



Cultivar fingerprinting and SNP-based pedigree reconstruction in Danish heritage apple cultivars utilizing genotypic data from multiple germplasm collections in the world

Bjarne Larsen · Nicholas P. Howard · Caroline Denancé · Charles-Eric Durel · Carsten Pedersen · Jonas Skytte af Sättra · Larisa Garkava-Gustavsson · Michela Troggio · Eric van de Weg

Received: 7 May 2024 / Accepted: 30 July 2024
© The Author(s) 2024

Abstract Heirloom Danish apple cultivars are historically and pomologically important, part of the cultural heritage, and have valuable adaptation to regional climate conditions. However, lack of information about their genetic identity and pedigree relatedness with other cultivars hampers proper cultivar identification, germplasm curation, genebank management, and future regional breeding efforts. Many Danish apple cultivars are maintained in the national collection “The Pometum”, maintaining around 850 apple accessions. Additional material is maintained in public or private Danish collections. However, no information exists regarding genotypic duplicates

between these collections and germplasm collections in other countries, pedigree inferences across collections, and genotypically unique accessions at the genebank level.

To provide such information, 976 accessions from Denmark were genotyped with simple sequence repeat (SSR) markers and the Illumina Infinium 20K single nucleotide polymorphism (SNP) array. The resulting genotypic data were compared to large databases of genotypic data from germplasm collections in multiple countries to identify genotypic duplicates and conduct pedigree reconstruction. The germplasm maintains 305 unique genotypic profiles which were not found in other germplasm collections. The study exposed previously unknown synonyms, accessions not true-to-type, and novel pedigree relationships involving accessions from multiple collection sites.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-024-02104-1>.

B. Larsen (✉) · N. P. Howard · E. van de Weg
Department of Plant Breeding, Wageningen University and Research, Wageningen, The Netherlands
e-mail: bjl@plen.ku.dk

B. Larsen · C. Pedersen
Department of Plant and Environmental Sciences,
University of Copenhagen, Frederiksberg C, Denmark

N. P. Howard
Fresh Forward Breeding and Marketing B.V., Huissen,
The Netherlands

N. P. Howard
Institut für Biologie und Umweltwissenschaften, Carl von
Ossietzky University, Oldenburg, Germany

C. Denancé · C.-E. Durel
Univ Angers, Institut Agro, INRAE, IRHS, SFR
QUASAV, F-49000 Angers, France

J. S. a. Sättra · L. Garkava-Gustavsson
Department of Plant Breeding, Swedish University
of Agricultural Sciences, Alnarp, Sweden

M. Troggio
Research and Innovation Centre, Fondazione Edmund
Mach, San Michele all'Adige, Trento, Italy

The most frequent parents of Danish germplasm were ‘Hvid Vinter Pigeon’ and ‘Cox’s Orange Pippin’ whereas ‘Reinette Franche’ was the most common grandparent. The accession-level information will benefit germplasm curation, cultivar identification, genebank management, and future breeding efforts, and shed new light on cultivar history and origin.

Keywords Cultivar identification · Cultivar parentage · Cultivar provenance · Genebank cross-collection comparison · Heirloom cultivars

Introduction

Apple (*Malus domestica* Borkh.) cultivars are an important and cherished part of the cultural heritage of Denmark. Apple cultivation likely dates to at least the Middle Ages (1000–1536 A.D.) in Denmark, though cultivars available today likely originated during the past few centuries (Bredsted 1893; Lange 1975). Many of the cultivars were grown in a narrow geographical region where they fitted local traditions of culinary uses, seasonal fresh consumption, or well-defined decoration purposes. Heirloom cultivars are frequently well-adapted to regional climate conditions and maintain traits that may be rare or absent in commercial cultivars, making them a fundamental resource for breeding efforts and related genetic research. In addition, numerous heirloom cultivars have been identified as pedigree ancestors of current market cultivars and elite individuals (Muranty et al. 2020; Howard et al. 2021a), making them valuable in pedigree reconstruction research and breeding.

Many Danish apple cultivars are today maintained in the germplasm collection “The Pometum” at the University of Copenhagen, Denmark, which is part of The Nordic Genetic Resource Center (NordGen) under the Nordic Council of Ministers. The Pometum was founded in 1863 and maintains the largest collection of historical apple cultivars in the Nordic region, including a “heritage collection” of approximately 273 accessions with recorded or speculated origins in Denmark (Bredsted 1893; Matthiessen 1913, 1924; Pedersen 1950, 1966) in addition to a set of around 457 accessions with various origins but generally representing germplasm that historically has been cultivated in Northern Europe. In addition, a trial collection of 113 accessions obtained from private

Danish collections maintained at The Pometum were included in this study. Other Danish heirloom germplasm is kept and preserved by public institutions such as the Bornholm Museum, the National Museum of Denmark, as well as private associations and personal orchards in Denmark. Such public and private collections are a potential resource for unique accessions of culturally significant and regionally adapted germplasm not maintained at The Pometum. Genotypic information could help identifying genotypic duplicates between collections and identify genetically unique accessions present in specific collections.

Part of The Pometum collection has previously been investigated for genotypic duplicates, pedigree relationships, and genetic structure using simple sequence repeat (SSR) markers and genotyping-by-sequencing (GBS) data (Larsen et al. 2017, 2018). However, a major limitation of these investigations was that they only considered accessions within The Pometum collection, making comparisons with accessions maintained in other collections in Denmark and other parts of the world not possible. Inclusion of genotypic data from multiple germplasm collections would allow for the identification of genotypically unique accessions maintained in each germplasm collection and expose duplicates and wrongly labelled accessions. It would also enable large-scale pedigree reconstruction by potentially including ancestors that are not present in the collection themselves, reveal differences in the frequency of parental genotypic profiles among germplasm collections, and in turn benefit preservation efforts of Danish apple cultivars by including genotypically unique accessions or important pedigree ancestors of heritage cultivars in The Pometum collection.

Cross-collection genotypic comparison of apple cultivars has been enabled through ongoing large international collaborative projects (Howard et al. 2018; Denancé et al. 2020), but The Pometum and other Danish collections have not yet benefited from such genotypic comparison with other germplasm collections in the world. An initial cross-collection comparison integrated genotypic data from multiple collections in Europe, enabling the comparison of accessions within and across collection sites based on 16 SSRs as described by Urrestarazu et al. (2016). This culminated in the ongoing *Malus* UniQue genotype (MUNQ) coding system (Denancé et al. 2020; Muranty et al. 2020) in which each unique SSR

genotypic profile was ascribed to a unique MUNQ code. The MUNQ project currently allows the comparison of more than 20,000 SSR-genotyped, MUNQ attributed accessions. The MUNQ code is a highly efficient tool for the identification of duplicates and synonyms across collection sites and facilitates communication about genetically unique accessions and accession identity. While the relatively small number of SSRs are highly efficient for cultivar identification, they give much less certainty and precision for pedigree validation and reconstruction and for elucidation of distant genetic relations. Consequently, single nucleotide polymorphism (SNP) array data has been preferred in recent large-scale apple diversity and pedigree reconstruction projects (Howard et al. 2018; Muranty et al. 2020), as they allow for high certainty in the confirmation of pedigree relationships, allow integration of data obtained from multiple sources, provide information on ploidy and aneuploidy (Chagné et al. 2015), as well as the ordering of parent–offspring duo relationships as in Howard et al. (2022).

The goal of this study was to provide accession level information on the genotypic identities, synonyms, and pedigrees of Danish apple germplasm held in The Pometum and other Danish collections. This was approached by (a) identifying unique genotypic profiles, genotypic duplicates, and synonyms by MUNQ attribution through the comparison of SSR profiles to the SSR profiles of the large ongoing collaborative apple SSR fingerprinting project (Urrestarazu et al. 2016; Denancé et al. 2020; Muranty et al. 2020), and (b) to reveal pedigree relations of the Danish germplasm through integrating genome-wide genotypic data from the 20K Infinium® SNP array (Bianco et al. 2014) with the ongoing large-scale collaborative apple pedigree reconstruction project (APR project, Howard et al. 2018). The overall purpose of this study was to provide information useful for germplasm collection curation and management, for historical and pomological research, and for future cultivar development.

Materials and methods

Germplasm

A set of 976 apple accessions were included in this study (Table S1). This set is composed of all apple

accessions maintained in the genebank collection “The Pometum” (University of Copenhagen, Taastrup, Denmark) (N=843), as well as various other accessions sampled from private collections, nurseries, gardens, and public open-air museums in Denmark (N=133). The Pometum is divided into the Danish heritage collection (N=273), a non-heritage collection (N=457), and a trial collection of germplasm more recently collected in Denmark (N=113).

SSR genotyping and MUNQ assignment

This study included SSR data for 976 accessions, of which 442 accessions came from Larsen et al. (2017). The remaining 534 accessions were genotyped for this study. Leaf sampling, DNA extraction, and SSR genotyping were performed as described by Larsen et al. (2017), where a core set of 15 SSR markers was used for genotyping allowing to merge the Danish dataset into the SSR-based database of the MUNQ project (Denancé et al. 2020; Muranty et al. 2020) using a common set of 37 diploid and triploid accessions to standardize calls and ascribe each accession to a MUNQ code (Table S1). This allowed the Danish accessions to be compared to over 20,000 unique *Malus* genotypic profiles in the MUNQ dataset.

SNP array genotyping

Genotyping with the Illumina Infinium 20K SNP array (Bianco et al. 2014) was conducted for accessions representing each unique MUNQ value that had not been previously genotyped on the 20K SNP array in the APR project (Howard et al. 2018), or the Affymetrix Axiom 480 K SNP array (Bianco et al. 2016) in Muranty et al. (2020) as the SNP data representing those MUNQs was made available for the present study.

Accessions chosen for 20K SNP array genotyping were harvested as young leaves and dried immediately after sampling, either by freeze drying or on silica gel. Subsequent DNA extraction was performed using a modified version of the protocol of Fulton et al. (1995). SNP genotyping was performed according to the standard Illumina protocol outlined in Chagné et al. (2012). Intensity data were analysed in GenomeStudio v2.0 (Illumina Inc.) and curated

following the workflow described in Vanderzande et al. (2019) and with the modifications, cluster definitions, and genetic positions for the “robust” SNPs identified in Howard et al. (2021c), but with some minor modifications to use the same set of 10,321 SNPs that were previously used in Volk et al. (2022). These modifications included the inclusion and exclusion of a small number of SNPs. All genotypic profiles in this study were compared to those in the APR project.

SNP data analyses

Ploidy levels were identified using the *B*-allele (BAF) frequency distribution histogram of each sample in GenomeStudio (GS) v2.0 (Illumina Inc.) (Chagné et al. 2015). The BAF distributions were used to confirm good DNA and iScan quality sensu Vanderzande et al. (2019). Samples were evaluated for the presence of aneuploidy by plotting the BAF frequency of each SNP against their cumulative genetic position as in Vanderzande et al. (2019).

The SNP data was integrated into the SNP dataset of the ongoing APR project. This allowed the Danish germplasm to be compared to over 5000 unique *Malus* genotypic profiles from accessions sampled from over 50 apple collections from various parts of the world. Part of these genotypic profiles were genotyped on the Axiom Affymetrix® 480 K SNP array (Muranty et al. 2020), for which integrated SNP data came from Howard et al. (2021c).

Accessions with more than 99.5% identical calls were considered genotypic duplicates (as per Howard et al. (2021c)) and were identified using a custom R script (R Core team 2022).

Mendelian exclusion methods were used to identify the pedigree relationships among individuals in this study. These methods have been well established and successful in previous research with SNP array data in apple (e.g. Vanderzande et al. 2017, van de Weg et al. 2017, Howard et al. 2018, Vanderzande et al. 2019, Muranty et al. 2020, Skytte af Sättra et al. 2020, Luby et al. 2022, Gilpin et al. 2023, Konjić et al. 2023).

Parent–offspring relationships were identified as described in Vanderzande et al. (2019) and Howard et al. (2023) using a threshold of 0.1% for Mendelian inconsistent errors among the highly curated

genotypic profiles in the ARP project. Parent–offspring relationships involving tetraploid genotypic profiles were identified using Mendelian inconsistent errors using homozygous calls only. Parents of Danish germplasm that were not present in the Danish dataset but that came from the APR project were recorded in Table S2 and their accession source information is provided in Table S3.

Phased SNP data was generated using FlexQTL™ v.0.9907 software (Bink et al. 2014) using pedigree information from the APR project, and segregation information from full-sib families including those used in the creation of the iGLMap (DiPierro et al. 2016). The phased SNP data was used for the curation of marker calls for non-Mendelian errors (Vanderzande et al. 2019), ordering parent–offspring relations (Howard et al. 2022), and, whenever relevant, for assessing distant genetic relationships using shared haplotype information (Howard et al. 2021a). Unordered parent–offspring duos were ordered whenever at least four direct relationships were available for at least one individual utilizing the parent–offspring order resolution (POR) tests, POR-1 and POR-2 using an Excel tool as described in Howard et al. (2022). Half-sibling (HSIB) and grandparent–grandchild (GPGC) pedigree relationships were examined using “summed potential lengths of shared haplotype” (SPLoSH) information obtained through a Python script as previously described (Howard et al. 2021a), using 20 cM as a minimum length for individual shared haplotype fragments. Unreduced gamete-donating parents (UGDPs) of triploids were identified or imputed using an excel tool as in Howard et al. 2023.

Groups of HSIBs consisting of at least five members sharing an unknown common parent were identified and the virtual genotypic profile of their common parent was imputed following the methods of Howard et al. (2023). Imputed genotypic profiles from the APR project that were parents of accessions in this study were recorded. These were previously not published. They were named “Unknown Founder X”, where “X” was a sequential number, following the first “Unknown Founder” genotypically described in Howard et al. (2021a) (Table S4). A singular genotypic profile imputed in this study was given the name “Unknown Danish Founder”. Additional dummy genotypic profiles were recorded of instances where a pair of grandparents were

identified that could account for the genotypic profile of an ungenotyped parent of an individual, as previously described in Howard et al. (2021a). These were named “UP_” (standing for “Unknown Parent”) followed by the name of their offspring. In some cases, the cultivar identity of this dummy profile was known from pedigree records but was ungenotyped. In those cases, the recorded name was reported instead of the “UP_” naming convention. Finally, in two cases the unknown, but deduced genotypic profiles of the un-reduced gamete donating parent of triploids (see Howard et al. 2023 for more information) were found to also be parents of diploid cultivars. In these cases, the profiles were only partially imputed and given the temporary name “UGDP_” followed by the name of the triploid used for the original deduction of the genotypic profile of the individual.

Potential GPGC relationships through an unknown parent were investigated for selected cultivars with major historical importance. Possible GPGC relationships were considered when individuals had a SPLoSH-based estimated coefficient of relatedness of approximately 0.25. GPGC relationships were considered probable when the extended haplotypes of the prospective grandchild were composed of extended haplotypes of the prospective grandparent on approximately at least half of its chromosome pairs, with the extended shared haplotypes covering roughly at least 25% of chromosome ends, and with evidence of some of these extended chromosomes being recombinant haplotypes from the prospective grandparent, as previously demonstrated in Howard et al. (2021b).

A network of parent–offspring relationships among cultivars recorded as originating in Denmark was generated using the package ggraph (Pedersen 2024) in R (R core team 2024).

Evaluation of true-to-typeness

Whenever needed, true-to-typeness (TTT) was evaluated by comparing observed phenotypes with literature descriptions as far as sufficient cultivar descriptions exist allowing cultivar identification (e.g. Bredsted 1893; Matthiessen 1913, 1924; Pedersen 1950), and by comparing confirmed or reconstructed pedigree relationships with pedigree records from the literature. In case an accession was found to be inconsistent with pedigree records, provenance, or

genotypic information it was recorded as not true-to-type (NTTT). For these NTTTT accessions, the suffix (NTTT) was added to the given preferred name (Muranty et al. 2020) (Table S1). For example, three genotypic profiles represented accessions labelled as ‘Broholm Rosenæble’ (‘Broholm’). One genotypic profile (MUNQ 5954) represented the true cultivar ‘Broholm Rosenæble’ and was therefore given the preferred name ‘Broholm Rosenæble’. The second genotypic profile (MUNQ 6350) was genetically identical to an accession labelled ‘Grevinde Ahlefeldt’. The accession labelled ‘Bronholm’ was therefore recorded as NTTTT while the best fitting name (the preferred name) for MUNQ 6350 was ‘Grevinde Ahlefeldt’. The third genotypic profile, (MUNQ 6091) represented a genetically unique accession. The identity of this genotypic profile could not be identified, so it was given the preferred name of its recorded name but with (NTTT) as a suffix.

A “preferred name” was assigned to each unique genotypic profile (Table S5) based on an evaluation of validated synonyms across the MUNQ project and the APR project, synonyms recorded in relevant pomological literature, SNP confirmed pedigrees, and when needed, a comparison of the accession phenotype with pomological cultivar descriptions. The preferred names chosen were those that were considered to most accurately reflect the historical provenance or in some cases the most commonly used name in (western) literature and pomological records.

Results

Genotypic duplication among accessions held in Denmark There were 667 (68%) unique genotypic profiles among the 976 studied accessions (Tables S1, S5). In total, 305 genotypic profiles were genotypically unique to Danish germplasm, of which 245 genotypic profiles were represented by single accessions in Danish collections. The germplasm in this study included 307 genotypic profiles that had not been previously identified in other collections within the SSR based MUNQ project and 342 genotypic profiles which were not identified from other sources in the SNP array-based APR project (Table S5). Seventeen accessions were matching the genotypic profile of six rootstock clones (‘M.4’, ‘M.7’, ‘M.9’, ‘M.16’, ‘MM.106’, and ‘MM.111’).

The Pometum maintains 843 accessions representing 623 (74%) genotypic profiles. There were 148 unique genotypic profiles in The Pometum heritage collection which were not identified in other collections within the MUNQ project and the APR project (Tables S1, S5).

Ploidy and aneuploidy

The dataset included 77 triploid genotypic profiles (7.9%) of which 24 were unique for the Danish germplasm within the MUNQ and APR projects. One tetraploid accession, ‘Alfa 68’ was identified (Figure S1). A possible case of partial or complete tetrasomy was observed for the proximal part of chromosome 9 on the accession ‘Lord Rosebery’ (Figure S2). The remaining samples were diploid without any identified chromosomal abnormalities.

Identity of accessions

Forty-three accessions were identified as NTTT (Table S1). The MUNQ results allowed for the identification of 35 NTTT accessions (labelled as “NTTT-MUNQ” in Table S1). In 17 of those cases, the MUNQ code corresponded to a known rootstock. In the other 18 cases, genotypically identical accessions helped to resolve the correct identities of the accessions. For example, four differentially named accessions from The Pometum, ‘Biesterfelder Reinet’, ‘Casseler Reinet’, ‘Claygate Pearmain’, and ‘Newton Wonder’, matched with the genotypic profile for ‘Kasseler Renette’ (MUNQ 629, syn. ‘Dutch Mignonne’ and ‘Reinette de Caux’) from the National Fruit Collection (NFC), Julius Kühn-Institut (JKI), Ökowerk, and INRAE. Thus, the identity of MUNQ 629 was considered to be ‘Kasseler Renette’ as suggested by Denancé et al. (2020) and Muranty et al. (2020).

Three accessions were determined to be NTTT via pedigree reconstruction (labelled as “NTTT-ped” in Table S1). In the first case, an accession labelled ‘Brøndæble’, had the SNP deducted parentage ‘James Grieve’ × ‘Maglemær’. However, ‘Brøndæble’ was described as a very old cultivar already by Matthiessen (1913), making it older than ‘James Grieve’, which originated in the last part of the nineteenth century (Smith 1971). Thus,

‘James Grieve’ could not be a parent of the true ‘Brøndæble’ of Matthiessen (1913). The second case was an accession named ‘Antonius’, which had the recorded parentage ‘Wealthy’ × ‘Cox’s Orange Pippin’ (Dullum 1961). However, the genotypic profile of this accession matched ‘Merton Beauty’ (‘Ellison’s Orange’ × ‘Cox’s Orange Pippin’) from NFC, JKI, INRAE, and The Pometum (MUNQ 2694). Thus, the correct identity of this genotypic profile was concluded to be ‘Merton Beauty’. The third case was an accession labelled ‘Ceres’. This cultivar originated in the Netherlands and had the recorded pedigree ‘Cox’s Orange Pippin’ × ‘Jonathan’ (Veldhuyzen van Zanten 1957) and matches the MUNQ 2118. The Pometum accession of ‘Ceres’ (MUNQ 5392) instead had the pedigree ‘Cox’s Orange Pippin’ × ‘Reinette Rouge Etoilee’ and was considered NTTT.

There were five cases of genotypic profiles for which the phenotype did not match the pomological description in the literature (labelled “NTTT-pheno”, Table S1). One example was MUNQ 5967, represented by two accessions labelled ‘Ydunsæble’, which did not match the pomological description of ‘Ydunsæble’ of Matthiessen (1913). Another example is The Pometum accession labelled ‘Grand Richard’ (MUNQ 376). The phenotype of this accession did not match the literature description of ‘Grand Richard’, a regional, historical cultivar first described in Hirschfeld (1788), which was not identified in this study. Instead, this accession matched the genotypic profile of ‘Blenheim Orange’ in the NFC, JKI, and INRAE collections (Denancé et al. 2020).

There was one case of four differently labelled accessions constituting the same genotypic profile where the correct identity was unclear due to lacking passport information and insufficient pomological cultivar descriptions. The four Pometum accessions, ‘Alsisk Citronæble’, ‘Æbeltoftæble’, ‘Vejløæble’, and ‘Møllers Venus’ (syn. ‘Venus’) (MUNQ 5972) were genetically identical. This genotypic profile was only identified in Danish germplasm and the original cultivar descriptions of the four cultivars made by Matthiessen (1913) have major similarities but do not allow for cultivar identification. The correct identity therefore remains uncertain. The name ‘Venus’ was chosen as the preferred name as it is the name under which the genotypic profile seems most commonly recognized in Denmark.

Ordering of parent–child duos

All parent–child duos identified in this study (Table S5) could be ordered using the POR-tests except the identified parent–offspring relationship between ‘Rolund’ (MUNQ5965) and ‘Frøbjergæble’ (MUNQ6129). This parent–offspring relation was not recorded in Table S5.

HSIB groups

A group of 17 HSIBs sharing an unknown common parent was identified (Table 1). This unknown individual has been tentatively named the “Unknown Danish Founder”. Both alleles were imputed for 10,272 SNPs of the “Unknown Danish Founder” (99.5% of SNPs included in this study) (Table S6). The 17 half-siblings were all unique to the germplasm of this study, except for the accessions ‘Drejæble’ and ‘Arreskov’, which were also present in the germplasm collection at the Swedish University of Agricultural Sciences (SLU).

A group of 11 HSIBs was identified as offspring from an unknown common parent that was previously imputed and named the “Unknown Founder 1”

(Table S5) (Howard et al. (2021b)). Another group of four HSIBs was identified that comprised the accessions ‘Askeæble’, ‘Kundbyæble’, ‘Marselisborg Sommeræble’, and ‘Ringkloster Kammerjunker’. These share an unknown common parent that previously has been imputed and given the name “Unknown Founder 3” through the APR project.

Pedigree validation and reconstruction results

Both parents were identified for 256 genotypic profiles, and one parent was identified for 185 genotypic profiles (Table S5). In addition, both parents were imputed for two genotypic profiles and one dummy parent was imputed for 56 genotypic profiles. There were 119 parents involving 205 pedigree relations that were not present in the Danish germplasm. The most common parent in this study was ‘Cox’s Orange Pippin’, whereas ‘Hvid Vinter Pigeon’ (MUNQ 1509) was the most frequent parent among genotypic profiles unique to Denmark (Fig. 1). ‘Hvid Vinter Pigeon’ was the parent of in 26 genotypic profiles in this study (Table S5) (Fig. 2). ‘Reinette Franche’ was the most frequent grandparent of Danish germplasm (Fig. 1), but the

Table 1 Half-sibling group of 17 genotypic profiles sharing a common unknown parent, tentatively named the “Unknown Danish Founder”

Preferred name	Other parent	Recorded geographic origin in Denmark	Recorded year of origin
Arreskov	–	Brændelydinge	> 1850
Bødkeræble	Prinzenapfel	Tjele	~ 1860
Bøgelundsæble	–	–	–
Bogense kålæble	–	–	–
Bredæble (NTTT)	–	–	–
Drejæble	Elmelund	Odense	> 1900
Faster Anes æble	–	–	–
Fru Glahns Æble	Calville_Blanc_dHiver	Strynø	~ 1865
Gråsten fra 1905	Reinette_de_Hollande	–	–
Gul Pigeon	Hvid_Vinter_Pigeon	–	–
Herschendsgave	–	Herschendsgave	> 1850
Langæble	–	–	–
Langeland Hvid Pigeon (2)	Hvid_Vinter_Pigeon	–	–
Nonnetit Bastard	Hvid_Vinter_Pigeon	Kolding	> 1850
Ormslevæble	–	Ormslev	> 1900
Rifbjerg Gylling	–	Langeland	~ 1800
Rifbjerg Skarlagan Pearmain	Broholm_Rosenaeble	Langeland	~ 1850

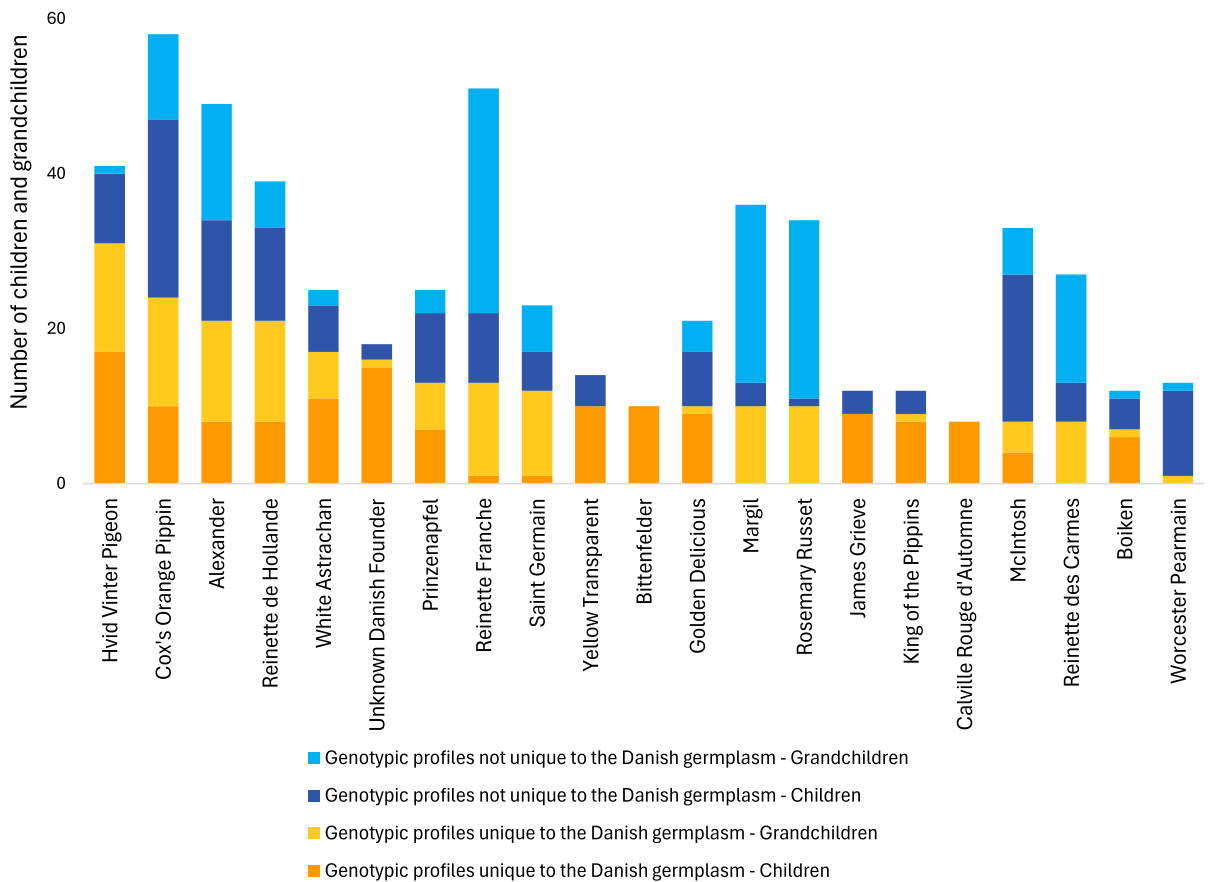


Fig. 1 The most common pedigree ancestors of Danish germplasm. The figure shows the count of parents and grandparents among 305 genotypic profiles that were unique to Denmark as well as count of parents and grandparents among 362 genotypic profiles that were present in the Danish dataset but also identified in other germplasm collections within the MUNQ and APR projects. The ancestors are sorted by their summed

number of children and grandchildren among genotypic profiles that were unique to Denmark. Cultivars included are having at least five direct offspring in the germplasm unique to Denmark, at least 10 direct offspring among germplasm not unique to Denmark, or at least 10 grandchildren in the germplasm unique and not unique to Denmark. The numbers underlying the figure are provided in Table S7

cultivar had no direct offspring in The Pometum heritage collection. No pedigree relationship was identified for 203 genotypic profiles, not considering dummy parents (Table S5). The most common parents of cultivars considered of Danish origin are all of non-Danish origin or of unknown origin (Fig. 3).

A relatively close relationship was identified between ‘Hvid Vinter Pigeon’ and ‘Passe Pomme Rouge’ using shared haplotype data (SPLoSH=842 cM). This SPLoSH value is indicative of a possible HSIB or GPGC relationship (Howard et al. 2021a). The exact relationship between the two cultivars was not clear because possible recombinant haplotypes of ‘Hvid

Vinter Pigeon’ were identified in ‘Passe Pomme Rouge’ and vice versa.

The historically important triploid cultivar Gravensteiner had no SNP validated parent–offspring relationship nor any grandparent–grandchild relationships identified in this study. The closest relative of ‘Gravensteiner’ was revealed to be the diploid cultivar ‘Prinzenapfel’ (SPLoSH=947 cM), but the exact relationship between them could not be determined.

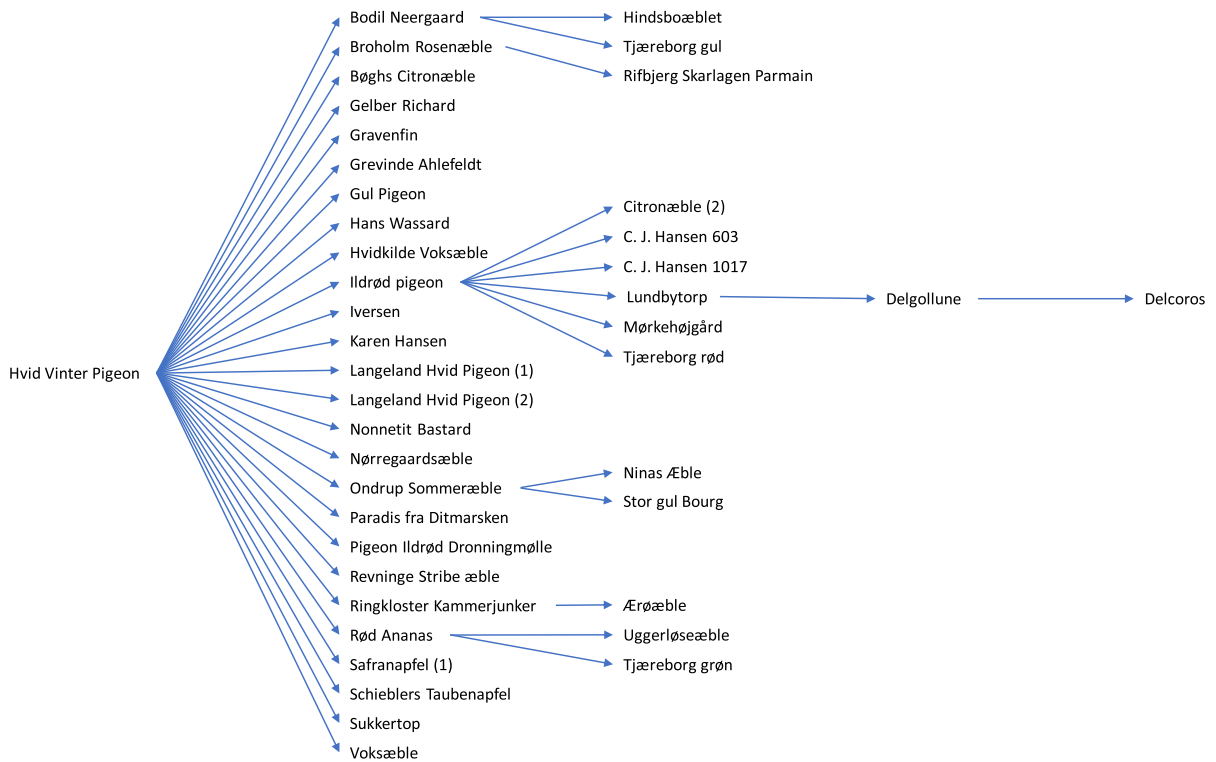


Fig. 2 Descendants of ‘Hvid Vinter Pigeon’ revealed among the 667 unique genotypic profiles. The arrows point from parent to offspring

Discussion

This study analysed for the first time the complete set of genotypic data from all apple accessions kept at The Pomatum and enabled a comparison of The Pomatum germplasm with other Danish germplasm collections, as well as various germplasm collections in Europe and other areas of the world. This international component of this study revealed 305 genotypic profiles unique to Denmark among more than 20,000 genotypic profiles in the MUNQ project or among more than 5,000 genotypic profiles in the APR project. In addition, the comparison of Danish and international genome-wide SSR and SNP data allowed for robust curation of cultivar identities, parentage validation and reconstruction, and the deduction of chromosome-wide SNP haplotypes for both the Danish and international germplasm. The data may strengthen future management of Danish germplasm and additionally provide a starting point for future research on marker-trait associations and

genetic components involved in for example climate adaptation.

Genotypic duplicates and accession identity

The results of this study have identified many genotypic duplicates in the Pomatum collection. This information can be used in future genebank management to select unwanted genotypic duplicates. However, not all duplication was unintended, as some genotypic duplicates were due to the maintenance of colour sports, such as ‘Ingrid Marie’ and ‘Rød Ingrid Marie’, and ‘Guldborg’ and ‘Rød Guldborg’. In addition, the genebank maintains 14 ‘Gravensteiner’ accessions and four ‘Cox’s Orange Pippin’ accessions with distinct accession names of which some such as ‘Red Gravensteiner’ refer to well-described colour sports. These colour sports have added to the relatively high number of genotypic duplicates (26%) present in The Pomatum

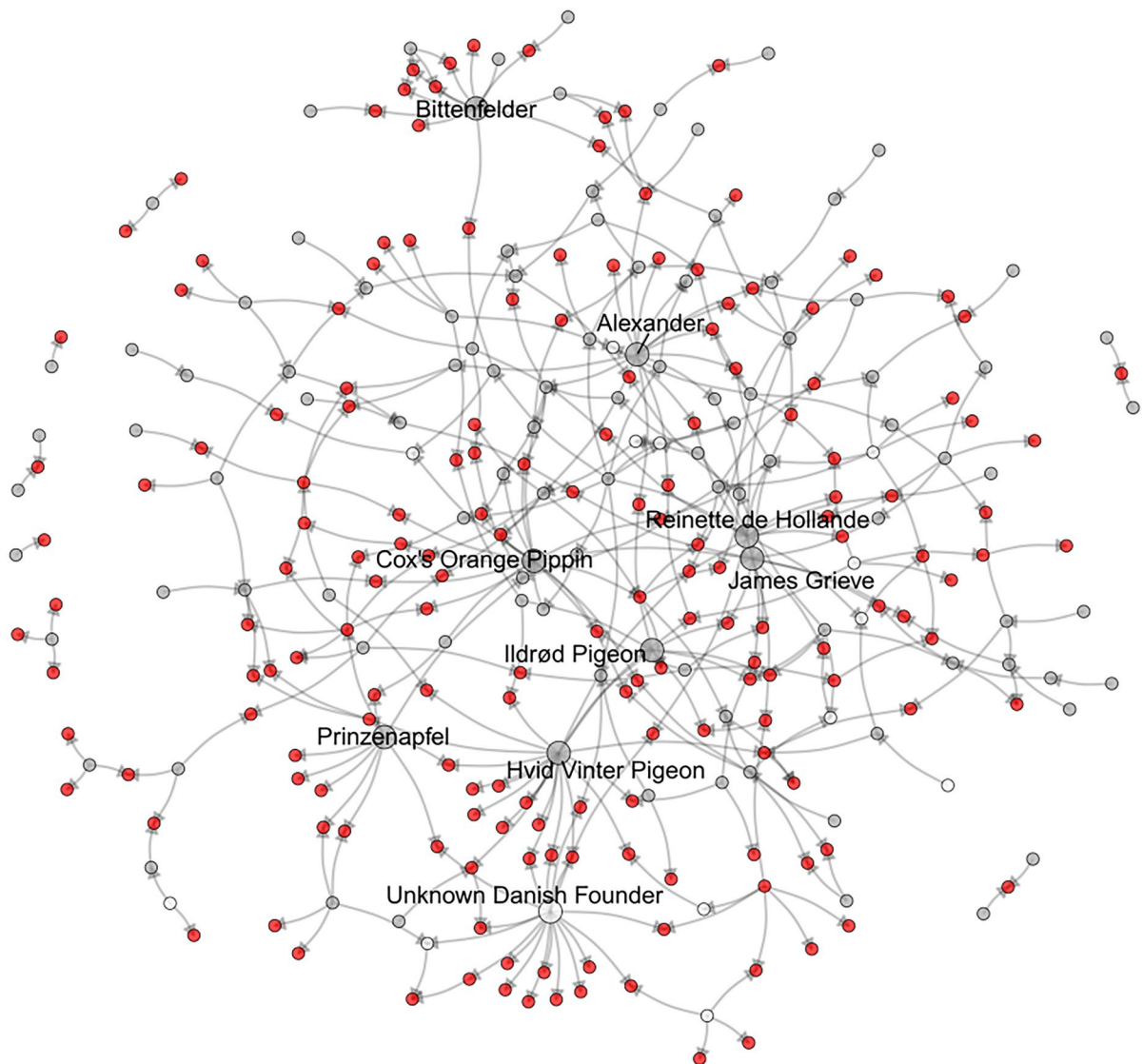


Fig. 3 Network of parent–offspring relations among cultivars considered to be of Danish origin. Arrows point from parent to offspring. Red circles represent cultivars that are considered being of Danish origin. Grey circles represent cultivars that are considered not to be of Danish origin while white circles rep-

resent ungenotyped individuals of unknown origin. Larger circles and their associated names represent cultivars with at least six offspring among the genotypic profiles that are considered to be of Danish origin. The figure includes only cultivars that have identified pedigree links

collection compared to the number of duplicates present in other European apple genebank collections, e.g. in the Swedish Central Collection (14%) (Skytte af Sättra et al. 2020) and the CGN collection in The Netherlands (10%) (Larsen et al. 2024). Even though genotypic duplication may be seen as a burden for the genebank, some duplication may still

be desired to serve as backup in order to help safeguarding particular valuable or rare germplasm.

The comparison of genotypic data across collections enabled the identification of synonyms that have not previously been reported in the literature. ‘Skovfoged’ (MUNQ 345) is a historically important cultivar in Denmark, which to our knowledge has no

previously recorded synonyms. MUNQ 345 existed under the name ‘Skovfoged’ at The Pometum, NFC, and SLU. In addition, MUNQ 345 was also found under the accession name ‘Rode Tulappel’ at Ökowerk and the CGN (Centre for Genetic Resources, the Netherlands) and under the name ‘Jepke’ at Ökowerk. Another example was ‘Skenkelsø æble’ (MUNQ 892) which was described as a possible unique Danish seedling (Matthiessen 1913) but was identified under multiple synonyms in the MUNQ project.

The integration of genotypic data from various collection sites also enabled the confirmation of previously recorded or speculated synonyms. For example, ‘Vallekilde æble’ (MUNQ 707) was reported as a synonym of ‘Beauty of Kent’ in Pedersen (1950) and its genotypic profile was identical to ‘Beauty of Kent’ from the NFC. Part of the genotypic duplicates in the Danish germplasm include regionally economically important colour mutants such as sports of ‘Cox’s Orange Pippin’, ‘Gravensteiner’, and ‘Ingrid Marie’.

Pedigree reconstruction results

The results elucidated many new pedigree relationships and networks among Danish apple germplasm. Some of these were identified or speculated based on SSR data (Larsen et al. 2017), though the results of this study confirmed many of these results and greatly expanded upon them. Important parents of Danish germplasm (Figs. 1, . 3), such as ‘Reinette de Hollande’ (syn. ‘Orleans Reinette’) and ‘Prinzenapfel’, were already introduced to Denmark in the eighteenth century and were commonly cultivated in fruit gardens in Denmark in the first half of the nineteenth century, whereas other common parents of Danish cultivars ‘Cox’s Orange Pippin’ and ‘Alexander’ were both introduced to Denmark around 1850 (Bredsted 1893). ‘Cox’s Orange Pippin’ is a parent of cultivars such as ‘Ingrid Marie’, ‘Auroavej’, and ‘Ditte Wiuff’, all of which likely arose during the twentieth century. The rootstock ‘Bittenfelder’ was a parent of 10 genotypic profiles, each represented by a single accession and all coming from two private Danish collections. None of these had names referring to any literature descriptions.

The French renaissance cultivar ‘Reinette Franche’, which has been identified as the most common parent in a diverse set of 1,400 mostly historical European

cultivars in Muranty et al. (2020), had minor importance as a parent in the Pometum heritage collection. In Denmark, it was cultivated primarily before 1850 (Bredsted 1893) and was the most common grandparent in the Danish germplasm, especially through ‘Reinette de Hollande’ (Fig. 1, Table S5). In addition, ‘Reinette Franche’ is represented in the genetics of numerous Danish cultivars through several more distant descendants that are themselves common ancestors of Danish cultivars: ‘Cox’s Orange Pippin’ (‘Reinette Franche’ is a great-grandparent; Howard et al. 2023), ‘Cox’s Pomona’ (‘Reinette Franche’ is double great-grandparent; Howard et al. 2021a), ‘James Grieve’ (‘Cox’s Orange Pippin’ is a parent). These contributions can now be quantified and localized at the level of small chromosome segments by using the inheritance of phased SNP data along genotyped pedigrees.

The Danish germplasm allowed for the identification of a half-sibling group consisting of 17 individuals sharing an unknown parent, the “Unknown Danish Founder” (Table 1). Fourteen of these half-siblings either have their reported origin in a relatively narrow part of the central to southern part of Denmark or were obtained from private collections located in the same region of the country. The “Unknown Danish Founder” might therefore potentially be a rather locally grown cultivar which might no longer exist, or which could be identified through further exploration of local material.

The Pometum germplasm allowed for reconstruction pedigrees for commercially important cultivars such as ‘Ortley’ being one parent of ‘Granny Smith’.

Future germplasm management and breeding

The study demonstrated the value of public and privately governed germplasm collections in Denmark as a resource for genetically unique apple cultivars (89) that were neither present at The Pometum heritage or non-heritage collections, nor at other investigated germplasm collections in the world (Table S1, S5). Such accessions should be considered in future conservation strategies if they are phenotypically or culturally significant, or in case they maintain desirable traits that are rare or not present in commercial germplasm, for example historical cultivars with a

long cultivation history reflecting probable regional climate adaptation.

The duplicate and pedigree reconstruction allow future germplasm management efforts to focus on accessions that are genotypically unique to the collection or ancestors of regional germplasm. Accessions for which both parents are included in the collection do not contribute additional genes (excluding mutations) and could be seen as genetically redundant as they do not contribute to the genetic diversity in the collection. However, such offspring might be critical for revealing the existence of desired traits (alleles) that are hidden in the parents but eventually might be expressed in some offspring, for example recessive genes or traits that require presence of desired alleles in more than one locus or allelic copy. Genetically redundant accessions might also be valuable for a collection in case they represent a specific cultural value, an extraordinary phenotype or genotype (resulting from a particular assemblage of genes and linkages), or constitute a specific step in a pedigree which is crucial for pedigree reconstruction or phasing. An example is ‘Ingrid Marie’ (‘Cox’s Orange Pippin’ × ‘Cox’s Pomona’), which is an important heritage cultivar that has been among the top-3 most frequently planted cultivars in Denmark for more than 40 years (DST 2018) and a parent of commercially relevant cultivars like ‘Aroma’ and ‘Elstar’. ‘Ingrid Marie’ might therefore be relevant to maintain in The Pometum heritage collection even though both of its parents are maintained simultaneously at The Pometum as well as in various other collections.

With the completion of this study, SNP array data is now available for most accessions maintained in the largest Nordic germplasm collections of heirloom cultivars in Sweden (Skytte af Sätra et al. 2020), Norway (Gilpin et al. 2023), and now Denmark. A set of 106 Finnish accessions kept at The Agricultural Research Centre Finland has been SNP genotyped, but the data has not yet been published. The germplasm for which SNP data is available has been cultivated around the northern-most apple cultivation region in the world around the 60° northern latitude in Norway (Fotirić Akšić et al. 2022; Gilpin et al. 2023) to material that has been cultivated around the major apple cultivation areas in Denmark around the 55° northern latitude, representing a span from North to South of around 600 km. The germplasm is therefore likely to be well adapted to a range of climatic

conditions at the northern border of apple cultivation making it suitable for genetic research and breeding efforts in adapted germplasm.

The available SNP array data and accession-level information may enhance future development of regionally adapted cultivars. Such breeding efforts are highly needed since most cultivars released from current breeding programmes in the world lack adaptation to the Nordic climate. In addition, several new cultivars are released as patented “club cultivars”, which are inaccessible to Nordic growers situated on the northern border of apple cultivation (Nybom et al. 2014). The development of cultivars with Nordic climate adaptation is currently hampered by limited regional apple breeding efforts even though regional apple breeding is conducted (Røen et al. 2000; Nybom 2019).

Parents of North European historical germplasm

Several apple cultivars have historically been distributed among various North European countries. Heirloom apple germplasm from different regions in Northern Europe could therefore be expected to share some common parents. Data from the APR project allowed us to compare the frequency of certain pedigree ancestors across North European germplasm collections. The two largest collections of historical Nordic apple germplasm are maintained at The Pometum and in the Swedish Central Collection, representing heirloom germplasm from the two neighbouring countries, Denmark and Sweden. However, the major parents identified in in The Pometum heritage collection did not or hardly occur as parents the Swedish Central Collection (Skytte af Sätra et al. 2020).

Four out of the seven most common parents in The Pometum heritage collection, ‘Hvid Vinter Pigeon’, ‘Cox’s Orange Pippin’, ‘Reinette de Hollande’, and ‘Prinzenapfel’ (Fig. 1, Table S7) were also among the five most common parents in the Ökowerk collection (Howard et al. 2021b). ‘Hvid Vinter Pigeon’ was the most common parent in The Pometum heritage collection. The cultivar has been cultivated in Denmark since at least 1795 and was in the nineteenth century appreciated as a dessert apple with good fruit quality under regional climate conditions (Bentzien 1861; Bredsted 1893). Twenty-seven out of the 34 identified offspring of ‘Pigeon Hvid Vinter’ were unique to

Nordic and German germplasm collections, namely The Pometum (Denmark), SLU (Sweden), Ökow-erk (Germany), and JKI (Germany), revealing ‘Hvid Vinter Pigeon’ to be a North European regionally specific ancestor. ‘Hvid Vinter Pigeon’ was identified as the parent of the two small-scale regional Danish market cultivars, ‘Ildrød Pigeon’ and ‘Rød Ananas’. The cultivar was also the great-grandparent of the French cultivar ‘Delgollune’, which was an offspring of the Danish cultivar ‘Lundbytorpæble’, and a double great-grandparent of ‘Delcoros’ (Fig. 2).

Conclusion

This study is the first Nordic study using SSR and SNP-array data and international reference databases (Denancé et al. 2020; Howard et al. 2018) for duplicate analysis and pedigree reconstruction across a complete genebank collection. This enabled us to identify 307 genotypic profiles which were unique to the Danish germplasm and to elucidate extended pedigrees by identifying multiple pedigree interferences involving genotypic profiles that were not present in the Danish dataset themselves. Similar efforts comparing apple germplasm and performing pedigree validation and reconstruction efforts combining genotypic information from different collections were performed by Muranty et al. (2020) and Luby et al. (2022). Such cross-collection comparisons should be done in a similar way for all Nordic accessions for which SSR data and SNP array data are available to further investigate genotypic duplicates and pedigree inferences within and between collection sites in the Nordic region as well as other parts of the world. The assignment of correct names and pedigrees for Danish germplasm accessions facilitates further use of the germplasm for breeding and pomological purposes.

Acknowledgements The authors would like to thank the following group and people: The French associations ‘Union Pomologique de France’ and ‘Croqueurs de pommes®’ for allowing MUNQ data of French germplasm collections to be available, Herma J.J. Koehorst-van Putten for help with data preparation, Charlotte Sommerlund Riis, Maren Korsgaard, Flemming Thorninger, Boi Jensen, and Bjarne Sandemand for help and a fruitful discussion revealing the identity of specific samples, and Anni og Jørgen Svendsen, Birgit Thinggaard, Frank Erichsen, Frode Jensen, Hanne Rasmussen, Jacob Østergård, Jes Romme, Marie Riegels Melchior, Mette Østergård, Niels Jørgen Poulsen, Peder Lollesgaard, Peter Løn,

Rikke Lock Harvig, Tina Bach Nygaard, and Villy Mougard for generously providing accessions that were included in this study.

Author Contributions Bjarne Larsen and Eric van de Weg conceived the study. Bjarne Larsen led the project, collected all samples, curated the accessions, performed the DNA-extraction, part of the SSR-genotyping, and part of the genetic analyses in collaboration with Eric van de Weg and Carsten Pedersen. Caroline Denancé and Charles-Eric Durel integrated the SSR data into the MUNQ dataset and assigned each accession to a MUNQ-code. Nicholas P. Howard curated the SNP data and integrated the data into the APR project database and assigned the pedigree interferences. Jonas Skytte af Sättra, Larisa Garkava-Gustavsson, and Michela Troggio contributed with SNP data and provided critical feedback on the methods, results, and interpretations in the study. Bjarne Larsen, Eric van de Weg, and Nicholas P. Howard wrote the majority of the manuscript, which was reviewed, edited, and approved by all authors.

Funding We are thankful for three generous grants from the Independent Research Fund Denmark (grant number 8027-00007B), The Carlsberg Foundation (grant number CF17-0576), and Foreningen PlanDanmark. These grants were given to the first author and enabled this work to be performed. SNP data that was used for comparison of the Danish germplasm SNP dataset was obtained through the collaborative apple pedigree reconstruction project. Part of this 20K SNP data were funded by the Niedersächsisches Ministerium für Wissenschaft und Kultur through the EGON project: “Research for a sustainable agricultural production: Development of organically bred fruit cultivars in creative commons initiatives” and by the FruitBreedomics Project No. 265582, (funded by the EU Seventh Framework Programme).

Data availability The standardized SSR calls and the curated SNP calls used in this study are deposited in the Genome Database for Rosaceae (rosaceae.org) under the accession number tGDR1078.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Reference

- af Sättra JS, Troglio M, Odilbekov F, Sehic J, Mattisson H, Hjalmarsson I, Ingvarsson P, Gustavsson L (2020) Genetic status of the Swedish central collection of heirloom apple cultivars. *Sci Hortic* 272:109599. <https://doi.org/10.1016/j.scienta.2020.109599>
- Bentzien JA (1861) *Haandbog for Frugttrædyrkere*. Trier, Copenhagen
- Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdak S, Di Guardo M, Salvi S, Jansen J, Viola R, Gut I, Laurens F, Chagné D, Velasco R, van de Weg E, Troglio M (2014) Development and Validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus × domestica* Borkh). *PLoS ONE* 9:e110377. <https://doi.org/10.1371/journal.pone.0110377>
- Bianco L, Cestaro A, Linsmith G, Muranty H, Denancé C, Théron A, Poncet C, Micheletti D, Kerschbamer E, Di Pierro EA, Llarger S, Pindo M, Van de Weg E, Davassi A, Laurens F, Velasco R, Durel C-E, Troglio M (2016) Development and validation of the Axiom@Apple480K SNP genotyping array. *Plant J* 86:62–74. <https://doi.org/10.1111/tpj.13145>
- Bink MCAM, Jansen J, Madduri M, Voorrips RE, Durel CE, Kouassi AB, Laurens F, Mathis F, Gessler C, Gobbin D, Rezzonico F, Patocchi A, Kellerhals M, Boudichevskaia A, Dunemann F, Peil A, Nowicka A, Lata B, Stankiewicz-Kosyl M, Jeziorek K, Pitera E, Soska A, Tomala K, Evans KM, Fernández-Fernández F, Guerra W, Korbin M, Keller S, Lewandowski M, Plochanski W, Rutkowski K, Zurawicz E, Costa F, Sansavini S, Tartarini S, Komjanc M, Mott D, Antofie A, Lateur M, Rondia A, Gianfranceschi L, van de Weg WE (2014) Bayesian QTL analyses using pedigreed families of an outcrossing species, with application to fruit firmness in apple. *Theor Appl Genet* 127:1073–1090. <https://doi.org/10.1007/s00122-014-2281-3>
- Bredsted HC (1893) *Haandbog i dansk Pomologi, 2. æbler*. Hempelske Bog- og Papirhandels Forlag, Odense
- Chagné D, Crowhurst RN, Troglio M, Davey MW, Gilmore B, Lawley C, Vanderzande S, Hellens RP, Kumar S, Cestaro A, Velasco R, Main D, Rees JD, Iezzoni A, Mockler T, Wilhelm L, van de Weg E, Gardiner SE, Bassil N, Peace C (2012) Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0031745>
- Chagné D, Kirk C, Whitworth C, Erasmus S, Bicknell R, Sargent DJ, Kumar S, Troglio M (2015) Polyploid and aneuploid detection in apple using a single nucleotide polymorphism array. *Tree Genet Genomes* 11:1–6
- Denancé C, Muranty H, Durel C-E (2020) MUNQ - *Malus* UNiQue genotype code for grouping apple accessions corresponding to a unique genotypic profile. V1 edn. Portal Data INRAE
- DST (2018) *Frugttræplantager 2017*. NYT Fra Danmarks Statistik 26:2018
- Dullum N (1961) Antonius-æblet. Særtryk af Tidsskrift for planteavl 65(3), Copenhagen
- Fotirić Akšić M, Cerović R, Hjeltnes SH, Meland M (2022) The effective pollination period of European plum (*Prunus domestica* L.) cultivars in Western Norway. *Horticulturae* 8:55
- Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol Biol Report* 13:207–209
- Gilpin L, Røen D, Schubert M, Davik J, Rumpunen K, Gardli KA, Hjeltnes SH, Alsheikh M (2023) Genetic characterization of the Norwegian apple collection. *Horticulturae* 9:575
- Hirschfeld CCL (1788) *Handbuch der Fruchtbauzucht*. Braunschweig
- Howard NP, Peace C, Silverstein KAT, Poets A, Luby JJ, Vanderzande S, Durel C-E, Muranty H, Denancé C, van de Weg E (2021a) The use of shared haplotype length information for pedigree reconstruction in asexually propagated outbreeding crops, demonstrated for apple and sweet cherry. *Hortic Res* 8:202. <https://doi.org/10.1038/s41438-021-00637-5>
- Howard NP, Troglio M, Durel C-E, Muranty H, Denancé C, Bianco L, Tillman J, van de Weg E (2021c) Integration of infinium and axion SNP array data in the outcrossing species *Malus × domestica* and causes for seemingly incompatible calls. *BMC Genom* 22:246. <https://doi.org/10.1186/s12864-021-07565-7>
- Howard NP, van de Weg E, Luby JJ (2022) A new method to reconstruct the direction of parent-offspring duo relationships using SNP array data and its demonstration on ancient and modern cultivars in the outcrossing species *malus × domestica*. *Hortic Res*. <https://doi.org/10.1093/hr/uhab069>
- Howard NP, Micheletti D, Luby JJ, Durel CE, Denancé C, Muranty H, Ordidge M, Albach DC (2023) Pedigree reconstruction for triploid apple cultivars using single nucleotide polymorphism array data. *Plants, People, Planet* 5:98–111
- Howard NP, Albach DC, Luby JJ (2018) The identification of apple pedigree information on a large diverse set of apple germplasm and its application in apple breeding using new genetic tools. *Foerdergemeinschaft Oekologischer Obstbau e. V. (FOEKO), Weinsberg*, pp 88–91
- Howard NP, Luby JJ, van de Weg E, Durel CE, Denancé C, Muranty H, Larsen B, Troglio M, Ristel M, Albach DC (2021b) Applications of SNP-based apple pedigree identification to regionally specific germplasm collections and breeding programs. 1307 edn. International Society for Horticultural Science (ISHS), Leuven, Belgium, pp 231–238
- Konjić A, Kurtović M, Grahić J, Pojskić N, Kalajdžić A, Gaši F (2023) Using high-density SNP array to investigate genetic relationships and structure of apple germplasm in Bosnia and Herzegovina. *Horticulturae* 9:527
- Lange J (1975) *Haveplanternes indførelseshistorie i Danmark I*. Fra Kvangaard Til Humlekule 5:17–28
- Larsen B, Toldam-Andersen TB, Pedersen C, Ørgaard M (2017) Unravelling genetic diversity and cultivar parentage in the Danish apple gene bank collection. *Tree Genet Genomes* 13:14. <https://doi.org/10.1007/s11295-016-1087-7>
- Larsen B, Gardner K, Pedersen C, Ørgaard M, Migicovsky Z, Myles S, Toldam-Andersen TB (2018) Population structure, relatedness and ploidy levels in an apple gene bank

- revealed through genotyping-by-sequencing. *PLoS ONE* 13:e0201889
- Larsen B, van Dooijeweert W, Durel C-E, Denancé C, Rutten M, Howard NP (2024) SNP genotyping Dutch heritage apple cultivars allows for germplasm characterization, curation, and pedigree reconstruction using genotypic data from multiple collection sites across the world. *Tree Genet Genomes* 20:21. <https://doi.org/10.1007/s11295-024-01655-9>
- Luby JJ, Howard NP, Tillman JR, Bedford DS (2022) Extended pedigrees of apple cultivars from the University of Minnesota breeding program elucidated using SNP array markers. *HortScience* 57:472–477
- Matthiessen C (1913) *Dansk Frugt*. H. Hagerup, Copenhagen
- Matthiessen C (1924) *Dansk Frugt*. H. Hagerup, Copenhagen
- Muranty H, Denancé C, Feugey L, Crépin J-L, Barbier Y, Tartarini S, Ordidge M, Troggio M, Lateur M, Nybom H, Paprstein F, Laurens F, Durel C-E (2020) Using whole-genome SNP data to reconstruct a large multi-generation pedigree in apple germplasm. *BMC Plant Biol* 20:2. <https://doi.org/10.1186/s12870-019-2171-6>
- Nybom H (2019) Apple production and breeding in Sweden. *Chron Horticult* 59:21–25
- Nybom H, RÄ, en D, Karhu S, Garkava-Gustavsson L, Tahir I, Haikonen T, RÄ, en K, Ahmadi-Afzadi M, Ghasemkhani M, Sehic J (2014) Pre-breeding for future challenges in Nordic apples: Susceptibility to fruit tree canker and storage diseases. XXIX International Horticultural Congress on Horticulture: Sustaining Lives, Livelihoods and Landscapes (IHC2014): 1127, pp 117–124
- Pedersen A (1950) *Danmarks Frugtsorter*, 1. del. æbler. Alm. Dansk Gartnerforening, Copenhagen
- Pedersen A (1966) *Frugtsorternes oprindelse*. DS-Tryk, Fredericia
- Pedersen T (2024) ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. R package version 2.2.1.9000.
- R Core Team (2022) The r project for statistical computing.
- Røen D, Moe S, Nornes L (2000) Early ripening apple cultivars from Norway. *Eucarpia Symp Fruit Breed Genetics* 538:685–688
- Smith MWG, Great Britain. Ministry of Agriculture F, Food (1971) National Apple Register of the United Kingdom. Ministry of Agriculture, Fisheries and Food
- Urrestarazu J, Denance C, Ravon E, Guyader A, Guisnel R, Feugey L, Poncet C, Lateur M, Houben P, Ordidge M, Fernandez-Fernandez F, Evans KM, Paprstein F, Sedlak J, Nybom H, Garkava-Gustavsson L, Miranda C, Gassmann J, Kellerhals M, Suprun I, Pikunova AV, Krasova NG, Torutaeva E, Dondini L, Tartarini S, Laurens F, Durel CE (2016) Analysis of the genetic diversity and structure across a wide range of germplasm reveals prominent gene flow in apple at the European level. *BMC Plant Biol* 16:130. <https://doi.org/10.1186/s12870-016-0818-0>
- van de Weg E, Di Guardo M, Jänsch M, Socquet-Juglard D, Costa F, Baumgartner I, Broggin GAL, Kellerhals M, Troggio M, Laurens F, Durel C-E, Patocchi A (2017) Epistatic fire blight resistance QTL alleles in the apple cultivar Enterprise and selection X-6398 discovered and characterized through pedigree-informed analysis. *Mol Breed* 38:5. <https://doi.org/10.1007/s11032-017-0755-0>
- Vanderzande S, Micheletti D, Troggio M, Davey MW, Keulemans J (2017) Genetic diversity, population structure, and linkage disequilibrium of elite and local apple accessions from Belgium using the IRSC array. *Tree Genet Genomes* 13:125. <https://doi.org/10.1007/s11295-017-1206-0>
- Vanderzande S, Howard NP, Cai L, Da Silva LC, Antanaviciute L, Bink MCAM, Kruisselbrink JW, Bassil N, Gasic K, Iezzoni A, Van de Weg E, Peace C (2019) High-quality, genome-wide SNP genotypic data for pedigreed germplasm of the diploid outbreeding species apple, peach, and sweet cherry through a common workflow. *PLoS ONE* 14:e0210928. <https://doi.org/10.1371/journal.pone.0210928>
- Veldhuyzen van Zanten N. (1957) *Rassenlijst voor fruitgewassen*. Instituut voor de Veredeling van Tuinbouwgewassen, The Netherlands
- Volk GM, Peace CP, Henk AD, Howard NP (2022) DNA profiling with the 20K apple SNP array reveals *Malus domestica* hybridization and admixture in *M sieversii*, *M orientalis*, and *M sylvestris* genebank accessions. *Front Plant Sci* 13:1015658

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.