Genotyping of pedigreed apple breeding material with a genome-covering set of SSRs: trueness-to-type of cultivars and their parentages

K. M. Evans, A. Patocchi, F. Rezzonico, F. Mathis, C. E. Durel, F. Fernández-Fernández, A. Boudichevskaia, F. Dunemann, M. Stankiewicz-Kosyl, et al.

Molecular Breeding
New Strategies in Plant Improvement

ISSN 1380-3743
Volume 28
Number 4

Mol Breeding (2011) 28:535-547
DOI 10.1007/s11032-010-9502-5

Springer

Mol Breeding (2011) 28:535–547
DOI 10.1007/s11032-010-9502-5

Genotyping of pedigreed apple breeding material with a genome-covering set of SSRs: trueness-to-type of cultivars and their parentages

K. M. Evans · A. Patocchi · F. Rezzonico · F. Mathis · C. E. Durel · F. Fernández-Fernández · A. Boudichevskaia · F. Dunemann · M. Stankiewicz-Kosyl · L. Gianfranceschi · M. Komjanc · M. Lateur · M. Madduri · Y. Noordijk · W. E. van de Weg

Received: 21 May 2010 / Accepted: 26 August 2010 / Published online: 2 October 2010

Abstract Apple cultivars and breeding lines that represent much of the diversity currently present in major European breeding programmes and are genetically related by their pedigree were examined for the trueness of their identity and parentage by consistency in marker scores using a genome-covering set of 80 microsatellite (SSR) markers and an 'identity-by-descent' approach. One hundred and twenty-five individuals were validated for the trueness-to-type of both their parents and 49 were validated for one of their parents, their second being unknown (23 individuals) or not available in this study (26 individuals). In addition, 15 individuals for which we lacked one of or both the direct parents were validated by consistency with tested parents of earlier generations. Furthermore, the identity of 28 founder cultivars was validated, their marker scores being consistent with descending cultivars and breeding

Electronic supplementary material The online version of this article (doi:10.1007/s11032-010-9502-5) contains supplementary material, which is available to authorized users.

K. M. Evans · F. Fernández-Fernández
East Malling Research, New Road, East Malling, Kent ME19 6BJ, UK

A. Patocchi · F. Rezzonico
Plant Pathology, Institute of Integrative Biology (IBZ), ETH Zurich, 8092 Zurich, Switzerland

F. Mathis · C. E. Durel
Genetics and Horticulture (GenHort), UMR1259 INRA/UA/AgroCampus-Ouest, INRA Centre d'Angers-Nantes, IFR 149 QUASAV, 60057, 49071 Beaucouzé Cedex, France

A. Boudichevskaia · F. Dunemann
Bundesanstalt für Züchtungsforschung an Kulturpflanzen, Institut für Obstzüchtung, Pillnitzer Platz 3a, 01326 Dresden, Germany

M. Stankiewicz-Kosyl
Laboratory of Basic Research in Horticulture, Faculty of Horticulture and Landscape Architecture, Warsaw Agricultural University (WAU), ul. Nowoursynowska 166, 02–787 Warsaw, Poland

L. Gianfranceschi
Department of Biomolecular Sciences and Biotechnology, University of Milan, Via Celoria 26, 20133 Milan, Italy

M. Komjanc
Istituto Agrario di San Michele all'Adige, 38010 San Michele all'Adige, Trento, Italy

M. Lateur
Centre Wallon de Recherches Agronomiques (CRA-W), Unité d'Amélioration & Biodiversité, Bâtiment Emile Marchal, Rue de Liroux, 4, 5030 Gembloux, Belgium

M. Madduri · Y. Noordijk · W. E. van de Weg
Department of Plant Breeding, Plant Research International, Wageningen University and Research Centre, P.O. Box 16, 6700 AA Wageningen, The Netherlands

K. M. Evans (✉)
WSU-TFREC, 1100 N. Western Ave, Wenatchee, WA 98801, USA
e-mail: kate_evans@wsu.edu

lines. Four of the eight triploids identified were clearly shown to have arisen from unreduced egg cells. The assumed pedigree of 15 further individuals was found to be incorrect; fully consistent pedigrees were suggested for three of the cultivars. The pedigrees of a further eight individuals were confirmed through inference from the molecular data.

Keywords *Malus* × *domestica* Borkh · Microsatellites · Founders · Triploid

Introduction

Apple breeding programmes have developed breeding lines based on improving traditional cultivars that were bred or discovered as chance seedlings decades or even centuries ago. Many such lines have focused on introducing resistance to disease from other *Malus* species into the better quality *Malus* × *domestica* Borkh. lines. Apple is a self-incompatible species with a long juvenile phase, so introgressing new characters into the *M.* × *domestica*-quality background requires a long-term commitment. The degree of control used by the breeder to avoid out-crossing has varied through the generations depending on both the breeder and the purpose of the cross; hence, although most new cultivars are released with known parentage, others can be less certain. The development of molecular markers has lead to the identification of some parentage inconsistencies already reported in the literature; for example, Nova Easygro (Gianfranceschi et al. 1996), F_2 26829-2-2 (Vinatzer et al. 2004), and Honeycrisp (Cabe et al. 2005). Also, the regular exchange of material between breeding programmes and variety collections has resulted in a number of unidentified naming errors (e.g. the accession GMAL 2473 donor of the *Rvi15* apple scab resistance gene; Patocchi et al. 2004).

The EU-funded HiDRAS project (High-quality Disease Resistant Apples for a Sustainable agriculture) (Gianfranceschi and Soglio 2004) used pedigree-based quantitative trait locus (QTL) analysis to follow the segregation of specific chromosomal regions with highly polymorphic co-dominant microsatellite (SSR) markers in related cultivars, breeding selections and small progenies. The pedigree-based analysis (PBA) system can analyse data from progenies from ongoing breeding programmes, thus reducing the need to establish large and expensive specific mapping progenies, as well as reducing the need for the high marker densities required for linkage disequilibrium-mapping in presumed genetically unstructured germplasm (van de Weg et al. 2004; Bink et al. 2008). Modern breeding lines with well-documented pedigrees linked by major founder cultivars from six European countries have been analysed within HiDRAS.

Although the principal aim of this 'identity-by-descent' approach was to identify QTL controlling apple fruit quality which could be applied as preselection in breeding programmes, pedigree information of the plant material had to be verified initially for this approach to succeed.

Genotyping the plant material with a set of highly polymorphic SSRs spanning the whole genome (Patocchi et al. 2009a) and verification of the consistencies of the generated data has enabled correct genotypes to be identified and several inconsistencies of reported parentage to be corrected with confidence.

In this paper we present a description of the plant material analysed, the strategy used to identify mistakes in the reported pedigrees and the individuals that are true-to-type (TTT), resolving or confirming some pedigrees for which uncertainties have been reported. Furthermore, we report the confirmation of eight triploids together with a possible explanation of their derivation in seven of the eight cases.

Materials and methods

Plant material

The plant material used in the study was as described in Patocchi et al. (2009b). In summary, a series of small, genetically related progenies and the cultivars

Present Address:
A. Patocchi · F. Rezzonico
Agroscope Changins-Wädenswil Research Station, Plant Protection and Fruit and Vegetable Extension, Phytopathology, Schloss, 8820 Wädenswil, Switzerland

A. Boudichevskaia
Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany

and breeding lines which were present in the pedigree of any of these progenies were analysed. Other pedigreed cultivars and breeding lines that were not in the pedigree of the progenies were also included. Cultivar samples were collected from more than one site when available, for more robust validation of identity. Samples were collected using a strategy allowing complete tracking from the plants in the field to the DNA aliquot used for genotyping (Antofie et al. 2007). Acronyms of six to nine letters were adopted for lengthy cultivar names, thus facilitating PediMap (Voorrips 2007) visualizations and making names compatible with FlexQTL™ (Bink et al. 2002; http://www.flexqtl.nl) format requirements. Published pedigrees used for verification with marker data are as cited in the text for specific examples or sourced from the following: GRIN (Germplasm Resources Information Network; http://www.ars-grin.gov), King et al. (1998), Morgan and Richards (2002), Stark (1974), http://www.pubs.ext.vt.edu/422/422-760/422-760.html#L2, http://www.applejournal.com/use.htm, http://www.hort.purdue.edu/newcrop/pri/, http://www.homeorchardsociety.org/ebooks/ and http://www.nyease.cornell.edu/pubs/fls/OCRPDF/151.pdf.

DNA analysis and validation

DNA was extracted and analysed with a set of 80 SSR markers that were evenly distributed across the apple genome (Patocchi et al. 2009a). Microsatellite data were validated in four successive steps as explained in Patocchi et al. (2009b). A small subset of 31 "common progenitor" accessions was nominated from four partners. In order to reduce the number of possible replicate samples to genotype, these samples were first analyzed with the multiplex PCR Hi12a (Patocchi et al. 2009a). When all the samples of an accession gave identical results for the six SSRs of this multiplex, a single representative sample was selected. In the cases where samples were shown to differ, a representative for each sub-set was maintained.

Pedigree analysis

Initial pedigree information for the plant material came from the participating breeders and was complemented by searches on the internet. As part of the validation process, genotyping data was first run through the xls-macro GOB-validator (unpublished) to check consistency between an accession and its direct parents when compared with pre-entered reported pedigrees. This was followed by FlexQTL™ (Bink et al. 2002) which checked consistency through the entire germplasm set. Errors of parentage became apparent when there was an inconsistency of marker scores between successive generations, e.g. a genotype having at least one marker allele that did not occur in either of its supposed parents. DNA samples that did not fit pedigrees were compared with samples from other sources where available. If the inconsistency remained after testing several samples of different origins and after re-examination of the marker scoring, the reported pedigree was presumed to be incorrect. In the rare case that there was only one source of the genotype and it was part of only one pedigree, it was not possible to establish whether the pedigree was incorrect or whether the sample was not true-to-type (NTTT). An incorrect pedigree was also assumed if a single sample did not fit to its presumed parents, but did fit correctly with its offspring.

Throughout this process, the Pedimap software (Voorrips 2007) was used to visualize the pedigree and marker information, thus supporting the understanding and in many cases the identification of the cause of the conflicting data, as in the presence of multiple inconsistencies FlexQTL™ usually indicates only the conflicting region. The score generating the inconsistency had to be manually traced.

Founders

Founders for the genetically related progenies, cultivars and breeding lines were defined as accessions from which allele flow cannot be traced further back to earlier generations. There are various reasons why this may be the case: for example, both parents are unknown (e.g. Winesap); or accessions that showed inconsistencies for their parentage or the parents are known but do not exist anymore (so consequently there is no DNA available) and that are also not part of other pedigrees within the HiDRAS germplasm as they were not needed to link multiple pedigrees (e.g. F_2 26829-2-2); or DNA is lacking for several successive generations which also cannot be indirectly genotyped through their involvement in multiple pedigrees (e.g. Prima). By introducing Prima as a new founder, its connection to F_2 26829-2-2 and Golden Delicious would be lost (Electronic

Supplementary Material Fig. 1). It is possible that earlier generations for some of the founders might be identified if a much wider source of germplasm was screened; however, as many of them were the result of chance seedlings rather than selective breeding, this is unlikely.

Identification of polyploids

Some accessions were known to be polyploid at the outset of HiDRAS. For others, a balanced amplification of more than two alleles from at least one of its parents and for multiple markers indicated polyploidy. Such samples could be distinguished from contaminated samples, where an unbalanced amplification leading to different intensities of amplified alleles would result. The evaluation of multiple SSRs avoids misclassification through the duplication of a chromosomal region within the same chromosome (possibly as a result of unequal crossing-over), such as the duplication of CH03d11 on one of the homologous regions of McIntosh and thereby many McIntosh offspring, or like a frequently occurring duplication of CH01h02. It also avoids misclassification through the occasional amplification on homoeologous chromosome segments. Where the parentage of the polyploid was known, the allele sizes were compared, which in some cases indicated the origin of the ploidy mutation.

Results

Consistency of DNA samples

In total, 490 DNA samples were available for 307 pedigreed cultivars and breeding lines. For 193 of these, a single DNA sample was available; therefore "true-to-type" (TTT) of these samples could only be assessed using pedigree information. For the remaining 114 individuals, at least one additional DNA sample from another source was available (297 samples in total). Differences in the SSR allele sizes between samples supposed to be identical were observed in 38 individuals (42.1%). Also in these cases, pedigree information was used to identify which sample was TTT.

Validation of pedigrees

During the validation process, it became evident that we had started with erroneous parentages for 24 accessions, due in part to administrative errors in the processing of data and also to errors in the information presented by the participating breeding programmes. These pedigrees were corrected and reanalyzed.

The marker data confirmed the parentage of multiple cultivars and breeding lines: (i) both parents could be confirmed for 125 accessions (Table 1); (ii) the female parent could be confirmed for 23 accessions for which the male parent is unknown and could not be suggested from the HiDRAS germplasm (Electronic Supplementary Material Table 1); (iii) one of the two reported parents could be confirmed for 26 accessions (Supplementary Table 2), and their other parent could not be confirmed as these were not included in the HiDRAS germplasm; and (iv) one of the two reported parents could be confirmed for 12 accessions, however their other parent was inconsistent with the marker data (Table 2). Moreover, parentages of 15 accessions for which one or both direct parents were not included could be indirectly confirmed either partially or entirely through matches with earlier generations (Table 3). This also led to parentages of one further cultivar and seven breeding lines being confirmed through inference (Table 3, note). The trueness-to-type of the DNA of 28 founders could be confirmed through the consistency of their marker scores with that of descending cultivars and breeding lines (Supplementary Table 3). Finally, fully matching alternative pedigrees are proposed for three accessions (Dayton, Liberty and X-2773) where the reported pedigrees are shown to be erroneous (Supplementary Table 4).

It was not possible to draw final conclusions on 75 accessions where only one plant of each accession was sampled and they had no links to other pedigrees analysed in this study, and therefore we could not distinguish between whether the accession was NTTT or whether a mistake was present in the reported pedigree or the TTT of the DNA sample itself. However, of these accessions, alternative matching parentages could be identified for Dukat (Dr Oldenburg × Cox), Elan (Golden Delicious × Ingrid Marie) and Melfree (Melrose × Tydeman's Early) (Supplementary Table 5). Molecular marker data suggested possible discrepancies with the parents but was

Table 1 Ninety-four cultivars and 31 breeding lines for which molecular marker data confirmed the TTT of their previously reported parents

Accession	Source	Mother	Father
Akane	4,7,9	Jonathan	Worcester Pearmain
Alkmene	4,7,10	Geheimrat Dr. Oldenburg	Cox's Orange Pippin
Apollo	7	Cox's Orange Pippin	Geheimrat Dr. Oldenburg
Arlet	7,9	Golden Delicious	Idared
Baujade	4,7,8	X-6799	Granny Smith
Cameo	8	Golden Delicious	Delicious
Chantecler	4	Golden Delicious	Clochard
Clivia	7	Geheimrat Dr. Oldenburg	Cox's Orange Pippin
Collina	4	Priscilla-NL	Elstar
Crandall	4	Rome Beauty	Jonathan
Cripps Pink	2,4	Golden Delicious	Lady Williams
Delcorf	8,10	Jongrimes	Golden Delicious
Delikates	9,10	James Grieve	Cortland
Delrouval	4,8	Delcorf	Akane
Discovery	2,4,7,9	Worcester Pearmain	Beauty of Bath
Dorianne	4	X-6823	X-4638
E210-80	6	Braeburn-EMR	Fiesta
E210-198	6	Braeburn-EMR	Fiesta
Early Geneva	9	Quinte	Julyred
Ecolette	10	Elstar	Prima
Elise	8,10	Septer	Cox's Orange Pippin
Elstar	2,4,7,9	Golden Delicious	Ingrid Marie
Empire	9,10	McIntosh	Delicious
Falstaff	7	James Grieve	Golden Delicious
Fantazja	7,9	McIntosh	Linda
Fiesta	2,4,7,9	Cox's Orange Pippin	Idared
Florina	2,4,8	PRI 612-1	Jonathan
Fuji	7	Ralls Janet	Delicious
Gala	2,4,7,9	Kidds Orange Red	Golden Delicious
Gloster	2,4,7,9	Weisser Winterglockenapfel	Delicious
Goldrush	4,7,10	Coop17	Golden Delicious
Goldsmith	4	Granny Smith	Golden Delicious
Greensleeves	7	James Grieve	Golden Delicious
Greenstar	9	Delcorf	Granny Smith
Himekami	4	Fuji	Jonathan
Holiday	9,10	Macoun	Jonathan
Honeygold	9,10	Golden Delicious	Haralson
Horei	7	Ralls Janet	Golden Delicious
Huagan	4	Golden Delicious	Fuji
Idagold	7	Wagenerapfel	Esopus Spitzenburg
Idajon	7	Wagenerapfel	Jonathan
Idared	2,4,7,9	Jonathan	Wagenerapfel
Ivette	7	Golden Delicious	Cox's Orange Pippin
Jamba	7,9,10	Melba	James Grieve
Jester	9	Worcester Pearmain	Golden Delicious

Table 1 continued

Accession	Source	Mother	Father
Jonadel	7	Jonathan	Delicious
Jonagold[3n]	8,10	Golden Delicious	Jonathan
Jonamac	2,4,7,9	McIntosh	Jonathan
Judaine	8	Reinette Du Mans	Priam
Judeline	8	Golden Delicious	Priam
Juga	10	Melrose	Idared
Julia	9	Quinte	Discovery
Jupiter[3n]	7,9,10	Cox's Orange Pippin	Delicious
Kanzi	8	Gala	Braeburn
Karmijn[3n]	7,8,10	Cox's Orange Pippin	Jonathan
Katja	7,10	James Grieve	Worcester Pearmain
Kidds Orange Red	4,7,10	Delicious	Cox's Orange Pippin
Ligol	9,10	Linda	Golden Delicious
Ligolina	9	Linda	Golden Delicious
Lord Lambourne	7,8	James Grieve	Worcester Pearmain
Melrose	7,8	Jonathan	Delicious
Merton Charm	7,9,10	McIntosh	Cox's Orange Pippin
Merton Worcester	10	Cox's Orange Pippin	Worcester Pearmain
Monroe	7,9	Jonathan	Rome Beauty
Murasaki	7	Jonathan	Delicious
Newtosh	9	McIntosh	Yellow Newton
Odin	7	Golden Delicious	Ingrid Marie
Odra	9	Primula	Bancroft
Pia	7	Idared	Helios
Pilot	2,4,7,9	Clivia	Undine
Pinova	2,4,7,9	Clivia	Golden Delicious
Pirol	7	Golden Delicious	Alkmene
Piros	7,9	Helios	Apollo
PRI 14-126	2, 11	Golden Delicious	F_2 26829-2-2
PRI 14-152	11	Golden Delicious	F_2 26829-2-2
PRI 612-1	4	Delicious	PRI 14-126
PRI 672-3	4	PRI 14-152	Golden Delicious
Priam	8,9,10	PRI 14-126	Jonathan
Quinte	7,9,10	Crimson Beauty	Melba
Redgold	9,10	Golden Delicious	Delicious
Reinette César	8	Golden Delicious	Cox's Orange Pippin
Rubens	2	Gala	Elstar
Rubin	2,4,7,9	Golden Delicious	Lord Lambourne
Rubinette	4, 8,9,10	Golden Delicious	Cox's Orange Pippin
Rubinstep	8	Clivia	Rubin
Sampion	4,10	Golden Delicious	Lord Lambourne
Sansa	7	Gala	Akane
Santana	3	Elstar	Priscilla-NL
Sawa	10	Fantazja	Primula
Selena	9	Britemac	Prima

Table 1 continued

Accession	Source	Mother	Father
Septer	4	Jonathan	Golden Delicious
Shinsei	7	Golden Delicious	Early McIntosh
SirPrize-GD4n	9	Golden Delicious	PRI 14-152
Slawa Pobjeditjeljam	9	White Transparent	McIntosh
Spencer	7,9,10	McIntosh	Golden Delicious
Summerland	7	McIntosh	Golden Delicious
Sundowner	7	Golden Delicious	Lady Williams
TN R10A8	4,8	Clochard	O53T136
Topaz	2,4,6,7	Rubin	Vanda
Tydemans Early Worcester	7,8	McIntosh	Worcester Pearmain
U1065	10	Idared	Primula
U1165	10	Siewka Gorjaczkowskiego	Priam
U636	9	Fantazja	Primula
U633	10	Fantazja	Primula
U641	9	Fantazja	Primula
X-3143	4	Winesap	X-2771
X-3177	4	Idared	Prima
X-3263	4	Red Winter	X-3177
X-3305	4	Chantecler	Baujade
X-3318	4	Fuji	X-3143
X-4295	4	Golden Delicious	Golden Delicious
X-4337	4	X-2773	Prima
X-4638	4	Clochard	PRI672-3
X-6064	4	Florina	Cloden
X-6417	4	Golden Delicious	TN R10A8
X-6564	4	Florina	Gala
X-6679	4	X-6823	Coop17
X-6808	4	Golden Delicious	X-4638
X-6823	4	Golden Delicious	Golden Delicious
X-6908	4	Liberty	X-3189
X-6911	4	Golden Delicious	NovaEasyGro-Liebh
Witos^{3n}	9,10	Fantazja	Primula
Z180	4,10	Golden Delicious	Anta34.16
Z185	4	Golden Delicious	Anta34.16
Z190	4	Golden Delicious	Anta34.16

Origin of the confirmed TTT accession is also noted

Source of accessions coded as follows: 2 = University of Bologna, Italy, 3 = PRI, Wageningen, The Netherlands, 4 = INRA, Angers, France, 6 = EMR, UK, 7 = JKI, Dresden, Germany, 8 = CRA, Gembloux, Belgium, 9 = Warsaw Agricultural University, Poland, 10 = RIPF Skierniewice, Poland, 11 = USDA-ARS Plant Genetic Resources Unit, Geneva, USA

3n: triploids

inconclusive for ten of these apple accessions (Auksis, Berlepsch, Ceres, Delorgue, Iduna, Laxton's Superb, Madame Colard, Novamac, Ruby and X-3305) due to the availability of just one DNA sample of each of these accessions; it could not be determined whether the DNA sample was NTTT, or the parentage of the accession is incorrect.

Speculative pedigrees

Several accessions, for which speculations on their parentage exist in the literature, could at least be partly confirmed: Rubinette (syn. Rafzubin) does indeed descend from Golden Delicious and Cox's Orange Pippin as speculated (Promo-Fruit 2009)

Table 2 Nine cultivars and three breeding lines for which one of the reported parents is confirmed (in bold) and the other is inconsistent with the marker data (italicised)

Accession	Source	Female parent	Male parent
Egeria	9	*Lobo*	**Red Spur Delicious**
Hokuto	4	*Fuji*	**Mutsu**
Kent	9,10	**Cox's Orange Pippin**	*Jonathan*
Lodel	9	*Lobo*	**Red Spur Delicious**
Medea	10	*Lobo*	**Red Spur Delicious**
Minjon	7	**Wealthy**	*Jonathan*
Rubinola	2,9,10	**Rubin**	*Prima*
Waleria	9,10	*Fantazja*	**Primula**
Winston	2,7,10	**Cox's Orange Pippin**	*Worcester Pearmain*
X-3189	4	**Winesap**	*X-2599*
Z207	4	*Jonathan*	**Anta.34.26**
Z638	4	*Golden Delicious*	**Anta.34.26**

Source of accessions coded as follows: 2 = University of Bologna, Italy, 4 = INRA, Angers, France, 7 = JKI, Dresden, Germany, 9 = Warsaw Agricultural University, Poland, 10 = RIPF Skierniewice, Poland

(Table 1). Haralson is assumed to be Malinda × Wealthy (Cabe et al. 2005) or Malinda × Ben Davis (http://www.applejournal.com/useall06.htm); we could confirm Wealthy as a parent but neither Malinda nor Ben Davis are present in the HiDRAS germplasm (Supplementary Table 2). Rajka is reported to be Sampion (syn. Champion) × Katka (=Jolana × Rubin) (Kruczyńsha and Rutkowski 2006); our data confirmed Sampion as a parent and Rubin as a grandparent, however Katka and Jolana were not tested (Table 3). Our data confirmed that the female parents of Realka and Reanda are Carola and Clivia, respectively (Supplementary Table 1), which does not agree with the recent data of Reim et al. (2009).

Other speculative pedigrees from the literature, for example that of Lady Williams (King et al. 1998), were incorrect; Granny Smith fits the pedigree as a possible parent but Jonathan does not, therefore Lady Williams was defined as a founder (Supplementary Table 3). King et al. (1998) also suggested that James Grieve was derived from either Pott's Seedling or Cox's Orange Pippin (Supplementary Table 1); the latter proved to fit the pedigree. Both samples of Dukat tested gave identical scores and could not have been derived from Golden Delicious, as reported by King et al. (1998) (Table 2).

Different accessions with the same name

Two different accessions of Priscilla (University of Bologna, Italy, and JKI Dresden, Germany), each represented by multiple samples, differed for 53% of their SSR loci. They were clearly two different accessions; however, both match with the reported pedigree for Priscilla (Supplementary Fig. 2) as they fit marker scores of one parent Starking Delicious and the founders of the second parent. The release of Priscilla was hampered by a mix-up during propagation (Janick, personal communication). Different full sibs may thus have been used for the release of Priscilla.

Tydeman's Early Worcester was also represented by several different accessions. Samples from JKI Dresden, Germany, and CRA Gembloux, Belgium, were identical and fit the pedigree of McIntosh × Worcester Pearmain; however a sample of Tydeman's Early, supposedly a synonym of Tydeman's Early Worcester (Morgan and Richards 2002), from RIPF Skierniewice, Poland, was inconsistent with the pedigree and differed greatly from the other samples. It was this sample of Tydeman's Early that fits the pedigree of Melfree (Supplementary Table 5). Seven accessions in Table 1, Akane, Discovery, Jester, Katja, Lord Lambourne, Merton Worcester and the previously mentioned Tydeman's Early Worcester, all show consistency with what must be the TTT Worcester Pearmain; however, although two samples labelled as Worcester Pearmain were in the HiDRAS germplasm set (from University of Bologna, Italy, and RIPF Skierniewice, Poland), neither of them proved to be TTT. The derived TTT Worcester Pearmain SSR profile was also consistent with Mio (Supplementary Table 2) but was inconsistent with Winston (Table 2). Subsequent analysis of non-HiDRAS germplasm identified a TTT sample of

Table 3 Ten cultivars and five breeding lines for which one or both of the reported parents are not tested, but for which the pedigree is (partly or entirely) confirmed through consistency of marker scores with earlier generations. Confirmed ancestors are in bold, lacking ancestors are in italics and pedigrees with inconsistent scores are underlined

Accession	Source	Female parent	Male parent
Burgundy	9	**Monroe**	*NY 18491* (**Macoun** × *Antonovka*)
Coop17	4	*Illinois_#2* (**Winesap** *open pollinated*)	*PRI 668-100* (Melrose × PRI 14-126)
Coop20	4	**Crandall**	*PRI 668-100* (Melrose × PRI 14-126)
Dalinbel	8	**Elstar**	*X-3191* (**Idared** × **Prima**)
Jerseymac	7,10	*NJ24* (*NJ 12* [Red Rome Beauty × **Melba**] × *NJ 117637* [*NJ 130* {**Wealthy** × **Starr**} × **Melba**])	**Julyred**
Moira	10	**McIntosh**	*PRI 47-147* (**Jonathan** × **F_2 26829-2-2**
Ozark Gold	7,9	**Golden Delicious**	*H1291* (Red Delicious × *Conrad* [*Ben Davis* × Jonathan])
Prima	4,7,9	*PRI 14-510* (**F_2 26289-2-2** × **Golden Delicious**)	*NJ 123249* (*NJ 12* [**Red Rome** × **Melba**] × *NJ 117637* [**Melba** × *NJ 130*{**Starr** × **Wealthy**}])
Rajka	9	**Sampion**	*Katka* (*Jolana* × **Rubin**)
Saturn	2	*PRI 1235-100* (*NJ 123249* [*NJ 117637* {**Melba** × *NJ 130* < **Wealthy** × **Starr** >} × *NJ 12* {Red Rome × **Melba**}] × PRI 47-147 [**Jonathan** × **F_2 26829-2-2**])	**Golden Delicious**
Spigold[3n]	9,10	*Red Spy* (**Northern Spy** × **Golden Delicious**)	**Golden Delicious**
Trent	4,10	**McIntosh**	*PRI 47-147* (**Jonathan** × **F_2 26829-2-2**)
X-2771	4	**Crandall**	*PRI 668-100* (Melrose × PRI 14-126)
X-6681	4	*X-3188* (**Winesap** × **X-2599**)	**X-3177**
X-6683	4	**X-4355**	*X-6820* (**Florina** x **Prima**)

Source of accessions coded as follows: 2 = University of Bologna, Italy, 4 = INRA, Angers, France, 7 = JKI, Dresden, Germany, 8 = CRA, Gembloux, Belgium, 9 = Warsaw Agricultural University, Poland, 10 = RIPF Skierniewice, Poland

3n: triploids

By inference, the pedigrees of Red Spy and the following breeding lines are also confirmed: NJ 12, NJ 130, NJ 117637, NJ 123249, PRI 14-510, PRI 47-147 and X-3191

Worcester Pearmain at INRA, France (Durel, personal communication).

Incorrect pedigrees

Eleven cultivars and four breeding lines had one incorrectly reported parent or a mistake in the pedigree, presented in Table 2 and Supplementary Table 4. Fully matching alternative pedigrees are proposed for three of these accessions (Dayton, Liberty and X-2773) (Supplementary Table 4). Macoun fits as a direct parent of Liberty; however, DNA of its other parent PRI 54-12 was unavailable. Both parents of PRI 54-12 were analyzed and while alleles of F_2 26829-2-2 fitted with Liberty, the other parent Jonathan showed many inconsistencies; the substitution of Wealthy instead of Jonathan gives a perfect fit.

Unravelling the pedigree of Dayton was more complex, as was frequently the case when both parents were breeding lines of which DNA samples were not available, rather than named varieties. One parent, NJ 123249, could be confirmed through a combination of its parents and grandparents; however, the other parent, PRI 1235-100, could only be partially confirmed through its pedigree. One of PRI 1235-100's parents, NJ 123249, could be confirmed by going back through four generations but the other, PRI 47-147 (Jonathan × F_2 26829-2-2), only partially fitted; replacing Jonathan with Golden Delicious, however, totally resolved the pedigree. To add to the complexity, PRI 47-147 is confirmed as Jonathan × F_2 26829-2-2 in the pedigree of Trent, so we can only assume that a different selection of the PRI 14 family (Golden Delicious × F_2 26829-2-2)

was used in the Dayton lineage (labeled as PRI 47-147a in Supplementary Table 4).

Golden Delicious and Jonathan and their related families PRI 14 and PRI 47 are the cause of further confusion in the HiDRAS germplasm. Golden Delicious was confirmed as the direct parent of the French breeding line X-2773, but the other supposed parent PRI 14-152, when evaluated indirectly through its assumed direct parents, did not fit X-2773 scores. F_2 26829-2-2 fitted with X-2773 as a grandparent, but Golden Delicious did not match; substituting Jonathan for Golden Delicious gave a full match, indicating that X-2773 actually derived from a member of the PRI 47 family (Supplementary Table 4).

Of the founders, F_2 26829-2-2 is thought to derive from two full sibs which no longer exist from the cross *M. floribunda* 821 × Rome Beauty. As a small proportion of the marker alleles of F_2 26829-2-2 cannot be explained by its two grandparents, one of the grandparents is likely to be false. Many SSR loci show an allele of *M. floribunda* 821, whereas 28% of the loci do not show an allele of either *M. floribunda* 821 or Rome Beauty. Rome Beauty is likely to be incorrect, as already postulated by Vinatzer et al. (2004). Considering the frequency of deviating alleles, Rome Beauty is still likely to be correct as one of the two full sibs.

None of the three examined progenies of PRI 668-100 matches its assumed parentage (Table 3), alternatives for which are currently under investigation.

Some samples proved to fit perfectly in some pedigrees (and were therefore TTT) but did not fit other specific pedigrees; for example, Braeburn-EMR fits with the offspring of the EMR breeding programme but not with the Braeburn offspring from other programmes; the Braeburn-EMR sample was shown to be NTTT. The Dutch sample of Priscilla, re-named Priscilla-NL, fits with the offspring of Dutch breeding programmes but with neither of the two previously mentioned accessions of Priscilla, nor with Priscilla's confirmed Delicious and PRI 610-2 parentage. The cross, from which Santana was derived, was made in 1978 on trees growing in The Netherlands. The scab resistant, *Vf*-carrying parent must therefore have been one of the earlier accessions from the American PRI breeding program. SSR scores of Santana match with Elstar, its other parent, completely, and all other alleles match with Priscilla-NL. Examination of the representation of these Priscilla-NL alleles in the entire HiDRAS germplasm established representation of *M. floribunda* 821 and McIntosh, and indicate the involvement of Jonathan and Lobo. McIntosh's representation is best explained through a direct contribution rather than through derived cultivars like Melba, Early McIntosh or Macoun, cultivars that occur in the pedigree of several other released PRI accessions. The *M. floribunda* 821 contribution may have been through F_2 26829-2-2 or F_2 26830, the latter giving a slightly better match. Santana might thus come from Elstar × (Lobo × [McIntosh × {Jonathan × F_2 26830}]), a pedigree which supports all examined non-Elstar alleles of Santana. All but two alleles of Collina, derived from the same parents as Santana, fit the postulated pedigree for Santana, indicating that this pedigree is not yet fully correct, but nevertheless close.

Polyploids

Eight polyploid samples were detected. For some samples, the origin of the ploidy mutation was clear from the comparison of the parental alleles; for example, both Jupiter (Supplementary Fig. 3) and Karmijn were the result of unreduced egg cells of Cox's Orange Pippin, and Jonagold and Mutsu were the result of unreduced egg cells of Golden Delicious (Table 1 and Supplementary Table 2, respectively). There were only seven SSR markers of Witos that appeared to be triploid; these fit with a partial unreduced egg cell of Fantazja (Table 1), specifically the bottom of linkage group (LG) 12 (SSR Hi07f01) and the entire LG13. Spigold (Table 3) is reportedly the result of a cross between Red Spy and Golden Delicious (Brooks and Olmo 1972). Red Spy was not available in the HiDRAS germplasm; however, as it is a red sport derived from Northern Spy (GRIN), allele sizes from Spigold were compared to those of Northern Spy. A double egg cell of Red Spy appears to be the predominant cause of the additional chromosomes in Spigold; however, there are additional alleles that cannot be explained by this study. Analysis of the Northern Spy red sport Red Spy should determine whether the different alleles found arose from mutations when Spigold was produced or whether they arose in Red Spy.

Cox's Orange Pippin is thought to be derived from the triploid Ribston Pippin (http://www.applejournal.

com/use.htm) which fits with our data; however we were unable to determine either parent of Ribston Pippin or the source of the polyploidy in this case. Obviously where the parentage is unknown, such comparisons cannot be undertaken, as in the case of Gravenstein.

Discussion

Plant material

Following the genotyping of the HiDRAS material and the validation of TTT examples of major founding cultivars, it is planned to set up core collections of this material within several national germplasm collections (e.g. France, Belgium, Switzerland and Great Britain). Such validated cultivar collections will be of great importance to molecular biologists, ensuring correct starting material for the development of further molecular tools for pre-selection. Breeders will also benefit from being able to access validated parental material and thus avoid what could be costly long-term mistakes arising from the use of incorrect germplasm. We expect that the marker data from this work will become publicly available in the near future in the European Cooperative Programme for Plant Genetic Resources (ECPGR) *Malus* Database currently hosted by the University of Reading (UK).

Data verification

With any project such as this where a large number of samples and data-points are involved, there are inevitable opportunities for errors. The robotics technology employed in distributing the samples to all parties will have greatly reduced any sampling error and the high level of validation of SSR scores thanks to the pedigree-based analysis would have highlighted any possible errors, leading to further verification of data (Patocchi et al. 2009b). Undoubtedly, structuring the work such that DNA samples were distributed to each genotyper and each SSR was tested at just a single location avoided all the difficulties of data alignment and integration. Thus we are confident that the data that is presented in this publication is accurate.

It should be noted that there are still a few alleles present in the data files that are represented by just one or two genotypes and that have just 1 bp difference from flanking alleles. However, following the in-depth survey of the accessions which had questionable TTT, most putative errors remaining are with accessions that have no clear pedigree relationships.

Polyploids

Whilst the partial triploid Witos fits its pedigree, its genotype appears unusual as all except one of its triploid loci are on the same linkage group. One possible explanation was that the data could have arisen as an artefact of the multiplex; however, the data are from several independent multiplexes. An examination of the cytology of Witos could clarify the situation. Our level of detected polyploids was much lower than that reported in previous molecular studies in apple (Pereira-Lorenzo et al. 2007; Van Treuren et al. 2010); it is possible that some of the cultivars previously reported as triploid could in fact have been misclassified due to additional alleles resulting from gene duplications and homoeologous chromosome segments. No evidence of chimaeras was found in this study although they are known to exist in apple (Dermen 1955) and could result in the occasional triallelic locus.

Pedigree validation

It is impressive that such a high proportion of the parentages of these cultivars are correct and well documented when many would have been produced with only a small measure of protection from outcrossing. Breeders will be reassured to have this level of confirmation of the accuracy of standard breeding techniques.

It is also, however, not surprising that with the movement of cultivars around the world to different collections, propagation errors have occurred resulting in different genotypes with the same name, as was the case with 'Priscilla'. Further literature research on the release of Braeburn revealed that there were originally two very similar 'Braeburn-type' apples that were tested concurrently in New Zealand (McKenzie 1979); it would appear that the tree accessed into the National Fruit Collection, Brogdale,

England, was not the one that was eventually commercialised.

The power of the pedigree analysis system for both pedigree validation, calculation of heritabilities from phenotypic data (Kouassi et al. 2009) and QTL mapping is inevitably decreased by the presence of open pollinations in the pedigree. A gap in the pedigree can be another problem, for example if the links to the founders were lost within the pedigree due to lack of DNA for various successive generations (e.g. pedigree of Prima) or if individuals were not available or not tested. The use of NTTT parentages is another factor that decreases the power of pedigree-based QTL mapping. This confirmation and re-defining of pedigrees and identification of TTT accessions will thus enforce PBA approaches. Moreover, this information will be particularly useful to breeders, especially with the progress in application of marker-assisted breeding. Being able to confidently source accessions of confirmed pedigree for both making crosses and further molecular analyses will improve the transferability of data between different groups and benefit the apple genomics community as a whole.

Acknowledgments Charles Simon of the USDA-ARS of the Plant Genetic Resources Unit, Geneva is acknowledged for supplying leaf samples of PR14-126, PR14-152 and Starr, and Stefano Tartarini of the University of Bologna and Margaret Korbin of RIPF for supplying DNA of all other cultivars and breeding lines not cited in the tables. This publication was carried out with the financial support from the Commission of the European Communities (contract no. QLK5-CT-2002-01492), Directorate-General Research—Quality of Life and Management of Living Resources Programme. It does not necessarily reflect the Commission's views and in no way anticipates its future policy in this area. Its content is the sole responsibility of the publishers.

References

Antofie A, Lateur M, Oger R, Patocchi A, Durel CE, van de Weg WE (2007) A new versatile database created for geneticists and breeders to link molecular and phenotypic data in perennial crops: the AppleBreed DataBase. Bioinformatics 23:882–891

Bink MCAM, Uimari P, Sillanpää MJ, Janss LLG, Jansen RC (2002) Multiple QTL mapping in related plant populations via a pedigree-analysis approach. Theor Appl Genet 104:751–762

Bink MCAM, Boer MP, ter Braak CJF, Jansen J, Voorrips RE, van de Weg WE (2008) Bayesian analysis of complex traits in pedigreed plant populations. Euphytica 161:85–96

Brooks RM, Olmo HP (1972) Register of new fruit and nut varieties, 2nd edn. Berkeley, Los Angeles, University of California Press, p 101

Cabe PR, Baumgaten A, Onan K, Luby JJ, Bedford DS (2005) Using microsatellite analysis to verify breeding records: a study of 'Honeycrisp' and other cold-hardy apple cultivars. HortScience 40:15–17

Dermen H (1955) A 2-4-2 chimera of McIntosh apple. J Wash Acad Sci 45:324–327

Gianfranceschi L, Soglio V (2004) The European Project HiDRAS: innovative multidisciplinary approaches to breeding high quality disease resistant apples. Acta Hortic 663:327–330

Gianfranceschi L, Koller B, Seglias N, Kellerhals M, Gessler C (1996) Molecular selection in apple for resistance to scab caused by *Venturia inaequalis*. Theor Appl Genet 93: 199–204

King G, Brown L, Ryder C, Periam N (1998) Microsatellite markers for accession identification, pedigree analysis and assessment of allelic diversity in *Malus* genetic resources. ECPGR report of a working group on *Malus*/*Pyrus*—1st meeting (May 1997, Dublin). In: Maggioni L, Janes R, Hayes A, Swinburne T, Lipman E (eds) Biodiversity international publications, pp 104–113

Kouassi AB, Durel CE, Costa F, Tartarini S, van de Weg E, Evans K, Fernández-Fernández F, Govan C, Boudichevskaja A, Dunemann F, Antofie A, Lateur M, Stankiewicz-Kosyl M, Soska A, Tomala K, Lewandowski M, Rutkovski K, Zurawicz E, Guerra W, Laurens F (2009) Estimation of genetic parameters and prediction of breeding values for apple fruit-quality traits using pedigreed plant material in Europe. Tree Genet Genomes 5:659–672

Kruczyńsha DE, Rutkowski KP (2006) Quality and storage of Czech scab resistant apple cultivars. Phytopath Pol 39:53–61

McKenzie D (1979) The two types of Braeburn. Orchard N Z 1979:182

Morgan J, Richards A (2002) The new book of apples. Ebury Press, London

Patocchi A, Bigler B, Koller B, Kellerhals M, Gessler C (2004) *Vr2*: a new apple scab resistance gene. Theor Appl Genet 109:1087–1092

Patocchi A, Fernández-Fernández F, Evans K, Gobbin D, Rezzonico F, Boudichevskaia A, Dunemann F, Stankiewicz-Kosyl M, Mathis F, Durel CE, Gianfranceschi L, Costa F, Toller C, Cova V, Mott D, Komjanc M, Barbaro E, Kodde L, Rikkerink E, Gessler C, van de Weg WE (2009a) Development and test of 21 multiplex PCRs composed of SSRs spanning 80% of the apple genome. Tree Genet Genom 5:211–223

Patocchi A, Fernández-Fernández F, Evans K, Silfverburg-Dilworth E, Matasci CL, Gobbin D, Rezzonico F, Boudichevskaia A, Dunemann F, Stankiewicz-Kosyl M, Mathis F, Durel CE, Soglio V, Gianfranceschi L, Costa F, Toller C, Cova V, Mott D, Komjanc M, Barbaro E, Voorrips RE, Rikkerink E, Yamamoto T, Cevik V, Gessler C, van de Weg WE (2009b) Development of a set of apple SSRs markers spanning the apple genome, genotyping of HiDRAS plant material and validation of genotypic data. Acta Hortic 814:603–608

Pereira-Lorenzo S, Ramos-Cabrer AM, Díaz-Hernández MB (2007) Evaluation of genetic identity and variation of

local apple cultivars (*Malus* × *domestica* Borkh.) from Spain using microsatellite markers. Genet Resour Crop Evol 54:405–420

Promo-Fruit (2009) http://www.promo-fruit.ch/de/rubinette.php

Reim S, Flachowsky H, Hanke M-V, Peil A (2009) Verifying the parents of the Pillnitzer apple cultivars. Acta Hortic 814:319–323

Stark P (1974) The 'Ozark Gold' apple. J Am Pomol Soc 28:20–21

Van de Weg WE, Voorips RE, Finkers R, Kodde LP, Jansen J, Bink MCAM (2004) Pedigree genotyping a new pedigree-based approach of QTL identification and allele mining. Acta Hort 663:45–50

Van Treuren R, Kemp H, Ernsting G, Jongejans B, Houtman H, Visser L (2010) Microsatellite genotyping of apple (*Malus* × *domestica* Borkh.) genetic resources in the Netherlands: application in collection management and variety identification. Genet Resour Crop Evol 57:853–865

Vinatzer BA, Patocchi A, Tartarini S, Gianfranceschi L, Sansavini S, Gessler C (2004) Isolation of two microsatellite markers from BAC clones of the *Vf* scab resistance region and molecular characterization of scab-resistant accessions in *Malus* germplasm. Plant Breeding 123:321–326

Voorrips RE (2007) Pedimap: software for visualization of genetic and phenotypic data in pedigrees. Plant Research International, Wageningen, the Netherlands http://www.plantbreeding.wur.nl/UK/software_pedimap.html