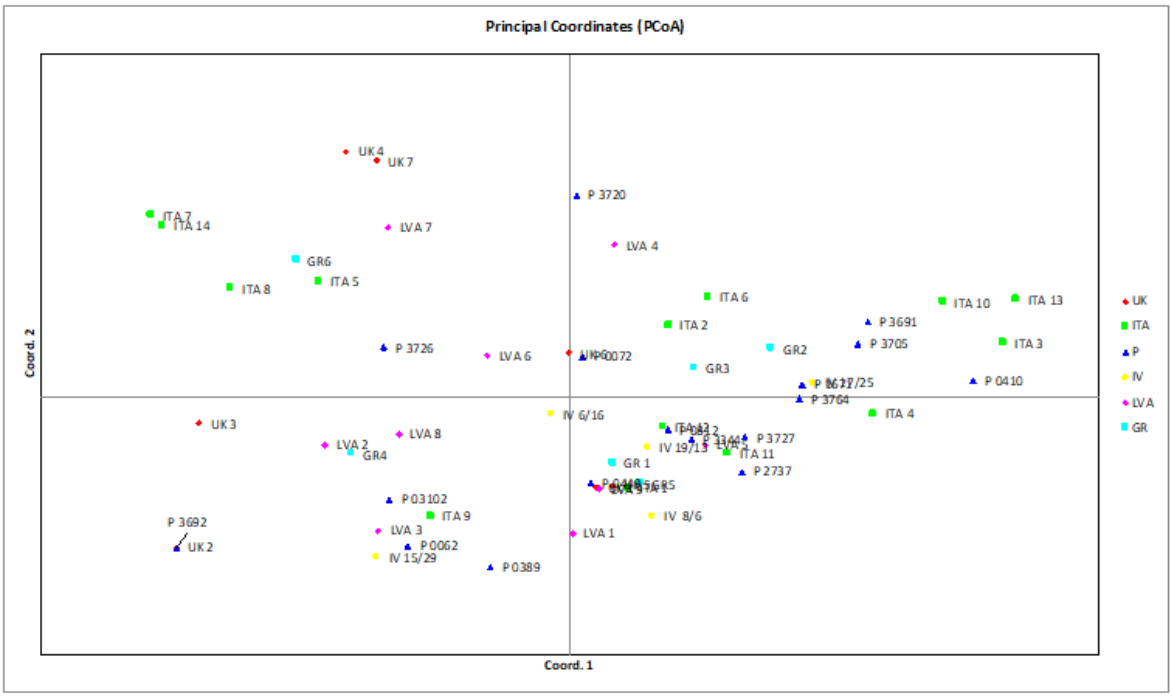


Figure 4. Principal coordinate scatter plot of 60 plum cultivars based on seven microsatellite markers.

A number of queries have been raised by the partners carrying out replication of the subsample (partners 1, 2 and 12) and these have been passed to partner 5, although it has not yet been possible to discuss them in detail. Again, it is expected that these will be considered fully in a scientific manuscript.

The intention of the partners is to produce and submit a manuscript that will assess the use and reproducibility of the technique and discuss any issues of aligning and comparing data between different laboratories, such that a full and detailed assessment will be available to ECPGR members and the wider community.

SSR analysis of cherry accessions and alignment of national datasets with EU.Cherry

Alleles detected with the 14 SSR markers are presented for all samples and standards in 'Annex 1 – Full Scores' of the partner 11 report. In many cases, low signal quality data was obtained for markers EMPa017 and CPSCT038. No PCR amplified alleles were observed in marker CPSCT038 for standard P. incisa E621, standard P. nipponica F1292, or sample SLU005 ('Pernilla', Sweden), despite repeated PCR amplification. In addition, the large allele sizes observed in marker PAV-Rf-SSR made it necessary for this marker to be scored separately from other markers in the GENESCAN software.
Of the 99 samples screened, 82 presented allelic patterns corresponding to diploid (2n) genotypes (i.e. one or two alleles per marker), while 17 presented additional alleles in various markers suggesting triploid (3n) or tetraploid (4n) genotypes ('Annex 1 – Full Scores').

A number of alleles were observed in this Prunus Alignment analysis that were not reported in the EUCHERRY project, in some cases expanding the allele range observed for markers. The full list of alleles observed in EUCHERRY and Prunus Alignment can be found in 'Annex 1 – Alleles Observed' of

the partner 11 report. In addition, this analysis suggests a number of changes to allele scores presented in EUCHERRY ('Annex 1 - Allele changes EUCHERRY' of the partner 11 report) that most likely result from differences in rounding and harmonisation.

Attempts at alignment of the data, carried out by partner 8, appear to have been reasonably successful. An initial output, detailing 46 groups of what appear to be indistinguishable accessions between collections in differing countries (i.e. either samples in two or more different national datasets, or samples in one or more national dataset of a country different to the source of matching data in EU.CHERRY) has been circulated between the partners for consideration. Eight further indistinguishable groups have since been identified, and it is expected that significant numbers of additional groups worthy of investigation will be revealed as the data is scrutinised and either the level of acceptance is lowered, or the alignment is further refined based on the initial findings.

One immediate observation is the apparent resolution of what would appear to be a potential labelling error in a small number of EU.CHERRY data entries; specifically, a set of the samples analysed from Germany. Through the process of alignment, it has been possible to establish that these samples are most likely mislabelled in the original EU.CHERRY dataset (the apparent error being associated with the data being labelled with the accessions in alphabetical order by name in the dataset, rather than the order of the samples as collected). This error has been notified to both NIAB-EMR, colleagues at INRAE (France) who are in the process of analysing EU.CHERRY data for potential publication and the German partner in Prunus Alignment.

RECOMMENDATIONS

We would recommend that all of the plum accessions previously listed as 'in principle' most appropriate accessions (MAAs) within PRUNDOC be considered for inclusion within AEGIS by the appropriate national representatives. We would also recommend that the newly identified accessions which were also confirmed to be unique, should be considered as MAA by the relevant national representatives.

A recommendation for a standard set of markers and reference accessions for the conventional SSR genotyping of plum accessions within ECPGR has been made. This has been detailed in a manuscript that was published in the first edition of the journal Genetic Resources (Nybom et al., 2020). We would recommend that members of ECPGR, and others attempt to follow this recommendation when carrying out conventional SSR analysis of plum germplasm.

We would currently recommend that further investigation and/or optimization would be required before HRM as a technique would be a valuable tool for producing comparable data for plum germplasm across ECPGR members. Further, more detailed, analysis and recommendations are expected to form the basis of a publication of results.

We would recommend that partners review the apparent duplication revealed through the alignment of SSR data for cherry. From the initial findings, it would appear that this alignment has potentially revealed a combination of duplication, mislabeling and potential synonymy within the overall ECPGR cherry germplasm. The full resolution of these indistinguishable groups will require further analysis of both morphological similarity (in order to check for potential sports) and of the provenance of both accessions and named cultivars (in order to check for potential mislabelling and/or to assess the potential that true synonymy has been revealed).

DATA AVAILABILITY

All data from conventional SSR analysis and the associated characterization data are available within the supplementary material of the publication by Gasi et al. (2020).

Additional characterization data for the accessions that were not genotyped have been submitted to the Prunus DB manager.

HRM data are expected to be made fully available through the publication of an associated manuscript. Without the detailed associated analysis that will be required for publication, these data would be of relatively little use to members.

The aligned SSR data are expected to be made fully available through an associated publication. They have also been submitted, with the approval of all relevant partners, to the Prunus DB manager.

COMPARISON OF REPORTED VS EXPECTED RESULTS

Conventional SSR analysis and associated characterization of plum accessions

We expected that a total of 59 accessions would be genotyped, with 37 of these being accompanied by SPD data and 22 being accompanied by FPD data (coming from the partners not included in PRUNDOC, or unable to be characterized due to regeneration). Twenty-eight additional accessions were expected to have FPD data collected/collated. In actuality, the genotyping ran to 64 samples in total and for all of these, a range of both FPD and SPD data were included in the analysis. In addition, SPD data were made available in the publication for 36 accessions that had been genotyped in PRUNDOC but only previously genotyped with FPD. At the time of reporting, the number of non-genotyped samples with FPD was only 22 (and this included samples from self-funded partner 1 that was not included in the original count), although these did all additionally have SPD scored for each of them.

Analysis of plum accessions by HRM and comparison between labs

We expected that a master set of 60 accessions would be genotyped by Partner 5 with 24 samples replicated by four other partners. In actuality, the master set ran to 58 accessions (due to the availability of samples) and only three partners were able to replicate subsets of material (due to the failure of machinery at the University of Reading). However, each of the three partners generated data for a larger subset of 36 accessions.

SSR analysis of cherry accessions and alignment of national datasets with EU.Cherry

We expected that a master set of approximately 140 accessions would be genotyped by partner 11 and an unknown number of data entries (minimum of approximately 800) would be available for alignment from the national datasets. In actuality, 113 accessions (plus 8 reference genotypes) were selected from the partner collections based on an initial attempt at aligning data against the EU.CHERRY dataset. It was not possible to obtain samples from all of these, and in the end 99 samples were received and genotyped alongside the 8 reference genotypes by partner 11 (i.e. 107 samples in total). The total number of national data entries that were aligned was 2,242.

ACKNOWLEDGEMENTS

We would like to specifically thank Jasna Sehic for her technical work in the analysis of plum SSR profiles and Ioannis Ganopoulos for his technical guidance and the analysis of the HRM work at HAO. We would like to thank Marco Pietrella for the analysis of HRM samples at CREA-OFA. We would like to thank Suzanne Litthauer for her technical work in the analysis of cherry SSR profiles for the project and Edward Venison for the sharing of the UK cherry SSR data. We would like to thank Stephanie Mariette, Markus Kellerhals and Christina Kaegi for collaborating to share French and Swiss national datasets with the project and Annalisa Marchese for agreeing to collaborate with the downstream analysis of cherry data. We would like to thank Georgios Merkouropoulos for collaborating in the analysis of HRM data and we would like to thank Fuad Gaši for leading the analysis of plum SSR data for publication.

BIBLIOGRAPHY

Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E and Testolin R (1999). AC/GT and AG/CT microsatellite repeats in peach [Prunus persica (L) Batsch]: isolation, characterisation and cross-species amplification in Prunus. Theoretical and Applied Genetics 99: 65–72.

Clarke JB, Tobutt KR (2009). 'A standard set of accessions, microsatellites and genotypes for harmonising the fingerprinting of cherry collections for the ECPGR' Acta Horticulturae 814:615-61
Dirlewanger E, Cosson P, Tavaud M, Aranzana M, Poizat C, Zanetto A, Arús P and Laigret F (2002). Development of microsatellite markers in peach [Prunus persica (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (Prunus avium L.). Theoretical and Applied Genetics 105: 127–138.

Downey SL and Iezzoni AF (2000). Polymorphic DNA markers in black cherry (Prunus serotina) are identified using sequences from sweet cherry, peach, and sour cherry. Journal of the American Society for Horticultural Science 125: 76–80.

Edge-Garza D.A, Rowland T.V, Haendiges S, Peace C (2014). 'A high-throughput and cost-efficient DNA extraction protocol for the tree fruit crops of apple, sweet cherry, and peach relying on silica beads during tissue sampling'. Molecular Breeding 34:2225–2228

Gaši F, Sehic J, Grahic J et al. (2020). Genetic assessment of the pomological classification of plum Prunus domestica L accessions sampled across Europe. Genetic Resources and Crop Evolution. https://doi.org/10.1007/s10722--020-00901-y

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

Hampl, V, Pavlícek, A, and Flegr, J (2001). Construction and bootstrap analysis of DNA fingerprinting-based phylogenetic trees with the freeware program FreeTree: application to trichomonad parasites. International Journal of Systematic and Evolutionary Biology 51:731-735.

Kalinowski, ST, Taper ML & Marshall, TC (2010). Corrigendum. Molecular Ecology 19: 1512. http://dx.doi.org/10.1111/j.1365-294X.2010.04544.x

Merkouropoulos, G., Ganopoulos, I., Tsaftaris, A., Papadopoulos, I., & Drogoudi, P. (2017). Combination of high resolution melting (HRM) analysis and SSR molecular markers speeds up plum

genotyping: Case study genotyping the Greek plum GeneBank collection. Plant Genetic Resources, 15(4), 366-375. doi:10.1017/S1479262116000022

Messina R, Lain O, Marrazzo MT, Cipriani G and Testolin R (2004). New set of microsatellite loci isolated in apricot. Molecular Ecology Notes 4: 432–434.

Nei M and Li W (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences of the United States of America 76: 5269–5273.

Nybom, H., Giovannini, D., Ordidge, M., Hjeltnes, S. H., Grahić, J. and Gaši, F. (2020). ECPGR recommended simple sequence repeat loci for analyses of European plum (Prunus domestica) collections. Genetic Resources, 1. pp. 40-48. ISSN 0925-9864 doi: https://doi.org/10.46265/genresj.2020.1.40-48

Ordidge M. et al. (2019). Prunus Alignment – integrating data for European plum and cherry germplasm (863,9 KB). Poster presented at the XV EUCARPIA Fruit Breeding and Genetics Symposium, 3-7 June 2019, Prague – Suchdol, Czech Republic.

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

Rosyara UR, Bink MC, Van de Weg E, Zhang G, Wang D, Sebolt A,Dirlewanger E, Quero-Garcia J, Schuster M, Iezzoni AF (2013). 'Fruit size QTL identification and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of sweet cherry' Molecular Breeding 32:875–887

Sandefur P, Oraguzie N, Peace C (2016). 'A DNA test for routine prediction in breeding of sweet cherry fruit color, Pav-Rf-SSR' Molecular Breeding 36: 33

Struss D, Ahmad R, Southwick SM and Boritzki M (2003). Analysis of sweet cherry (Prunus avium L.) cultivars using SSR and AFLP markers. Journal of the American Society for Horticultural Science 128: 904–909.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731–2739.

Xanthopoulou A, Ganopoulos I, Tsaballa A, Nianiou-Obeidat I, Kalivas A, Tsaftaris A and Madesis P (2013). Summer squash identification by high-resolution-melting (HRM) analysis using gene-based EST-SSR molecular markers. Plant Molecular Biology Reporter 32: 395–405.

ANNEXES

Annex I. Plum accessions analysed by conventional SSR and/or HRM

Accession details are modified from Gasi et al. (2020).
Entries in grey were not analysed by conventional SSR within Prunus Alignment (having been previously analysed in PRUNDOC) but were included in the subsets analysed by HRM.
Entry in grey italic was not analysed by conventional SSR in either project but was included to allow comparison with results of previous analysis by HRM.

Country	Sample	Accession	Cultivar	Inclusion in HRM (Large or Small subset)
Germany	DEU5	PFL0022	Mirabelle de Flotow (FRA)	
	DEU6	PFL0012	Gräfin Cosel	
	DEU7	PFL0014	Haferpflaume	
Denmark	DNK1	POM B197	Gul Havreblomme	
	DNK2	POM B141	Gul Rosinblomme	
	DNK3	POM B25	Gul Sveskeblomme	
	DNK4	POM B70	Kongeblomme	
	DNK5	POM B205	Spilling (DEU)	
Estonia	EST1	EST2169	Amitar	
	EST2	EST2170	Ave	
	EST3	EST2176	Liisu	
	EST4	EST2186	Tartu Punane	
	EST5	EST2250	Kihelkonna	
France	FRA1	P3691	Reine-Claude Diaphane	L
	FRA2	P3727	D'Ente Double	L
	FRA3	P0302	Impériale Murat	L
	FRA4	P3720	Oustenque Bleue	L
	FRA5	P0062	Abricotée Jaune	L
	FRA6	P0072	Mirabelle Parfumée de Septembre	L
	FRA7	P0328	Madame Guttin	L
	FRA8	P0389	Reine-Claude Davion	L
	FRA9	P0410	Quetsche de Wagenstadt	L
	FRA10	P0449	Impériale Epineuse	L
	FRA11	P0812	Double Robe	L
	FRA12	P1671	Verdanne	L
	FRA13	P2737	Prune de Vars	L
	FRA14	P3344	Prune de Chien	L
	FRA15	P3692	Bonjour	L
	FRA16	P3705	Saint-Léonard	L
	FRA17	P3726	Prune de Chien	L

	FRA18	P3764	Briquetch	L
Great Britain	GBR1	1977-186	Czar	L/S
	GBR2	1977-187	Marjorie's Seedling	L/S
	GBR3	2003-002	Grand Duke	L/S
	GBR4	1949-225	Winesour	L/S
	GBR5	2000-127	Late Orange	L/S
	GBR6	1949-255	Utility	L/S
	GBR7	1976-046	Victoria	L/S
Greece	GRC3	PD0006	Ksina Skopelou	L/S
	GRC4	PD0007	Avgata Skopelou	L/S
	GRC5	PD0001	Asvestochoriou	L/S
	GRC6	PD0004	Mpardaki Circular (BEL)	L/S
	GRC7	PD0002	Praousti	L/S
	GRC8		Damaskino Lefko	L/S
Hungary	HUN1	prudo20	Duránci	L
	HUN2	prudo229	Tölcsér koronájú	L
	HUN3	prudo436	Potyó fehér	L
	HUN4	prudo484	Besztercei 5/a (Balkan?)	L
	HUN5	prudo522	Potyó szilva	L
	HUN6	prudo530	Tarka perdrigon (FRA?)	
Italy	ITA1	264	Susino Secondo	L/S
	ITA2	220	Ramassin Giallo	L/S
	ITA3	216	Ramassin Ramassin	L/S
	ITA4	438	Agostana	L/S
	ITA5	128	Gaiotti	L/S
	ITA6	147	Lazzarinu	L/S
	ITA7	189	Muninca	L/S
	ITA8	196	Paradisu	L/S
	ITA9	214	Prunella	L/S
	ITA10	219	Ramassin di Pagno	L/S
	ITA11	240	Sanacore	L/S
	ITA12	249	Sighera	L/S
	ITA13	59	Caleca	L/S
	ITA14	62	Cariadoggia	L/S
Latvia	LVA1	LVA01016	Zemgale	L/S
	LVA2	LVA01006	Aizputes	L/S
	LVA3	LVA01013	Latvijas Sarkanā Olplūme	L/S
	LVA4	LVA01010	Latvijas Dzeltenā Olplūme	L/S
	LVA5	LVA01008	Kārsavas	L/S
	LVA6	LVA01009	Lāse	L/S

	LVA7	LVA02549	Zilā Lāse	L/S
	LVA8	LVADPru1	Suhkruploom (EST)	L/S
	LVA9	LVADPru2	Julius (EST)	L/S
Norway	NOR5	1398	Tråneplomme	
Sweden	SWE2	BF0229	Hackman	
	SWE3	BF0237	Herman	
	SWE4	BF0243	Ive	
	SWE5	BF0249	Jubileum	
	SWE6	BF0337	Opal	

Annex II. Additional plum samples characterized but not genotyped

Accession name	Country of origin	European, Greengage, Bullace/damson, Mirabelle, 'Wildtype'
Altesse Dorée	BEL	European
Belle de Louvain	BEL	European
Belle de Thuin	BEL	European
Prune de Prince	BEL	Mirabelle
Reine-Claude Souffriau	BEL	European
Sainte-Catherine	BEL	European
Wignon	BEL	European
Early Laxton	UK	European
Early Rivers	UK	European
Monarch	UK	European
Pond's Seedling	UK	European
President	UK	European
Bjelica	HRV	European (Prunus domestica)
Bistrica	HRV	European (Prunus domestica)
Mandalenka	HRV	European (Prunus domestica)
Motičanka	HRV	European (Prunus domestica)
Pasjara	HRV	European (Prunus domestica)
Torgulja plava	HRV	European (Prunus domestica)
Trnovača	HRV	European (Prunus domestica)
Turkinja	HRV	European (Prunus domestica)
Bjelica jajara	HRV	Mirabelle (P. insititia var. syriaca)
Bijela kasna mirisava	HRV	Mirabelle (P. insititia var. syriaca)
Bijela sitna	HRV	Mirabelle (P. insititia var. syriaca)
Cerićanka	HRV	Mirabelle (P. insititia var. syriaca)
Pintara	HRV	Mirabelle (P. insititia var. syriaca)
Sitna bijela	HRV	Mirabelle (P. insititia var. syriaca)
Valpovka	HRV	Mirabelle (P. insititia var. syriaca)
Debeljara	HRV	Damson (P. insititia var.damascena)
Ružica	HRV	Damson (P. insititia var.damascena)