Genome-wide Association Study of Individual Sugar Content and Sugar Conversion in Fruit of Japanese Pear (Pyrus spp.)

Sogo Nishio (nishios@affrc.go.jp)
Institute of Fruit Tree and Tea Science, NARO

Takeshi Hayashi
Research Center for Agricultural Information Technology, NARO

Kenta Shirasawa
Kazusa DNA Research Institute

Toshihiro Saito
Institute of Fruit Tree and Tea Science, NARO

Shingo Terakami
Institute of Fruit Tree and Tea Science, NARO

Norio Takada
Institute of Fruit Tree and Tea Science, NARO

Yukie Takeuchi
Institute of Fruit Tree and Tea Science, NARO

Shigeki Moriya
Apple Research Station, Institute of Fruit Tree and Tea Science, NARO

Akihiko Itai
Graduate School of Life and Environmental Sciences, Kyoto Prefectural University

Research Article

Keywords: Pyrus pyrifolia Nakai, Fruit quality, Sucrose, Fructose, Glucose, Sorbitol, Apple

Posted Date: February 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-170038/v1

Version of Record: A version of this preprint was published at BMC Plant Biology on August 16th, 2021.
See the published version at https://doi.org/10.1186/s12870-021-03130-2.

Abstract

Background: Sweetness is one of the most important traits determining fruit quality. Sweetness is controlled not only by the total sugar content but also by the contents of individual sugars. The major sugars in mature Rosaceae fruits are sucrose, fructose, glucose, and sorbitol, which have different levels of sweetness. Among these, sucrose and fructose have high sweetness, whereas glucose and sorbitol have low sweetness. The objective of this study was to identify the quantitative trait loci (QTLs) associated with fruit traits including individual sugar accumulation and conversion, to infer the candidate genes underlying the QTLs, and to assess the potential of genomic selection for breeding pear fruit traits.

Results: We evaluated 10 fruit traits and conducted genome-wide association studies (GWAS) for 106 cultivars and 17 breeding populations (1112 F1 individuals) using 3484 tag single-nucleotide polymorphisms (SNPs) genotyped by double-digest restriction-site associated DNA sequencing (ddRAD-Seq). By implementing a mixed linear model and a Bayesian multiple-QTL model in GWAS, 56 SNPs associated with fruit traits were identified. Four loci were presumed to be associated with sugar conversion because the SNPs were significant for more than one individual sugar and the individual sugar contents associated with each SNP genotype were negatively correlated. In particular, a SNP located close to acid invertase gene *PPAIV3* on chromosome 7 and a newly identified SNP on chromosome 11 had quite large effects on sugar conversion. We used 'Golden Delicious' doubled haploid (GDDH) 13, an apple reference genome to infer the candidate genes for the identified SNPs. In the region flanking the SNP on chromosome 11, there is a tandem repeat of *early responsive to dehydration* (*ERD6*)-like sugar transporter genes which might play a role in the phenotypes observed.

Conclusions: SNPs associated with sugar accumulation and conversion were newly identified at several loci, and candidate genes underlying QTLs were inferred using advanced apple genome information. Several QTLs showed clear effects with more than 10% of phenotypic variance explained by those SNPs in the breeding populations. By combining the effects of multiple QTLs, breeders would be able to select seedlings that will later bear fruit with high sucrose and fructose content.

Background

Pears (*Pyrus* spp.) belong to the subtribe Pyrinae of the Rosaceae and are one of the most important fruit crops in temperate regions. The origin of pear is presumed to be in the mountainous regions of southwestern China and to date back to the Tertiary period (65 to 55 million years ago) [1]. Asian pears have fruit with crisp, juicy, and sandy texture that is edible just after harvesting. Records of pear cultivation in China have been found from 2000 to 3300 years ago [1, 2], and major cultivated pears are traditionally classified into three species: *P. ussuriensis* Maxim., *P. bretschneideri* Rehder, and *P. pyrifolia* (Burm. f.) Nakai [3, 4]. Among the three species, *P. pyrifolia* is presumed to have been introduced into Japan prehistorically and became the major species in Japan [5]. Previous reports suggested that there were opportunities for ancient cultivar exchange between Japan and eastern China [6, 7], but the varieties currently cultivated in Japan and China are genetically different from each other, suggesting that they

have different breeding histories. In Japan, local cultivar ‘Nijisseiki’ has been one of the leading cultivars, and ‘Nijisseiki’ and its relatives have been repeatedly used as parents in breeding programs, suggesting that recent cultivars are deficient in genetic diversity [7].

Some fruit traits that are important for pear breeding programs and have been genetically studied are fruit harvesting day, fruit weight, fruit hardness, acid content, and sweetness [8–11]. Among these, sweetness is the most important factor determining fruit quality [9]. Fruit sweetness is controlled not only by the total sugar content but also by individual sugar composition. While sucrose is the major individual sugar in carbohydrate translocation from source to sink in most crops, the Rosaceae are unique in that sorbitol plays an important role in this process [12]. After the sugar loaded to fruit is converted by several enzymes that play a critical role in sugar metabolism during fruit development [12–17], sucrose, fructose, glucose, and sorbitol accumulate in mature Rosaceae fruit. These sugars have different levels of sweetness: if sucrose is rated 1, then fructose is 1.50–1.75, glucose is 0.70–0.80, and sorbitol is 0.55–0.70 [18–20]. In a study of various Rosaceae species, pear had a large variation in individual sugar contents in mature fruit [21], whereas cultivar collections of apple and peach, had less variation in individual sugar contents than pear. Fructose is dominant in the fruit of most apple cultivars [22, 23], while sucrose is dominant in most peach cultivars [15, 24]. Quantitative trait loci (QTLs) associated with the conversion of sucrose to hexose (i.e., fructose and glucose) in mature fruit were identified on chromosomes 1 and 7 in Japanese pear [9]. Large-effect QTLs that control individual sugar contents were also identified on apple chromosome 1 [25]. Members of the acid invertase family, a key enzyme in the conversion of sucrose to hexose, were presumed to be candidates for these QTLs. Moreover, QTLs for soluble solids concentration (SSC) have been mapped on pear chromosomes 2, 4, 5, 6, and 8 [11, 26], though the effects of these QTLs fluctuated from year to year.

Whole-genome duplications are suggested to have occurred in pear and apple, as their genome sequences have extensive syntenic blocks covering much of the chromosomes ($2n = 2x = 34$) [27]. In addition to synteny of the whole genomes, they also have interesting genes and QTLs in common. The *S* haplotypes that control gametophytic self-incompatibility [28] and susceptibility genes to *Alternaria alternata* [29, 30] are located at the same positions in both genomes A member of the 1-aminocyclopropane-1-carboxylate synthase (ACS) gene family is related to several important fruit traits including fruit harvesting day, storage ability, and fruit drop [11, 31–33]. QTLs for harvesting date on chromosome 3 were commonly identified in several studies [11, 34–36]. Because of the high similarity between *Pyrus* and *Malus* genome sequences, pear genetic studies were conducted using advanced apple genome information [9, 37]. A de novo assembly of a ‘Golden Delicious’ doubled-haploid tree (GDDH13) composed of 280 assembled scaffolds and arranged into 17 pseudomolecules is now the most widely used reference genome in apple genetic studies [38]. Although draft genomes of Chinese pear and European pear are available [39, 40], a draft genome of Japanese pear has not been available yet.

Several useful DNA markers have been developed and applied to Japanese pear breeding programs: these include DNA markers to identify self-compatibility [41, 42], a molecular marker associated with

disease resistance genes [43, 44], and markers associated with fruit harvesting day [8, 9, 11]. Unlike some traits controlled by a single gene or large-effect QTLs, marker-assisted selection (MAS) for traits controlled by multiple minor genes has not been applied in pear breeding programs. Currently, genomic selection (GS) is gaining attention as an efficient breeding method for such traits in fruit trees. GS utilizes predicted breeding values given by prediction models based on genome-wide single-nucleotide polymorphism (SNP) data to enable selection of superior individuals. The potential of GS for use in Japanese pear breeding was assessed by Minamikawa et al. [10], who used 86 varieties (84 Japanese pear, 2 Chinese pear) and 765 F1 trees from 16 breeding populations (full-sib families) genotyped for SNPs to compare the accuracy of genomic prediction obtained using 12 different methods. The accuracy of the mean prediction for these models was about 0.6 for fruit quality traits, but those for physiological disorders such as heart rot and watercore were low.

The major challenges to improvement of fruit and nut trees are the long juvenile phase and large plant size [45–47]. The fruit traits can only be phenotypically assessed once the plants are physiologically mature [48], and the cost of growing plants to that stage remains quite high. On the other hand, the cost of genotyping by next-generation sequencing technology is dramatically decreasing. As each pear fruit generally contains about five to ten seeds, obtaining the number of seeds necessary for cross-breeding is not as difficult as it is in stone fruit species, which have only one seed per fruit. Therefore, pear is a suitable fruit tree for introducing GS in future breeding programs. In this study, we updated Minamikawa et al.'s [10] study by further evaluating sugar component, by increasing the numbers of cultivars and individuals and by applying a more powerful genotyping method. Here, we used double-digest restriction-site associated DNA sequencing (ddRAD-Seq) to genotype 106 cultivars and 17 breeding populations (1112 F1 individuals), representing about 40% more genotypes than in the previous study. To characterize the sugar metabolism in these materials and to increase the content of specific sugars, here we measured individual sugar contents as well as total sugar content. The objective of this study was to develop molecular markers associated with individual sugar contents and to identify additional QTLs for fruit traits. Candidate genes for those QTLs were inferred by using the advanced apple reference genome. We also examined the potential of GS for future use in pear breeding programs.

Results

Phenotypic distribution of individual sugars

The average contents of sucrose (SUC), fructose (FRU), glucose (GLU), sorbitol (SOR), and total sugar content (TSC) averaged over 1218 individuals (Table 1) were 43.1, 40.6, 14.9, 32.2, and 130.7 mg/ml, respectively. SUC showed the greatest range of phenotypic variation (1.6–117.0 mg/ml, variance 355.6; Table S1). The distributions of FRU, GLU, SOR, and TSC were much narrower: 9.6–70.7 mg/ml (variance 91.9), 0–35.0 mg/ml (variance 64.3), 11.5–56.3 mg/ml (variance 56.5), and 101.3–173.8 mg/ml (variance 125.2), respectively. Fruit harvest time (HarT) ranged over 93 days (July 23 to October 23, mean August 31; Table S1). The range of fruit weight (FruW) was 92.5–1016.1 g (mean 394.2 g), fruit hardness

(FruH) was 2.2–9.5 lb (mean 4.8 lb), SSC was 10.3–18.1% (mean 13.8%) and Aci was 3.4–5.6% (mean 4.9).

Phenotypic correlation coefficients and their significances for all trait combinations were calculated (Table 2). SUC had strong negative correlations with FRU, GLU, and SOR (r = −0.57, −0.76, and −0.42, respectively), and GLU had positive correlations with FRU and SOR (r = 0.39 and 0.36, respectively). TSC was correlated positively with SUC, FRU, and SOR (r = 0.36, 0.22, and 0.29, respectively) but was not significantly correlated with GLU. TSC had a very strong positive correlation with SSC (r = 0.92), indicating that the SSC of juice from mature Japanese pear was composed almost entirely of the individual sugar contents.

ddRAD genotyping

A total of 1.7 billion reads were obtained from the 1218 individuals by Illumina HiSeq 4000 (average of 1.4 M reads per individual). After trimming low-quality data and adapter sequences, 94.5% of the high-quality reads were successfully mapped onto the apple GDDH13 reference genome. After selecting SNP loci by the criteria of VCFtools described in the “Materials and Methods” section and removing individual SNP markers with <25% missing data, we obtained 9011 SNPs. Missing data were imputed using Beagle 4.0, and 7463 SNPs with >0.8 imputation accuracy were selected. To normalize the SNP density, 3484 tag SNPs were selected from the 7463 SNPs using Haploview.

Linkage disequilibrium and population structure

Linkage disequilibrium (LD) between pairs of loci was estimated using all 7463 SNPs, before selecting the tag SNPs. The r^2 values (estimates of LD) were plotted against marker distances (bp) in the apple genome GDDH13. The average r^2 values dropped below 0.2 at a marker distance of 250 kb and below 0.1 at 1750 kb (Figure S1). Population structure was estimated by Bayesian clustering analysis and principal component analysis (PCA; Fig. 1). Bayesian clustering analyses at K = 3 and K = 4 showed an overview of the genetic relationships among the breeding populations. Although cultivars showed admixed genetic structure, each breeding population had a simple genetic structure. Populations derived from crosses between ‘Kanta’ and other cultivars (population IDs 581, 589, 590, 591, and 592) carried “orange” genetic clusters, those derived from crosses between ‘Hoshiakari’ and other cultivars (IDs 538, 541, 543, 578, 581, and 588) carried “light blue” genetic clusters, and those derived from crosses between ‘Akizuki’ and other cultivars (IDs 523, 538, 539, and 546) carried “dark blue” genetic clusters. When K = 4, the “dark blue” genetic clusters in some cultivars and populations became separated into “dark blue” and “green” clusters. The new “green” cluster seemed to be derived from relatively old cultivars such as ‘Okuroku’ and ‘Inagi’ that are genetically somewhat different from recent cultivars.

PCA showed a similar classification to the Bayesian clustering analysis. The first principal component, which explained 16.3% of the total variation, reflected the difference between old cultivars or populations

derived from relatively old cultivars (green), and populations derived from crosses among new cultivars (light blue, orange and dark blue). The second principal component, which explained 12.5% of the total variation, reflected the genetic difference between 'Kanta' and 'Hoshiakari', i.e., the populations derived from 'Kanta' (orange) were distributed towards positive values of the second principal component and those of 'Hoshiakari' (light blue) were distributed towards negative values.

Genome-wide association study (GWAS)

GWAS analysis was conducted for 10 fruit traits by applying two methods: one using a mixed linear model (MLM) to test the significance of association between a single SNP and each trait, and the other using a Bayesian multiple-QTL model. The latter was regarded a variational approximation version of BayesB method [49], in which all SNPs were simultaneously fitted in the model with a variational approximation [76] (hereafter referred to as vBayesB).

In the MLM-based GWAS, 16 significant SNPs were identified for five fruit traits (SUC, FRU, GLU, HarT, and acid content [Aci]; Table 3 and Fig. S2). No significant SNPs were identified for SOR, TSC, FruW, FruH, or SSC. The relationships between genotype and phenotype for SNPs that explained ≥10% of the phenotypic variation for a trait are shown in box plots (Fig. 2). Different sugar content traits that were significantly associated with the same SNP are grouped in Fig. 2 in order to compare the relationships between marker genotype and the content of each individual sugar, possibly revealing SNPs involved in sugar conversion. Two QTLs for SUC were detected on chromosomes 7 and 15; four QTLs for FRU were detected on chromosomes 6, 7, 11, and 15; and four QTLs for GLU were detected on chromosomes 0, 6, 7, and 11. Chr07_33139082 was identified as a significant SNP for SUC, FRU, and GLU, with 23.5%, 9.2%, and 25.2% of the phenotypic variance, respectively, explained by the SNP effects. The average values of genotypes at Chr07_33139082 for AA, AC, and CC were 54.0, 39.1, and 27.4 mg/ml for SUC; 36.8, 43.0, and 42.2 mg/ml for FRU; and 10.4, 16.0, and 22.8 mg/ml for GLU, respectively (Fig. 2, Figure S3), suggesting that the effect of the QTL on SUC had strong negative correlations with its effects on both FRU and GLU. On the other hand, this SNP seemed to have little effect on TSC (131.5, 130.5, and 129.5 mg /ml for the average values of AA, AC, and CC, respectively; Figure S3). For FRU and GLU, two SNPs (Chr06_7938399 and Chr11_41197041) were significant for both traits. Chr11_41197041 showed effects with negative correlation between FRU and GLU. The effect of Chr11_41197041 on GLU had the highest −log10 (p) value, with 21.7% of the variance explained by the SNP. The average value for GLU of each genotype at Chr11_41197041 was 10.8 mg/ml for GG, 14.3 mg/ml for GT, and 21.4 mg/ml for TT, whereas the average values for FRU were 42.8 mg/ml for GG, 40.5 mg/ml for GT, and 38.1 mg/ml for TT (Fig. 2). On the other hand, this SNP seemed to have a relatively small effect on TSC (129.4, 130.9, and 132.1 mg /ml for the average value of GG, GT, and TT, respectively; Figure S3). Chr00_30710088, located on fictive chromosome 0 in the apple GDDH13 genome, was also significant for GLU, but was strongly linked to Chr11_41197041 in several populations (for example, r^2 =1.00 between Chr00_30710088 and Chr11_41197041 for population 589, derived from 'Kanta' and 'Rinka'). Thus, the effect of Chr00_30710088 was excluded from further analysis and discussion. Also, Chr15_17923340, which was

significant for SUC, and Chr15_16568005, which was significant for FRU, were located at close positions, suggesting that these two SNPs are detecting the same QTL. For HarT, two significant SNPs were identified. Chr03_31587739 and Chr15_16568005 showed high −log10 (p) values (8.9 and 12.2, respectively) and explained 16.5% and 20.3% of the phenotypic variance, respectively. The difference between the average values of the CT and TT genotypes was 14 days for Chr03_31587739, while the number of CC individuals was too low to include in the comparison. For Chr15_16568005, the average values of the AG and GG genotypes were comparable (52.7 and 52.0, respectively), but that of AA was 66.2. Out of the four significant SNPs associated with Aci, Chr06_20720912 and Chr16_3043313 showed relatively high percentages of variance explained (13.1% and 11.0%, respectively). The average values for Aci at Chr06_20720912 were 4.85 for AA, 4.96 for AT, and 5.20 for TT (Fig. 2), and those at Chr16_3043313 were 4.85 for CC, 4.90 for CG, and 5.08 for GG (Fig. 2).

In the vBayesB-based GWAS, 40 SNPs showed posterior probabilities exceeding 0.85 for all 10 fruit traits (Table 4, Fig. S2). Out of those 40 SNPs, 4 were also significant in the MLM-based GWAS: Chr07_33139082 for SUC and GLU, Chr11_41197041 for FRU, and Chr06_20720912 for Aci. Chr11_41197063 (for GLU) is close to Chr11_41197041, suggesting that they are associated with the same gene. Although the posterior probability values were high, the percentages of phenotypic variance explained by the SNPs were not large for most of the detected SNPs. Out of the 40 SNPs, only 7 had percentages of phenotypic variance explained of ≥10%, including the four that were also significant in MLM-based GWAS (considering Chr11_41197063 to be the same locus as Chr11_41197041). Chr04_29871378 explained 12.7% of the phenotypic variance for FRU, and the average value of each genotype was 45.3 mg/ml for CC, 39.6 mg/ml for CG, and 35.6 mg/ml for GG. These values were positively associated with those for TSC (134.2 mg/ml for CC, 129.8 mg/ml for CG, and 127.3 mg/ml for GG; Figure S3), though this SNP was not detected in the vBayesB-based GWAS for TSC. Chr06_7938399 (for HarT) and Chr12_4383743 (for FruH) each accounted for 10% or more of the phenotypic variance for the indicated trait, but these SNP have minor genotype frequency (<0.05) and did not segregate in most of the populations.

To illustrate the effects of the 56 SNPs from the two GWAS analyses in each population, the percentages of phenotypic variance explained by SNPs in each population were calculated and displayed in a heatmap (Figure S4). While the SNPs identified in MLM-based GWAS tended to show effects across the 1218 individuals, some SNPs identified in vBayesB-based GWAS showed effects specific to a population. For example, Chr04_5362900 (for HarT) and Chr12_4383743 (for FruH) explained high percentages of variance only in the cultivar collection.

Genomic selection

We attempted to predict the breeding values of individuals with genomic best linear unbiased prediction (GBLUP), in which a polygenic effect included in MLM-based GWAS was regarded as the breeding value of an individual, and with the vBayesB-based method, which was also used in GWAS. We used each F1

family as a tested population and the remaining families and cultivars as the training population, where a prediction model was constructed with both observed phenotypes and SNP genotypes of a training population and the breeding values of a tested population were predicted from SNP genotypes with the constructed model. We calculated correlation coefficients between predicted breeding values and observed phenotypes over all individuals in F1 families (Table 5). The prediction accuracies in GBLUP were higher than those in the vBayesB-based method for SUC, FRU, SOR, TSC, HarT, FruW, FruH, and SSC, whereas the prediction accuracies in GBLUP for GLU and Aci were slightly lower than those in the vBayesB-based method. GLU showed the highest values of prediction accuracy in both GBLUP and the vBayesB-based method (0.74 and 0.75, respectively). TSC, FruW, FruH, and SSC showed lower values of prediction accuracy, especially with the vBayesB-based method (0.39–0.46).

Discussion

Several phenotypic correlations among individual sugars were identified in 1218 individuals and cultivars. SUC had strong negative correlations with both FRU and GLU (r = −0.57 and −0.76, respectively), and FRU had a positive correlation with GLU (r = 0.39). This result was quite similar to those in previous studies of QTLs for individual sugar traits [9] and genetic differences in individual sugars among leading or promising cultivars [50]. In addition, SOR had a negative correlation with SUC in this study (r = −0.42). These results suggested that sugar conversion had a strong influence on sugar composition in mature fruit. TSC had quite a strong correlation with SSC, suggesting that it would be possible to use SSC for the first screening of individuals: it is much easier to measure SSC (the amount of sugar in fruit) than TSC, which requires high-performance liquid chromatography. In a previous study, the broad-sense heritability of TSC was only 0.33, whereas those of SUC, FRU, GLU, and SOR were 0.64, 0.69, 0.71, and 0.76, respectively [50]. Segregating TSC into individual sugars would be an effective way to identify the genes associated with sugar accumulation and conversion in mature fruit, because the low heritability of TSC would make genetic analyses of that trait difficult. HarT had positive correlations with FruW, TSC, and SSC, as reported in previous studies [10, 51]. This suggests that it might be difficult to develop early-ripening cultivars with large fruit size and high sugar content. On the other hand, FruH and Aci did not show clear correlations with other fruit traits, suggesting that they are independent and controlled by different genes.

In previous studies, LD values of Japanese pear had been calculated based on genetic linkage maps (centimorgan distance) [8, 10]. In this study, we calculated LD based on the apple physical map, enabling us to compare the LD blocks with those in apple. The average r^2 values dropped below 0.2 at 250 kb. Our material had a smaller LD block size than the apples studied by Moriya et al. [52] and Kumar et al. [53], but a larger LD block size than in the study of Leforestier et al. [54]. We expected that the LD in Japanese pear would be high because genetic bottlenecks and breeding history would have increased the extent of LD by eliminating recombinant lineages [8]. Nevertheless, the LD block size was smaller than in several apple genetic studies, suggesting that apple experienced a strong bottleneck similar to that presumed to have occurred during pear breeding and domestication. In the present study, the cultivar collection had

admixed structure in Bayesian structure analyses and showed broad distribution in PCA compared to individual populations, whereas each population had smaller genetic diversity than the cultivar collection. In the Bayesian structure analysis at $K = 3$, the populations were roughly divided into three groups: those derived from crosses between 'Kanta' and other cultivars ("orange"), those derived from crosses between 'Hoshiakari' and other cultivars ("light blue"), and those derived from crosses between 'Akizuki' and other cultivars ("dark blue"). 'Kanta' has high sweetness, with high TSC and FRU [50], 'Hoshiakari' has a scab resistance gene inherited from local cultivar 'Kinchaku' [55], and Akizuki has a good fruit texture with excellent fruit shape. Since each of these cultivars has specific desirable characteristics, they have been used as parents in Japanese pear breeding programs. In essence, the populations used in this study were based on only 13 founder cultivars [55]. Because of the loss of genetic diversity among modern cultivars [33], the genetic diversity of the populations used in this study was not very large. The genetic structure identified in this study reflects the small genetic differences between the parental cultivars used to create the populations, which basically represent the same gene pool, rather than ancestral populations derived from different gene pools.

In this study, we identified a QTL associated with conversion of SUC to FRU and GLU on chromosome 7. A previous study also identified a QTL associated with conversion of SUC to FRU and GLU in this region (chromosome 7) and on chromosome 1, and further suggested that acid invertase genes *PPAIV1* and *PPAIV3* were the candidate genes [9]. In apple, a similar QTL for conversion of individual sugars was also identified on chromosome 1 in both mature fruit and fruit after storage [25]. Thus, it is possible that the AIV gene family plays an important role in determining individual sugar contents in mature fruit of Rosaceae fruit species. Other candidate QTLs that were significant for SUC and FRU on chromosome 15 (Chr15_17923340 and Chr15_16568005) might be sucrose synthase 4 (MD15G1223500, Chr15:18156590–18162358) and sucrose phosphate synthase 1F (MD15G1164900, Chr15: 12407116–12413753). Because sucrose synthase catalyzes the reversible conversion of sucrose and uridine diphosphate (UDP) to UDP-glucose, and sucrose phosphate synthase catalyzes the synthesis of sucrose from sucrose 6-phosphate, it is reasonable that these genes are related to the QTL on chromosome 15. Chr11_41197041 was identified as a significant SNP for both FRU and GLU, showing a strong effect in several populations (Figure S4). Because a negative correlation between FRU and GLU was observed for this SNP, a gene associated with conversion from FRU to GLU and vice versa would likely be involved. This SNP would potentially enable us to select genotypes with high FRU and low GLU, thus improving fruit sweetness. The QTL identified at Chr04_29871378 for FRU explained 12.7% of the variance among 1218 individuals and showed relatively high percentages of variance in several populations (more than 10% of the variance explained in nine populations; Figure S4). The effect of the QTL was limited to FRU; other individual sugars seemed to be unaffected, suggesting that an increase in FRU is directly associated with an increase in TSC (Figure S3). Interestingly, an *early responsive to dehydration* (*ERD6*)-like gene, encoding a sugar transporter, was also located near those SNPs (Chr11_41197041 and Chr04_29871378). Whereas a single *ERD6*-like gene is located near Chr04_29871378 (MD04G1211000; Chr4:29523287–29525923), a tandem repeat of *ERD6*-like genes (MD11G1293100, MD11G1293200, MD11G1293300, MD11G1293400; Chr11:41256007–41275197) is located near Chr11_41197041.

Arabidopsis ERD6 is related to sugar transporters in bacteria, yeasts, other plants, and mammals [56], and expression of its gene is induced by cold treatment. Tandem duplications of the *ERD6*-like family were found to be conserved in citrus, grape, apple, and Chinese jujube [57]. Zhang et al. suggested that duplication of sugar transporter genes plays crucial roles in sugar accumulation [58].

Several QTLs for HarT and Aci similar to those identified here were identified in previous studies of apple. The NAC18.1 transcription factor (MD03G1222600; Chr3:30696191–30698216) is a promising functional candidate for fruit ripening [59] and is close to Chr03_31587739 (for HarT) in this study. A NAC-family genes were also associated with maturity date and slow ripening in peach [60, 61]. One member of the ACS gene family (MD15G1203500) is located at Chr15:16180415, which is close to Chr15_16568005 (for HarT) in this study. This QTL and the function of the ACS gene have already been analyzed in several studies and found to affect fruit harvesting day, storage ability, and fruit drop [11, 31–33]. The early-harvesting genotype is correlated with increased fruit drop and short storage ability, suggesting that pear breeders need to consider carefully whether they use this marker for MAS, depending on their breeding objectives [33]. In apple, an aluminum-activated malate transporter-like gene (*ma1*, MDP0000252114) was identified as the candidate for a gene associated with low fruit acidity [62]. In the apple GDDH13 genome, *ma1* is located at Chr16:3176495–3179279, which is close to Chr16_3043313 (for Aci) in this study. In apple, two major loci, *Ma3* on chromosome 8 and *Ma1* on chromosome 16, play an important role in fruit acidity [63]. On the other hand, another major locus for Aci in this study was identified on chromosome 6: this locus showed quite a large effect and explained more than 30% of the phenotypic variance in seven populations (Chr06_20720912 in Figure S4).

In this study, we focused on detecting QTLs and inferring putative candidate genes for those QTLs, whereas Minamikawa et al. [10] focused on the accuracy of various methods for genomic prediction rather than on inferring candidate genes. We estimated prediction accuracy for GS using GBLUP and a vBayesB-based method for GWAS analyses. Most traits showed higher prediction accuracies with GBLUP than with the vBayesB-based method; on the other hand, the prediction accuracies for GLU and Aci were slightly higher when the vBayesB-based method was used. It will be important to choose the appropriate model depending on the objective and the traits. As the prediction accuracy for TSC, which is the most important trait to determine fruit quality, is not high enough for use in GS (0.56 in GBLUP), it is reasonable to first test GS in a limited number of crosses rather than to apply it to all of the seedlings in breeding programs. The accuracy of genomic prediction can be improved further when full-sib data for the target family are available [10]. It would be a good idea to apply GS to the populations used in this study, i.e., those derived from the crosses between ‘Hoshiakari’, ‘Kanta’, ‘Akizuki’, and other cultivars. In future, it is possible that further phenotypic data accumulation and further genotyping using whole-genome sequencing will accelerate the introduction of GS into practical breeding programs for Japanese pear.

Conclusions

In this study, we collected phenotypic data for fruit traits and conducted GWAS and GS in Japanese pear. Several important QTLs for fruit traits were identified, and genes associated with sugar accumulation and

sugar conversion were predicted by GWAS using a large number of individuals. The SNP located closest to *PPAIV3* on chromosome 7 and a newly identified SNP at chromosome 11 had large effects on sugar conversion. The SNP on chromosome 4 that was associated with FRU would be useful for increasing the contents of FRU and TSC without decreasing the contents of other individual sugars. By combining those QTLs, breeders would be able to select seedlings that will have high sucrose and fructose content in the fruit. We used the apple GDDH13 reference genome to locate the SNPs, thus enabling us to compare the effects of QTLs to those from previous studies, to infer the candidate genes, and to understand the accuracy of GS values for individual sugars, which were relatively higher than that of TSC. The traits evaluated in this study covered the principal fruit traits in pear breeding programs, so the results obtained illustrate the feasibility of GS for fruit traits in pear.

Methods

Plant materials

A cultivar collection including 106 cultivars and 17 breeding populations (consisting of 1112 F1 individuals) were used in this study (Table 1, Table S1). The population ID numbers are indicated in Table 1, and each family contained 28 to 121 individuals. Among those materials, 74 cultivars and 498 individuals were the same as those used in Minamikawa et al. [10]. The breeding populations originated from local cultivar ‘Nijisseiki’ and are the product of about five to seven generations of crossing in the NIFTS pear breeding program. The cultivars and F1 individuals were grown with cultural techniques used in commercial production in Japan [64]. The trees were trained on horizontal trellises, pruned annually in winter, and treated for pests and diseases. Fruits were thinned to one fruit per three fruit clusters in mid-May and harvested during late July to early November according to a color chart that indicates the optimum color for picking Japanese pear [65].

Ten fruit traits were evaluated: SUC, FRU, GLU, SOR, TSC, HarT, FruW, FruH, SSC, and Aci. The contents of each sugar (SUC, FRU, GLU, SOR) were analyzed in fruit from each individual. To do this, the juice from two fruits per sampling date was extracted and the samples were combined. Sampling was performed on two days in each of the years that the individual or cultivar was analyzed. The analysis of sugar components was based on the method described by Nishio et al. [9]. TSC was calculated by summing the contents of the four sugars. The harvest date for each fruit was expressed as the number of days after June 30 (i.e., July 1 = day 1), and the average value of harvest date for each fruit was used as its phenotypic score. Fruit weight (g) per fruit (FruW) was measured on a digital scale on each harvest date. Fruit hardness (FruH) was measured by the Magness–Taylor pressure test (lb). Total soluble solids concentration was determined with a digital refractometer (DBX-55; Atago, Japan) by adding a few drops of juice onto the lens of the measuring device, and the results were recorded as SSC (%). Juice pH was determined with a pH meter (IQ240; Scientific Instruments, USA) to estimate acid content (Aci).

Phenotypic data were collected from 2014 to 2018. The majority of individuals and cultivars were evaluated for more than one year. The average phenotypic values over five years (2014–2018) were

generated by the allEffects function from the package “effects”. Phenotypic correlation coefficients and their significance were calculated for all trait combinations using R (4.0.0) (R Development Core Team).

SNP genotyping

Genomic DNA was extracted from young leaves with a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. ddRAD-Seq libraries were constructed as described in Shirasawa et al. [66]. A total of 200 ng of genomic DNA for each individual was double digested with PstI and MspI (FastDigest restriction enzymes; Thermo Fisher Scientific, Waltham, MA, USA), ligated to adapters using the LigaFast Rapid DNA Ligation System (Promega, Madison, WI, USA), and purified using Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA) to eliminate short (<300 bp) DNA fragments. Purified DNA was diluted with H_2O and amplified by 20 cycles of PCR with indexed primers. Amplicons were pooled and separated on a BluePippin 1.5% agarose cassette (Sage Science, Beverly, MA, USA), and fragments of 300–900 bp were purified using a Mini Elute Kit (Qiagen). The library was then sequenced using a HiSeq 4000 (Illumina, Inc., San Diego, CA, USA).

SNPs and indels were identified according to Acquadro et al. [67]. Illumina reads were de-multiplexed on the basis of the Illumina TruSeq index. Raw reads were analyzed with Scythe (https://github.com/vsbuffalo/scythe) for filtering out contaminant substrings and Sickle (https://github.com/najoshi/sickle), which removes reads with poor-quality ends (Q<30). Alignment to the reference apple genome (GDDH13 v1.1; https://www.rosaceae.org/species/malus/malus_x_domestica/genome_GDDH13_v1.1) was carried out using BWA aligner [68] (e.g., mem command) with default parameters and avoiding multiple-mapping reads. SNP and indel mining was conducted by adopting a Samtools-based pipeline [69]. SNPs and indels (hereafter called SNPs) detected from the alignments were filtered with VCFtools (version 0.1.13; parameters: –minQ 20 –minDP 10 –maf 0.01 –maf 0.99 –min-allele 2 –max-allele 2). Individual SNP markers with >25% missing data were removed from further analysis, and the remaining missing values were phased and imputed using Beagle v4.1 [70]. Only the SNPs with an imputation accuracy >0.8 were adopted. The tagger software implemented in Haploview (http://www.broad.mit.edu/mpg/haploview) was used to select tag SNPs using its pairwise function, with a minimum r^2 of 0.8.

LD and population structure

LD was estimated using data after the imputation. Based on the apple GDDH13 genome, the detected SNPs were aligned according to their positions. LD between pairs of SNPs was calculated using R (4.0.0) (R Development Core Team). The SNPs mapped on the fictive chromosome 0, which contains all unassigned scaffolds, were removed to calculate the LDs. Average LD values (r^2) were plotted against physical distances in increments of 10 kb.

Population structure was estimated for each cross and for the cultivar collection in ADMIXTURE 1.30 [71] using the 3484 tag SNPs. The software PLINK v1.90 [72] was used to generate an input file from a vcf file. The analysis was run 10 times for each value of K (number of inferred ancestral populations) from 2 to 8. The CLUMPAK online tool [73] was applied to calculate the ΔK values and to graphically display the results produced by ADMIXTURE at $K = 3$ and $K = 4$. PCA was performed in PLINK v1.90. Prior to PCA, an input file was created by conducting linkage pruning using the –indep-pairwise option in PLINK (plink –file data –indep-pairwise 50 10 0.1). PCA was performed using the –pca option in PLINK. The plot was drawn with the R package “ggplot” [74].

Statistical methods in GWAS

For the MLM-based GWAS, the effect of a SNP and a polygenic effect affected by genetic background of an individual were included in the model as a fixed effect and a random effect, respectively. The covariance matrix of polygenic effects between individuals was established with a kinship matrix calculated from SNP genotypes. The effects of population structure were also included as fixed effects using the first three principal components obtained from the SNP genotype data. We used the R package rrBLUP ver. 4.3 [75] for this MLM-based GWAS and evaluated the effects of the significant SNPs on the basis of the mean phenotypic value for each genotype of the SNP.

For the vBayesB-based GWAS, a Bayesian multiple-QTL model was applied in which all SNPs were simultaneously fitted in the model assuming a specific probability for each SNP included in the model, called SNP weight, and a specific variance for each SNP effect as well as the progeny effect specific to each F1 family. The progeny effect specific to each F1 family was also included in the model to control for the difference of genetic background between F1 families. The model was fitted with a variational approximation method proposed by Hayashi and Iwata [76] using a custom program written with Fortran. This Bayesian procedure was regarded as a variational approximation of BayesB [49] and was previously applied to GWAS in chestnut [77], in which adopted estimation methods were described (we refer the reader to this paper for computational details). The settings of hyperparameter values of the prior distributions were conducted following this paper.

To validate the effect of the detected SNPs, the average values of each genotype for the SNPs and the distribution of the phenotypic values in the SNP genotypes were plotted. These plots were drawn with the R packages “ggplot” [74] and “gplots” [78]. The phenotypic variance explained by each SNP was calculated according to Nishio et al. [33]. The percentage of the phenotypic variance explained by each SNP was calculated by dividing the variance of the SNP by the total phenotypic variance.

Accuracies of predicted genomic breeding values in GS

The two methods used in GWAS, MLM-based and vBayesB-based, were also used for prediction of breeding values based on SNP genotypes. In the MLM-based method, a polygenic effect was regarded as

a breeding value that was predicted from SNP genotypes through a kinship matrix. This MLM-based method was referred to as GBLUP when applied to GS. In this study, an intercept but no fixed effects were included in MLM for GS. In the vBayesB method, we obtained estimates of a SNP effect and a SNP weight for each SNP, regarded as an approximate posterior probability of each SNP included in the model; accordingly, the predicted breeding value of an individual was calculated as the sum of the estimated SNP effects multiplied by the SNP weights over all SNPs.

Generally, in practical breeding programs, new elite cultivars and selections have been used as parents to create breeding populations and to select superior genotypes superior to the established leading cultivars. Thus, GS would be applied in breeding populations obtained by crossing between new combinations of parental genotypes. To evaluate the accuracy of genomic prediction for practical pear breeding programs, we conducted a cross validation in the following way. Each F1 family was regarded as a tested family, and the set of cultivars and F1 families with the tested family excluded was regarded as a training population. A prediction model was constructed using data for both phenotypes and SNP genotypes of a training population, and breeding values of individuals of a tested F1 family were predicted with a model using only their SNP genotypes. This process was repeated until all F1 families were selected just once as a tested population and the prediction accuracy was evaluated with the Pearson's correlation coefficient (r) between observed phenotypic values and predicted genotypic values. When estimated r was less than 0, it was regarded as 0. The prediction accuracy for eachtested family based on the training population (which excluded the target family) was calculated and averaged to understand the accuracies of predicted genomic breeding values for the breeding populations used in this study.

Abbreviations

Aci: Acid content; **ACS:** 1-aminocyclopropane-1-carboxylate synthase; **BLUP:** Best linear unbiased prediction; **FRU:** Fructose content; **FruH:** Fruit hardness; **FruW:** Fruit weight; **GBLUP:** Genomic best linear unbiased prediction; **GDDH:** 'Golden Delicious' doubled haploid; **GLU:** Glucose content; **GS:** Genomic selection; **GWAS:** Genome-wide association study; **HarT:** Fruit harvest time; **MLM:** mixed linear model; **NIFTS:** Institute of Fruit Tree and Tea Science, NARO; **PCA:** principal component analysis; **SNP:** single-nucleotide polymorphism; **SOR:** Sorbitol content; **SSC:** Soluble solids concentration; **SUC:** Sucrose content; **TSC:** Total sugar content

Declarations

Ethics approval (and consent to participate)

Not applicable.

Consent for publication

Not applicable.

Availability of supporting data

The datasets supporting the conclusions of this article are included within the article and its supplementary information files. Sequence reads are available from the Sequence Read Archive (DRA) of DNA Data Bank of Japan (DDBJ) under the accession number of DRA011324 (https://ddbj.nig.ac.jp/DRASearch/submission?acc=DRA011324).

Competing interests

The authors declare no competing interests.

Funding

This work was supported by the Cabinet Office, Government of Japan, Cross-ministerial Strategic Innovation Promotion Program (SIP), “Technologies for Smart Bio-industry and Agriculture” (funding agency: Bio-oriented Technology Research Advancement Institution, NARO).

Authors’ contributions

SN participated in the design of the experiments, genotyped SNPs, interpreted the data, and wrote the paper. TH, KS, and SM contributed reagents and analysis tools. TS, NT, ST, and YT prepared the plant materials. AI interpreted the data and contributed funding. All the authors read and approved the manuscript.

Acknowledgments

We are deeply indebted to all the people involved in the Japanese pear breeding program at the Institute of Fruit Tree and Tea Science, NARO, Japan.

References

1. Silva GJ, Souza TM, Barbieri RL, Oliveira ACD: Origin, domestication, and dispersing of pear (*Pyrus* spp.). Adv Agr. 2014; 2014: e541097.
2. Kikuchi A. Speciation and taxonomy of Chinese pears. Collected Records of Hort Res Kyoto Univ. 1946; 3:1–8.
3. Gu C, Sponberg SA. *Pyrus*: In: Wu Zy, Raven PH, Honf DY, editors. Flora of China. Vol. 9. Beijing: Science Press; 2003. p.173-19.
4. Yü TT, Ku TC. *Pyrus*: In: Yü TT, editor. Flora Reipublicae Popularis Sinicae. Vol. 36. Beijing: Science Press; 1974. p.354-72.

5. Iketani H, Yamamoto T, Katayama H, Uematsu C, Mase N, Sato Y. Introgression between native and prehistorically naturalized (archaeophytic) wild pear (Pyrus spp.) populations in Northern Tohoku, Northeast Japan. Conserv Genet. 2010; 11:115–26.
6. Jiang ZW, Tang FY, Huang HW, Hu HJ, Chen QL. Assessment of genetic diversity of Chinese sand pear landraces (*Pyrus pyrifolia* Nakai) using simple sequence repeat markers. HortScience. 2009;44:619-26.
7. Nishio S, Takada N, Saito T, Yamamoto T, Iketani H. Estimation of loss of genetic diversity in modern Japanese cultivars by comparison of diverse genetic resources in Asian pear (*Pyrus* spp.). BMC Genet. 2016; 17:81.
8. Iwata H, Hayashi T, Terakami S, Takada N, Sawamura Y, Yamamoto T. Potential assessment of genome-wide association study and genomic selection in Japanese pear *Pyrus pyrifolia*. Breeding Sci. 2013;63:125-140.
9. Nishio S, Saito T, Terakami S, Takada N, Kato H, Itai A, et al. Identification of QTLs associated with conversion of sucrose to hexose in mature fruit of Japanese pear. Plant Mol Biol Rptr. 2018;36:643-52.
10. Minamikawa MF, Takada N, Terakami S, Saito T, Onogi A, Kajiya-Kanegae H, et al. Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia* Nakai). Sci Reports. 2018;8:11994.
11. Yamamoto T, Terakami S, Takada N, Nishio S, Onoue N, Nishitani C, et al. Identification of QTLs controlling harvest time and fruit skin color in Japanese pear (*Pyrus pyrifolia* Nakai). Breeding Sci. 2014;64:351-61.
12. Yamaki S. Metabolism and accumulation of sugars translocated to fruit and their regulation. J Jpn Soc Hort Sci. 2010;79:1-15.
13. Kliewer WM. Sugars and organic acids of *Vitis vinifera*. Plant Physiol. 1966;41:923-31.
14. Moriguchi T, Abe K, Sanada T, Yamaki S. Levels and Role of sucrose synthase, sucrose-phosphate synthase, and acid invertase in sucrose accumulation in fruit of Asian pear. J Am Soc Hort Sci. 1992;117:274-8.
15. Moriguchi T, Ishizawa Y, Sanada T. Differences in sugar composition in *Prunus persica* fruit and the classification by the principal component analysis. J Jpn Soc Hort Sci. 1990; 59:307-12.
16. Moriguchi T, Sanada T, Yamaki S. Seasonal fluctuations of some enzymes relating to sucrose and sorbitol metabolism in peach fruit. J Am Soc Hort Sci. 1990; 115:278-81.
17. Yamaki S, Moriguchi T. Seasonal fluctuation of sorbitol-related enzymes and invertase activities accompanying maturation of Japanese Pear (*Pyrus serotina* Rehder var. *culta* Rehder) fruit. J Jpn Soc for Hortic Sci. 1989;57:602-7.
18. Doty T. Fructose sweetness: a new dimension. Cereal Foods World. 1976; 21:62-3.
19. Pancoast HM, Junk WR. Handbook of sugars. Westport, Connecticut, USA: AVI Publishing Company; 1980:387-9.

20. Pangborn R. Relative taste intensities of selected sugars and organic acids. J Food Sci. 1963;28:726-33.
21. Kajiura I, Yamaki S, Omura M, Akihama T, Machida Y. Improvement of sugar content and composition in fruits, and classifications of East Asian pears by the principal component analysis of sugar compositions in fruits. Jpn J Breed 1979;29:1-12.
22. Hecke K, Herbinger K, Veberic R, Trobec M, Toplak H, Stampar F, et al. Sugar-, acid- and phenol contents in apple cultivars from organic and integrated fruit cultivation. Eur J Clin Nutr. 2006;60:1136-40.
23. Wu J, Gao H, Zhao L, Liao X, Chen F, Wang Z, Hu X. Chemical compositional characterization of some apple cultivars. Food Chem. 2007;103:88-93.
24. Byrne DH, Nikolic AN, Burns EE. Variability in sugars, acids, firmness, and color characteristics of 12 peach genotypes. J Am Soc Hort Sci. 1991;116:1004-6.
25. Guan Y, Peace C, Rudell D, Verma S, Evans K. QTLs detected for individual sugars and soluble solids content in apple. Mol Breed. 2015;35:135.
26. Zhang RP, Wu J, Li XG, Khan MA, Chen H, Korban SS, et al. An AFLP, SRAP, and SSR genetic linkage map and identification of QTLs for fruit traits in Pear (*Pyrus* L.). Plant Mol Biol Rptr. 2013; 31:678-87.
27. Li H, Huang C-H, Ma H. Whole-genome duplications in pear and apple. In The Pear Genome. Edited by Korban SA. New York: Springer; 2019: 279-99.
28. Sassa H. Molecular mechanism of the S-RNase-based gametophytic self-incompatibility in fruit trees of Rosaceae. Breeding Sci. 2016;66:116-21.
29. Moriya S, Terakami S, Okada K, Shimizu T, Adachi Y, Katayose Y, et al. Identification of candidate genes responsible for the susceptibility of apple (*Malus × domestic*a Borkh.) to Alternaria blotch. BMC Plant Biol. 2019;19:132.
30. Terakami S, Moriya S, Adachi Y, Kunihisa M, Nishitani C, Saito T, et al. Fine mapping of the gene for susceptibility to black spot disease in Japanese pear (*Pyrus pyrifolia* Nakai). Breeding Sci. 2016;66:271-80.
31. Itai A, Kotaki T, Tanabe K, Tamura F, Kawaguchi D, Fukuda M. Rapid identification of 1-aminocyclopropane-1-carboxylate (ACC) synthase genotypes in cultivars of Japanese pear (*Pyrus pyrifolia* Nakai) using CAPS markers. Theor Appl Genet. 2003; 106:1266-72.
32. Sato T, Kudo T, Akada T, Wakasa Y, Niizeki M, Harada T. Allelotype of a ripening-specific 1-aminocyclopropane-1-carboxylate synthase gene defines the rate of fruit drop in apple. J Am Soc Hort Sci. 2004;129:32-6.
33. Nishio S, Hayashi T, Yamamoto T, Yamada M, Takada N, Kato H, et al. Validation of molecular markers associated with fruit ripening day of Japanese pear (*Pyrus pyrifolia* Nakai) using variance components. Sci Horti. 2016;199:9-14.
34. Kenis K, Keulemans J, Davey MW. Identification and stability of QTLs for fruit quality traits in apple. Tree Genet Genomes. 2008;4:647-61.

35. Kunihisa M, Moriya S, Abe K, Okada K, Haji T, Hayashi T, et al. Identification of QTLs for fruit quality traits in Japanese apples: QTLs for early ripening are tightly related to preharvest fruit drop. Breeding Sci. 2014;64:240-51.
36. Liebhard R, Kellerhals M, Pfammatter W, Jertmini M, Gessler C. Mapping quantitative physiological traits in apple (*Malus × domestica* Borkh.). Plant Mol Biol. 2003;52:511-26.
37. Cappai F, De Franceschi P, Ciriani A, Collina M, Dondini L. QTLs for susceptibility to *Stemphylium vesicarium* in pear. Mol Breed. 2018;38:24.
38. Daccord N, Celton J-M, Linsmith G, Becker C, Choisne N, Schijlen E, et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet. 2017;49:1099.
39. Chagné D, Crowhurst RN, Pindo M, Thrimawithana A, Deng C, Ireland H, et al. The draft genome sequence of European pear (*Pyrus communis* L. 'Bartlett'). PLoS One. 2014;9:e92644.
40. Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, et al. The genome of the pear (*Pyrus bretschneideri* Rehd.). Genome Res. 2013; 23:396-408.
41. Ishimizu T, Inoue K, Shimonaka M, Saito T, Terai O, Norioka S. PCR-based method for identifying the S-genotypes of Japanese pear cultivars. Theor Appl Genet. 1999;98:961-7.
42. Okada K, Tonaka N, Moriya Y, Norioka N, Sawamura Y, Matsumoto T, et al. Deletion of a 236 kb region around S4-RNase in a stylar-part mutant S4sm-haplotype of Japanese pear. Plant Mol Biol. 2008;66:389.
43. Terakami S, Adachi Y, Iketani H, Sato Y, Sawamura Y, Takada N, et al. Genetic mapping of genes for susceptibility to black spot disease in Japanese pears. Genome. 2007;50:735-41.
44. Terakami S, Shoda M, Adachi Y, Gonai T, Kasumi M, Sawamura Y, et al. Genetic mapping of the pear scab resistance gene Vnk of Japanese pear cultivar Kinchaku. Theor Appl Genet. 2006; 113:743-52.
45. Dirlewanger E, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P, Howad W, et al. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc of the Nat Acad Sci of the USA. 2004; 101:9891-6.
46. Ru S, Main D, Evans K, Peace C. Current applications, challenges, and perspectives of marker-assisted seedling selection in Rosaceae tree fruit breeding. Tree Genet Genomes. 2015;11:1-12.
47. Iwata H, Minamikawa MF, Kajiya-Kanegae H, Ishimori M, Hayashi T. Genomics-assisted breeding in fruit trees. Breeding Sci. 2016;66:100-115.
48. Luby JJ, Shaw DV. Does marker-assisted selection make dollars and sense in a fruit breeding program? HortScience. 2001;36:872-9.
49. Meuwissen T, Hayes B, Goddard M. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157:1819-29.
50. Saito T, Takada N, Kato H, Terakami S, Nishio S. Genotypic variation in and environmental variance components of sugar composition in Japanese pear fruit. HortScience. 2019;54:1465-9.

51. Abe K, Sato Y, Saito T, Kurihara A, Kotobuki K. Genetic correlation between ripening time and weight of fruits in Japanese pear (*Pyrus pyrifolia* Nakai). Jpn J Breed. 1993;43:439-47.
52. Moriya S, Kunihisa M, Okada K, Iwanami H, Iwata H, Minamikawa M, et al. Identification of QTLs for flesh mealiness in apple (*Malus × domestica* Borkh.). Hortic J. 2017;86:159-70.
53. Kumar S, Garrick DJ, Bink MC, Whitworth C, Chagné D, Volz RK. Novel genomic approaches unravel genetic architecture of complex traits in apple. BMC Genomics. 2013;14:393.
54. Leforestier D, Ravon E, Muranty H, Cornille A, Lemaire C, Giraud T, et al. Genomic basis of the differences between cider and dessert apple varieties. Evol Appl. 2015;8:650-61.
55. Saito T. Advances in Japanese pear breeding in Japan. Breeding Sci. 2016;66:46-59.
56. Kiyosue T, Abe H, Yamaguchi-Shinozaki K, Shinozaki K. ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encodes a putative sugar transporter. Biochimica et Biophysica Acta (BBA)-Biomembranes. 1998; 1370:187-91.
57. Afoufa-Bastien D, Medici A, Jeauffre J, Coutos-Thévenot P, Lemoine R, Atanassova R, Laloi M. The Vitis vinifera sugar transporter gene family: phylogenetic overview and macroarray expression profiling. BMC Plant Biol. 2010;10:245.
58. Zhang C, Bian Y, Hou S, Li X. Sugar transport played a more important role than sugar biosynthesis in fruit sugar accumulation during Chinese jujube domestication. Planta. 2018;248:1187-99.
59. Larsen B, Migicovsky Z, Jeppesen AA, Gardner KM, Toldam-Andersen TB, Myles S, et al. Genome-wide association studies in apple reveal loci for aroma volatiles, sugar composition, and harvest date. Plant Genome. 2019; 12:180104.
60. Eduardo I, Picañol R, Rojas E, Batlle I, Howad W, Aranzana M, et al. Mapping of a major gene for the slow ripening character in peach: co-location with the maturity date gene and development of a candidate gene-based diagnostic marker for its selection. Euphytica. 2015;205:627-36.
61. Pirona R, Eduardo I, Pacheco I, Linge CDS, Miculan M, Verde I, et al. Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. BMC Plant Biol. 2013;13:166.
62. Bai Y, Dougherty L, Li M, Fazio G, Cheng L, Xu K. A natural mutation-led truncation in one of the two aluminum-activated malate transporter-like genes at the Ma locus is associated with low fruit acidity in apple. Mol Genet Genomics. 2012; 287:663-78.
63. Verma S, Evans K, Guan Y, Luby JJ, Rosyara U, Howard NP, et al. Two large-effect QTLs, Ma and Ma3, determine genetic potential for acidity in apple fruit: breeding insights from a multi-family study. Tree Genet Genomes. 2019; 15:18.
64. Tamura F. Japanese pear. In: Jpn Soc Hort Sci, editor. Horticulture in Japan 2006. Kyoto: Shoukadoh Publication; 2006. p.50-8.
65. Kajiura I, Suzuki K, Yamazaki T. Color chart for Japanese pear (*Pyrus serotina* var. *culta* Rehder). HortScience. 1975;10:257-258

66. Shirasawa K, Hirakawa H, Isobe S. Analytical workflow of double-digest restriction site-associated DNA sequencing based on empirical and in silico optimization in tomato. DNA Res. 2016; 23:145-53.
67. Acquadro A, Barchi L, Gramazio P, Portis E, Vilanova S, Comino C, et al. Coding SNPs analysis highlights genetic relationships and evolution pattern in eggplant complexes. PLoS One. 2017;12:e0180774.
68. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-60.
69. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078-9.
70. Browning BL, Browning SR. Genotype imputation with millions of reference samples. Am J Hum Genet. 2016;98:116-26.
71. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009; 19:1655-64.
72. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559-75.
73. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources. 2015;15:1179-91.
74. Wickham H. Ggplot2: Elegant Graphics for Data Analysis. New York: Springer; 2009.
75. Endelman JB. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome. 2011;4:250-5.
76. Hayashi, T., Iwata, H., 2013. A Bayesian method and its variational approximation for prediction of genomic breeding values in multiple traits. BMC Bioinformatics 14, 34.
77. Nishio S, Hayashi T, Yamamoto T, Terakami S, Iwata H, Imai A, Takada N, Kato H, Saito T. Bayesian genome-wide association study of nut traits in Japanese chestnut. Mol Breed. 2018; 38:99.
78. Warnes G, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B: gplots: Various R programming tools for plotting data. The Comprehensive R Archive Network. 2009

Tables

Due to technical limitations, tables are only available as a download in the Supplemental Files section.

Figures

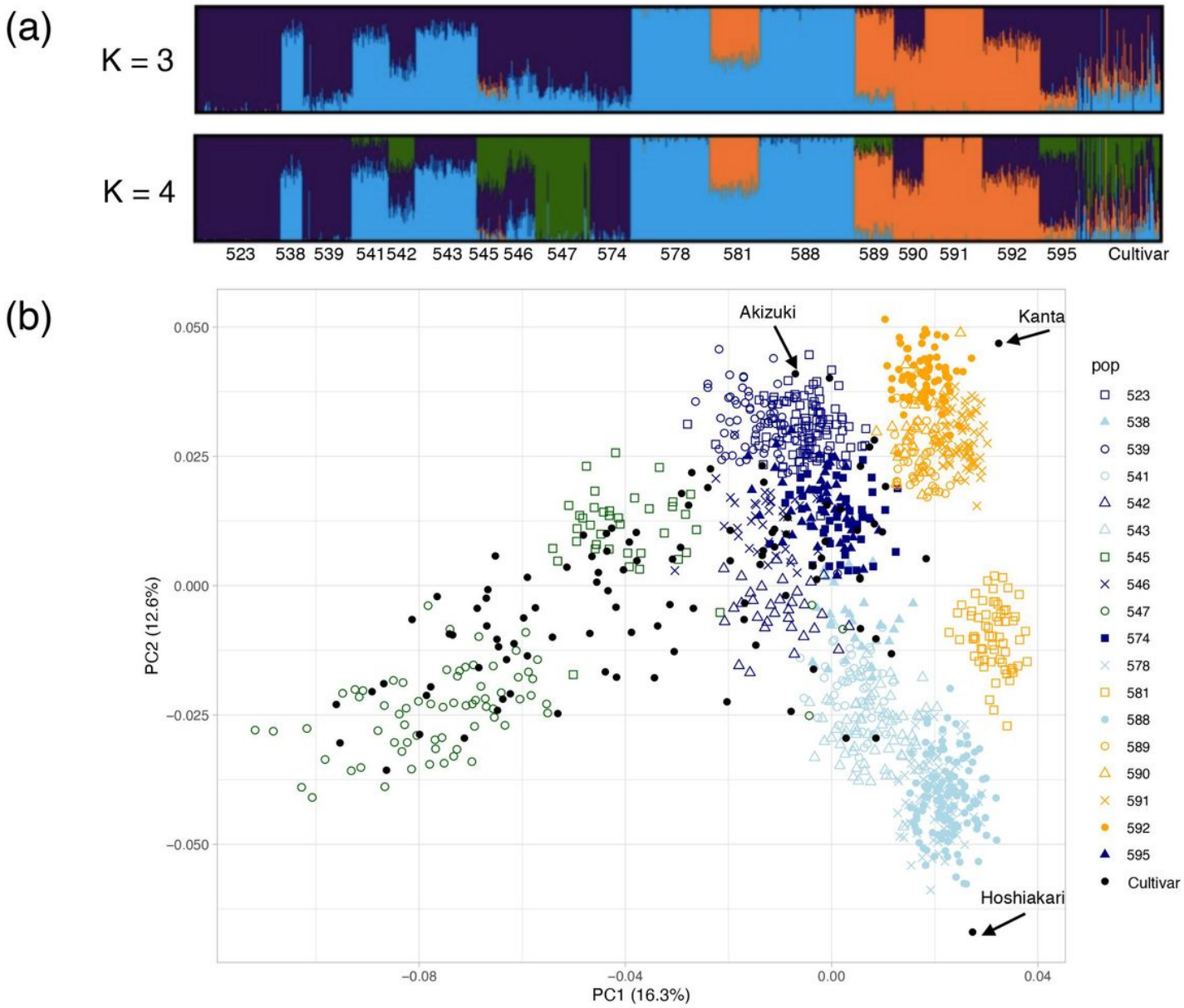

Figure 1

Genetic structure of plant materials used in this study. a Population structure estimated for each cross and the cultivar collection in software ADMIXTURE 1.30 at K = 3 and K = 4. b Principal component analysis was performed in software PLINK v1.90 using 3484 tag single-nucleotide polymorphisms (SNPs). The color used for each population in b corresponds to the color of its predominant cluster at K = 4 in a.

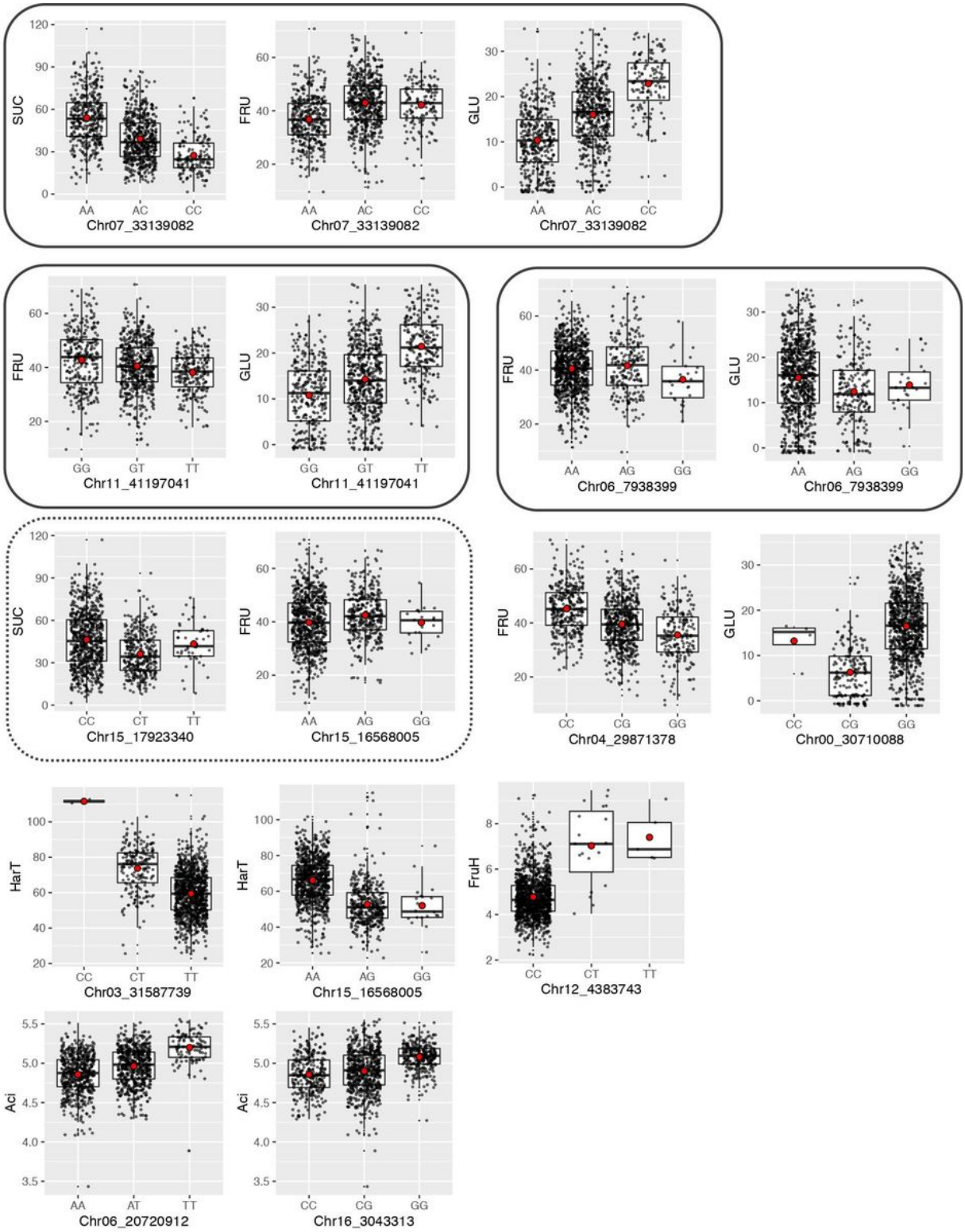

Figure 2

Box plots showing the association of SNP genotypes with 6 fruit traits. The SNPs shown explained more than 10% of the phenotypic variance explained for the indicated trait(s). Box plots for different traits associated with the same SNP are enclosed by rounded rectangles. Chr15_1792340 for SUC and Chr15_16568005 for FRU, which are close to one another, are enclosed by a rounded rectangle with a dotted line. Chr04_29871378 for FRU and Chr12_4383743 for FruH were identified only in the vBayesB-

based GWAS analysis (Table 4); the rest were identified in the MLM-based GWAS analysis (Table 3) or in both. The fictive chromosome 0 contains all unassigned scaffolds

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.pdf
- Table2.pdf
- Table3.pdf
- Table4.pdf
- Table5.pdf
- 210202SUPFigs.pptx
- 210202SUPTable1.xlsx