Development of SSR Databases Available for Both NGS and Capillary Electrophoresis in Apple, Pear and Tea

Sogo Nishio 1,*, Miyuki Kunihisa 1, Fumiya Taniguchi 1, Hiromi Kajiya-Kanegae 2, Shigeki Moriya 3, Yukie Takeuchi 1 and Yutaka Sawamura 3

1 Institute of Fruit Tree and Tea Science, NARO, Tsukuba 305-8605, Japan; miyuky@affrc.go.jp (M.K.); fumiya@affrc.go.jp (F.T.); takeuchiy120@affrc.go.jp (Y.T.)
2 Research Center for Agricultural Information Technology, NARO, Tokyo 105-0003, Japan; kanegae976@affrc.go.jp
3 Institute of Fruit Tree and Tea Science, NARO, Morioka 020-0123, Japan; moriyas@affrc.go.jp (S.M.); ysawa@affrc.go.jp (Y.S.)
* Correspondence: nishios@affrc.go.jp

Abstract: Developing new varieties in fruit and tea breeding programs is very costly and labor-intensive. Thus, establishing a variety discrimination system is important for protecting breeders’ rights and producers’ profits. Simple sequence repeat (SSR) databases that can be utilized for both next-generation sequencing (SSR-GBS) and polymerase chain reaction–capillary electrophoresis (PCR-CE) would be very useful in variety discrimination. In the present study, SSRs with tri-, tetra- and pentanucleotide repeats were examined in apple, pear and tea. Out of 37 SSRs that showed clear results in PCR-CE, 27 were suitable for SSR-GBS. Among the remaining markers, there was allele dropout for some markers that caused differences between the results of PCR-CE and SSR-GBS. For the selected 27 markers, the alleles detected by SSR-GBS were comparable to those detected by PCR-CE. Furthermore, we developed a computational pipeline for automated genotyping using SSR-GBS by setting a value “α” for each marker, a criterion whether a genotype is homozygous or heterozygous based on allele frequency. The set of 27 markers contains 10, 8 and 9 SSRs for apple, pear and tea, respectively, that are useful for both PCR-CE and SSR-GBS and suitable for automation. The databases help researchers discriminate varieties in various ways depending on sample size, markers and methods.

Keywords: variety discrimination; microsatellite; automation

1. Introduction

Fruit and tea breeding requires tremendous cost and time to release varieties because of their long juvenile stage and large plant size [1]. Once elite genotypes are selected and popularized by producers and consumers, the genotypes are able to be duplicated easily by clonal propagation. For this reason, it is difficult to protect breeders’ rights because breeders cannot control illegal spreads of scions. The period of protection for perennial plants is long, 25 years, as determined by the International Union for the Protection of New Varieties of Plants. Unlike vegetable and flower breeding programs, which are mainly supported by private companies, fruit and tea breeding programs have been mainly supported by governments, universities and individuals interested in plant breeding [2–4]. In Japan, because some elite new varieties registered by public research organizations were transported to other countries without authorization, Japanese organizations and breeders have started to register varieties in other countries and have engaged in developing variety identification methods that must be robust and reliable [5]. To further develop varieties for fruit and tea, it is important to protect breeders’ rights and continue organized breeding programs.
Simple sequence repeats (SSRs) are neutral markers that are used mainly in population genetics as well as in conservation ecology and phylogenetic analysis [6–9]. These markers have also been applied in sensitive studies such as forensic analyses [10], parentage analyses [11–13] and variety discrimination [14,15]. However, SSR detection has been difficult to automate because fragments from polymerase chain reaction–capillary electrophoresis (PCR-CE) have usually been evaluated by eye to determine genotypes. Due to limited multiplexing in PCR, the cost and time to determine genotypes has remained high. Slippage and stutter bands in fragment analysis have made allele determination difficult, resulting in the possibility of increased mis-genotyping [16]. Recently, more and more single-nucleotide polymorphism (SNP) analyses based on genotyping by sequencing (GBS) or double-digest restriction-site-associated DNA sequencing (ddRAD-seq) have been applied to the genotyping of plant materials [17,18]. Even so, SSRs have some advantages. First, genetic databases for SSRs were developed in many species [19–22], and some of the materials applied in the preparation of these databases were blighted and became unavailable. By using these same SSR marker sets, researchers can compare past materials with current ones. Second, if the quality and quantity of DNA are too low to construct libraries by GBS or ddRAD-seq, PCR-CE can be applied to SSRs. Therefore, SSRs can even be used to discriminate varieties in processed products or to genotype single pollen grains [23,24]. In addition, PCR-CE with SSRs can be applied easily and flexibly to plant collections of any size without utilizing next-generation sequencing (NGS).

Recently, some methods that apply NGS techniques for genotyping SSRs, called SSR-GBS [25] or SSR-seq [26], have been developed, and some of the disadvantages of SSRs have been overcome. These methods have speeded up genotyping, determined allele size precisely and identified SNPs in SSR flanking regions. However, manual scoring is still necessary to determine genotypes for some markers, especially for dinucleotide repeats with stutter bands [27]. Some software or pipelines have been developed for automated SSR allele calling [25,26,28], but no studies have tested whether the genotypes of a variety collection determined by SSR-GBS and PCR-CE are comparable for variety discrimination. Additionally, the repeatability necessary for variety discrimination has not been adequately confirmed for SSR-GBS.

To overcome these disadvantages of SSRs, databases that can be utilized for both PCR-CE and SSR-GBS would be very helpful for use in variety discrimination. Depending on the objective, researcher preference and sample size, researchers can select a suitable way to identify genotypes by using these databases. The objective of the present study was to select SSR markers that are suitable for both PCR-CE and SSR-GBS and to validate their repeatability in SSR-GBS. In order to develop automatic allele calling and to create a pipeline that is easy to implement, we mainly used tri-, tetra- and pentanucleotide repeats that do not produce stutter bands. To assess stability and repeatability, we replicated the SSR-GBS analysis in three different species: apple (Malus domestica Borkh.), pear (Pyrus pyrifolia Nakai.) and tea (Camellia sinensis (L.) O. Kuntze).

2. Results

2.1. Genotyping by PCR-CE

Prior to determining the SSR genotypes in each variety, SSR positions were confirmed using BLAST+ [29] against the ‘Golden Delicious’ double-haploid (GDDH13) apple genome [30], the Japanese pear ‘Nijisseiki’ genome [31] and a draft genome of C. sinensis var. sinensis ‘Shuchazao’ (CSS-SCZ) [32] (Table S1). All of the markers except Hi15h12 in apple and TM350 in tea were matched to specific mapped regions of reference genome sequences within repeat-containing regions. Hi15h12 was mapped to fictive chromosome (0) of GDDH13, and TM350 was mapped to contig805 of CSS-SCZ. The remaining markers were mapped to different chromosome regions within each genome, which is beneficial for variety discrimination.

The markers identified for each species were successfully genotyped by PCR-CE in all of the corresponding varieties (43 apple, 29 pear and 44 tea) without any missing
values (Table 1 and S2–S4). The SSR fragments were clear, and signals for each allele were constant for each marker. The average observed heterozygosity, expected heterozygosity and polymorphism information content were, respectively, 0.62, 0.59 and 0.52 in apple, 0.46, 0.41 and 0.35 in pear and 0.56, 0.52 and 0.46 in tea (Table 2). In each of the three variety collections, all of the varieties used in this study were successfully differentiated based on genotypes by using those marker sets. The unweighted pair group method with arithmetic mean (UPGMA) dendrograms were created in each variety collection, showing some cultivars that contain the genetic structure of different species are genetically distant from the main varieties used in this study (Figure S1). For example, JM varieties developed from crossing between M. prunifolia Borkh and M. pumila Miller in apple formed a subgroup which is distant from M. domestica varieties. ‘Cha Chukanbohon No. 3’, which is C. sinensis var. assamica, ‘Cha Chukanbohon No. 6’, which is a chance seedling of C. taliensis and their relatives formed a subcluster that is distant from the main tea varieties used in this study (C. sinensis var. sinensis).

Table 1. ID numbers, variety names and accession numbers for the 116 varieties used in this study, including 43 apple, 29 pear and 44 tea varieties.

| ID     | Variety | Accession No. | ID     | Variety | Accession No. | ID     | Variety | Accession No. |
|--------|---------|---------------|--------|---------|---------------|--------|---------|---------------|
| APPLE01| Akane   | JP169082      | PEAR01| Akikake  | JP118536      | TEA01  | Ryofu   | NA            |
| APPLE02| Akibae  | NA            | PEAR02| Akizuki  | JP118538      | TEA02  | Cha     | C. sinensis var. assamica |
| APPLE03| Iwakami | JP114453      | PEAR03| Atago    | JP113570      | TEA03  | Cha     | C. sinensis var. sinensis |
| APPLE04| Indou   | JP169681      | PEAR04| Aranegawa| JP113562      | TEA04  | Cha     | C. sinensis var. sinensis |
| APPLE05| Orin    | JP172644      | PEAR05| Oushu    | JP118539      | TEA05  | Cha     | C. sinensis var. sinensis |
| APPLE06| Ozenokurena | NA | PEAR06| Okusankichi | JP113634 | TEA06  | Cha     | C. sinensis var. sinensis |
| APPLE07| Gala    | JP114085      | PEAR07| Natsushizuku| JP230439 | TEA07  | Cha     | C. sinensis var. sinensis |
| APPLE08| Kie     | NA            | PEAR08| Kamo     | JP113623      | TEA08  | Cha     | C. sinensis var. sinensis |
| APPLE09| Kizashi | JP114525      | PEAR09| Kousui   | JP113619      | TEA09  | Cha     | C. sinensis var. sinensis |
| APPLE10| Kitakame| JP114148      | PEAR10| Shuagyouko| JP113707      | TEA10  | Cha     | C. sinensis var. sinensis |
| APPLE11| Kitaro  | NA            | PEAR11| Shuante  | JP118537      | TEA11  | Cha     | C. sinensis var. sinensis |
| APPLE12| Kinsei  | JP172634      | PEAR12| Shinkou  | JP113657      | TEA12  | Cha     | C. sinensis var. sinensis |
| APPLE13| Jonathan| JP169679      | PEAR13| Shinsei  | JP113660      | TEA13  | Cha     | C. sinensis var. sinensis |
| APPLE14| Kotoro  | NA            | PEAR14| Shinsei  | JP113694      | TEA14  | Cha     | C. sinensis var. sinensis |
| APPLE15| Golden  | JP172622      | PEAR15| Suisei   | JP113665      | TEA15  | Cha     | C. sinensis var. sinensis |
| APPLE16| Genda   | JP114453      | PEAR16| Chikusai | JP113716      | TEA16  | Cha     | C. sinensis var. sinensis |
| APPLE17| Santaa  | NA            | PEAR17| Choujuurou| JP113574     | TEA17  | Cha     | C. sinensis var. sinensis |
| APPLE18| Shinano | NA            | PEAR18| Nittaika | JP113630      | TEA18  | Cha     | C. sinensis var. sinensis |
| APPLE19| Sweet   | NA            | PEAR19| Nijisseki | JP113631      | TEA19  | Cha     | C. sinensis var. sinensis |
| APPLE20| Sensyu  | JP114288      | PEAR20| Hakko    | JP113585      | TEA20  | Cha     | C. sinensis var. sinensis |
| APPLE21| Chinatsu | NA            | PEAR21| Hayatama  | JP113591      | TEA21  | Cha     | C. sinensis var. sinensis |
| APPLE22| Tsugaru  | JP172651      | PEAR22| Houtetsu  | JP113720      | TEA22  | Cha     | C. sinensis var. sinensis |
| APPLE23| Delicious| JP114041     | PEAR23| Housui   | JP113598      | TEA23  | Cha     | C. sinensis var. sinensis |
| APPLE24| Hatsuzuki| JP169705      | PEAR24| Yatsuzo  | JP113718      | TEA24  | Cha     | C. sinensis var. sinensis |
| APPLE25| Himekami| JP169716      | PEAR25| Hatsumaru| NA            | TEA25  | Cha     | C. sinensis var. sinensis |
| APPLE26| Fuji    | JP114078      | PEAR26| Rinka    | NA            | TEA26  | Cha     | C. sinensis var. sinensis |
| APPLE27| Hokuto  | JP169715      | PEAR27| Hoshiaika| NA            | TEA27  | Cha     | C. sinensis var. sinensis |
| APPLE28| Maypole | NA            | PEAR28| Narumi   | NA            | TEA28  | Cha     | C. sinensis var. sinensis |
| APPLE29| Mori-no-| NA            | PEAR29| Kanta    | NA            | TEA29  | Cha     | C. sinensis var. sinensis |
| APPLE30| Yoko    | JP114416      | PEAR30| Kanta    | NA            | TEA30  | Cha     | C. sinensis var. sinensis |
| APPLE31| Ruby    | NA            | PEAR31| Kanta    | NA            | TEA31  | Cha     | C. sinensis var. sinensis |
| APPLE32| Rose    | NA            | PEAR32| Kanta    | NA            | TEA32  | Cha     | C. sinensis var. sinensis |
| APPLE33| JM1     | NA            | PEAR33| Kanta    | NA            | TEA33  | Cha     | C. sinensis var. sinensis |
| APPLE34| JM2     | NA            | PEAR34| Kanta    | NA            | TEA34  | Cha     | C. sinensis var. sinensis |
| APPLE35| JM3     | NA            | PEAR35| Kanta    | NA            | TEA35  | Cha     | C. sinensis var. sinensis |
| APPLE36| JM4     | NA            | PEAR36| Kanta    | NA            | TEA36  | Cha     | C. sinensis var. sinensis |
| APPLE37| JM5     | NA            | PEAR37| Kanta    | NA            | TEA37  | Cha     | C. sinensis var. sinensis |
| APPLE38| Morikota65| NA | PEAR38| Kanta    | NA            | TEA38  | Cha     | C. sinensis var. sinensis |
| APPLE39| Benimaru| NA            | PEAR39| Kanta    | NA            | TEA39  | Cha     | C. sinensis var. sinensis |
| APPLE40| Morioka | NA            | PEAR40| Kanta    | NA            | TEA40  | Cha     | C. sinensis var. sinensis |
| APPLE41| Morioka | NA            | PEAR41| Kanta    | NA            | TEA41  | Cha     | C. sinensis var. sinensis |
| APPLE42| Kirshu  | NA            | PEAR42| Kanta    | NA            | TEA42  | Cha     | C. sinensis var. sinensis |
| APPLE43| Morikota71| NA | PEAR43| Kanta    | NA            | TEA43  | Cha     | C. sinensis var. sinensis |

Table 1. ID numbers, variety names and accession numbers for the 116 varieties used in this study, including 43 apple, 29 pear and 44 tea varieties.

accession no. for NARO Genebank (www.gene.affrc.go.jp). NA means that the variety is not registered in NARO Genebank.
Table 2. Number of alleles, observed and expected heterozygosity and polymorphism information content of the SSR loci determined by PCR-CE.

| Marker         | Plant    | No. of Alleles | $H_O$ | $H_E$ | PIC  |
|----------------|----------|----------------|-------|-------|------|
| CH01b09b<sup>a</sup> | Apple    | 3              | 0.59  | 0.67  | 0.52 |
| Hi08h08        | Apple    | 3              | 0.56  | 0.54  | 0.47 |
| Hi09f01        | Apple    | 4              | 0.79  | 0.63  | 0.57 |
| Hi15h12        | Apple    | 3              | 0.26  | 0.29  | 0.26 |
| Hi22d06        | Apple    | 5              | 0.79  | 0.74  | 0.69 |
| Mdo.chr1.18    | Apple    | 5              | 0.63  | 0.64  | 0.57 |
| NZmsCN943067   | Apple    | 5              | 0.61  | 0.51  | 0.46 |
| NZmsEB116209   | Apple    | 3              | 0.70  | 0.66  | 0.58 |
| NZmsEB119405<sup>a</sup> | Apple | 4              | 0.52  | 0.61  | 0.48 |
| NZmsEB146613   | Apple    | 3              | 0.74  | 0.63  | 0.56 |
| SamsCN944528   | Apple    | 5              | 0.65  | 0.61  | 0.55 |
| SamsEB132187   | Apple    | 3              | 0.65  | 0.60  | 0.52 |
| Averages for apple markers |         | 3.8            | 0.62  | 0.59  | 0.52 |
| TsuGNH124      | Pear     | 2              | 0.55  | 0.49  | 0.37 |
| TsuGNH161      | Pear     | 3              | 0.62  | 0.60  | 0.52 |
| TsuGNH164<sup>a</sup> | Pear | 3              | 0.57  | 0.41  | 0.49 |
| TsuGNH179      | Pear     | 3              | 0.83  | 0.59  | 0.52 |
| TsuGNH184<sup>a</sup> | Pear | 2              | 0.50  | 0.48  | 0.37 |
| TsuGNH194      | Pear     | 2              | 0.35  | 0.37  | 0.30 |
| TsuGNH204      | Pear     | 2              | 0.03  | 0.03  | 0.03 |
| TsuGNH207      | Pear     | 2              | 0.28  | 0.24  | 0.21 |
| TsuGNH208      | Pear     | 3              | 0.45  | 0.52  | 0.46 |
| TsuGNH250      | Pear     | 2              | 0.41  | 0.33  | 0.27 |
| Averages for pear markers |         | 2.4            | 0.46  | 0.41  | 0.35 |
| CsFM1097       | Tea      | 4              | 0.43  | 0.49  | 0.45 |
| CsFM1206       | Tea      | 3              | 0.50  | 0.57  | 0.50 |
| CsFM1566       | Tea      | 4              | 0.61  | 0.49  | 0.45 |
| CsFM1595<sup>a</sup> | Tea | 5              | 0.62  | 0.63  | 0.57 |
| MSE0348        | Tea      | 5              | 0.71  | 0.61  | 0.56 |
| MSE0354<sup>a</sup> | Tea | 4              | 0.64  | 0.80  | 0.57 |
| TM043<sup>a</sup> | Tea      | 5              | 0.45  | 0.46  | 0.40 |
| TM107          | Tea      | 6              | 0.80  | 0.69  | 0.64 |
| TM336<sup>a</sup> | Tea     | 3              | 0.35  | 0.30  | 0.31 |
| TM348          | Tea      | 5              | 0.64  | 0.59  | 0.51 |
| TM350<sup>a</sup> | Tea     | 4              | 0.36  | 0.36  | 0.32 |
| TM464<sup>a</sup> | Tea     | 7              | 0.51  | 0.57  | 0.45 |
| TM485          | Tea      | 3              | 0.39  | 0.35  | 0.30 |
| TM553          | Tea      | 3              | 0.80  | 0.50  | 0.39 |
| TM626          | Tea      | 4              | 0.55  | 0.46  | 0.41 |
| Averages for tea markers |         | 4.3            | 0.56  | 0.52  | 0.46 |

$H_O$ = Observed heterozygosity, $H_E$ = Expected heterozygosity, PIC = Polymorphism information content.<sup>a</sup> This marker was not suitable for SSR-GBS because of the presence of allele dropout.

2.2. Genotyping by SSR-GBS

Illumina short-read sequencing data from 9.5 M reads, 7.7 M reads and 7.7 M reads were obtained from the 2 replicates of 43 varieties of apple (average of 110.2 K reads), 29 varieties of pear (average of 132.9 K reads) and 44 varieties of tea (average of 133.0 K reads), respectively. After combining the reads from forward and reverse sequences, 4.2 M, 3.8 M and 3.7 M reads were obtained, respectively. The average number of reads per variety was 49.0 K reads for apple, 66.5 K for pear and 51.4 K for tea. The reads were then demultiplexed based on the primer sequences. The average number of reads per marker ranged from 14 (TM336) to 25,406 (TsuGNH194), with an average of 3866 (Table 3). All 12 SSR markers in apple had an average of more than 50 reads, as did 9 of 10 SSRs in pear and 14 of 15 SSRs in tea. Because the markers that had a small number of reads...
had lower reliability, they were excluded from further analyses. The numbers of reads in the replicated analysis were similar to those in the original analysis. For example, in tea, the number of reads was highest in MSE0354 (10,510 and 8783) and lowest in TM336 (14 and 12) for the 2 replicated analyses. There were several markers that had an insufficient number of reads in a specific variety even though the overall average number of reads was high. The numbers of reads obtained for TsuGNH164 in ‘Oushuu’ and CsFM1595 in ‘Cha Chukanbohon No. 6’ were quite low, making it impossible to determine the genotypes of these markers in these varieties.

Table 3. Summary of each marker tested for SSR-GBS. Average depth indicates the number of the reads obtained after demultiplexing based on forward and reverse primer sequences in each marker. Replicated analyses with a slight change in the multiplex PCR set were conducted for validation.

| Marker      | Plant       | Average Depth (Number of Reads) | Average Depth in Replicate Analysis | Difference in Size from the Alleles in PCR-CE (bp) | Result in SSR-GBS                      |
|-------------|-------------|---------------------------------|-------------------------------------|---------------------------------------------------|---------------------------------------|
| CH01b09b    | Apple       | 6834                            | 5724                                | 1                                                 | Errors due to stutter band in ‘Jonathan’ and ‘Sensyu’ |
| Hi08h08     | Apple       | 342                             | 470                                 | 1                                                 | Comparable to the result of PCR-CE    |
| Hi09f01     | Apple       | 395                             | 405                                 | -4                                                | Comparable to the result of PCR-CE    |
| Hi15s12     | Apple       | 5329                            | 4331                                | 0                                                 | Comparable to the result of PCR-CE    |
| Hi22d06     | Apple       | 3611                            | 4070                                | 3                                                 | Comparable to the result of PCR-CE    |
| Mdo-chr1.18 | Apple       | 1206                            | 1108                                | -2                                                | Comparable to the result of PCR-CE    |
| NZmsCN943067| Apple       | 1777                            | 994                                 | -3                                                | Comparable to the result of PCR-CE    |
| NZmsEBI16209| Apple       | 747                             | 534                                 | 3                                                 | Comparable to the result of PCR-CE    |
| NZmsEBI19405| Apple       | 2442                            | 2481                                | 1                                                 | Comparable to the result of PCR-CE    |
| NZmsEBI146613| Apple      | 8023                            | 4663                                | 1                                                 | Comparable to the result of PCR-CE    |
| SAmCN944528 | Apple       | 6249                            | 4003                                | -1                                                | Comparable to the result of PCR-CE    |
| SAmEBI312817| Apple       | 6621                            | 4034                                | -3                                                | Comparable to the result of PCR-CE    |
| TsuGNH124   | Pear        | 75                              | 760                                 | -2                                                | Comparable to the result of PCR-CE    |
| TsuGNH161   | Pear        | 16,758                          | 23,807                              | 2                                                 | Comparable to the result of PCR-CE    |
| TsuGNH164   | Pear        | 1671                            | 1682                                | -3                                                | The read depths were too low to determine genotype in ‘Cha Chukanbohon No. 6’ |
| TsuGNH179   | Pear        | 21                              | 116                                 | -1                                                | Comparable to the result of PCR-CE    |
| TsuGNH184   | Pear        | 25,406                          | 12,843                              | 1                                                 | Comparable to the result of PCR-CE    |
| TsuGNH194   | Pear        | 25                              | 116                                 | -1                                                | Comparable to the result of PCR-CE    |
| TsuGNH204   | Pear        | 3877                            | 1957                                | -2                                                | Comparable to the result of PCR-CE    |
| TsuGNH207   | Pear        | 2718                            | 1231                                | 2                                                 | Comparable to the result of PCR-CE    |
| TsuGNH208   | Pear        | 7348                            | 3735                                | -1                                                | Comparable to the result of PCR-CE    |
| TsuGNH250   | Pear        | 5021                            | 2847                                | -2                                                | Comparable to the result of PCR-CE    |
| CsFM1097    | Tea         | 1217                            | 1868                                | -4                                                | Comparable to the result of PCR-CE    |
| CsFM1206    | Tea         | 1259                            | 1396                                | -2                                                | Comparable to the result of PCR-CE    |
| CsFM1566    | Tea         | 857                             | 1174                                | -1                                                | Comparable to the result of PCR-CE    |
| CsFM1595    | Tea         | 195                             | 268                                 | 0                                                 | The read depths were too low to determine genotype in ‘Cha Chukanbohon No. 6’ |
| MSE0334     | Tea         | 3580                            | 2559                                | 0                                                 | Comparable to the result of PCR-CE    |
| MSE0354     | Tea         | 10,510                          | 8783                                | 1                                                 | Allele dropout was observed in some varieties |
| TM3043      | Tea         | 3469                            | 2526                                | 0                                                 | Allele dropout was observed in some varieties |
| TM107       | Tea         | 2249                            | 1645                                | -3                                                | Comparable to the result of PCR-CE    |
| TM336       | Tea         | 14                              | 12                                  | 1                                                 | The read depths were too low to determine genotypes |
| TM348       | Tea         | 691                             | 577                                 | -2                                                | Comparable to the result of PCR-CE    |
| TM350       | Tea         | 514                             | 357                                 | -3                                                | Allele dropout was observed in some varieties |
| TM364       | Tea         | 546                             | 371                                 | 0                                                 | Allele dropout was observed in some varieties |
| TM485       | Tea         | 2365                            | 2947                                | -2                                                | Comparable to the result of PCR-CE    |
| TM553       | Tea         | 6800                            | 8285                                | 0                                                 | Comparable to the result of PCR-CE    |
| TM626       | Tea         | 1605                            | 2288                                | 0                                                 | Comparable to the result of PCR-CE    |

Using the pipeline described in Materials and Methods, we calculated the allele frequencies for each marker in each variety. For each marker, we retrieved allele frequencies for only the four most common alleles, which provided enough information to call the genotypes and understand the details of the PCR products (Figure 1). Then, each genotype was estimated manually based on allele frequency. In most cases, the allele frequency of the first allele was >0.8 when the genotype was homozygous and around 0.4 when the genotype was heterozygous. The distribution of allele frequencies fluctuated depending on the marker owing to background arising during amplification and sequencing. For CH01b09b, which has dinucleotide repeats, genotyping errors were found in ‘Jonathan’ and ‘Sensyu’ because of the presence of stutter bands. We used tri-, tetra- and pentanucleotide repeats for all markers except for CH01b09b and Mdo.chr1.18, so there were few stutter bands for the other markers. The detected sizes of some alleles in SSR-GBS were obviously different from those detected in PCR-CE (Figure 2). In particular, NZmsEBI19405 in apple, and MSE0354, TM043, TM350 and TM464 in tea, showed allele dropout (null alleles) in SSR-GBS, meaning that some alleles detected in PCR-CE were undetected in SSR-GBS (Figure 2).
Except for the cases of allele dropout, the alleles detected by SSR-GBS corresponded to those detected by PCR-CE. There were some small differences between digital sizes in SSR-GBS and the artificially determined sizes in PCR-CE, which ranged from −4 bp to 3 bp, but this difference could be corrected by adding or subtracting the indicated value for each marker (Table 3). Following this analysis, 10 SSRs for apple, 8 for pear and 9 for tea were selected as suitable markers to use in SSR-GBS for variety discrimination (Tables S5–S7). All of the varieties used in this study could be successfully differentiated based on genotypes determined by those marker sets.

| Allele frequencies and genotypes of Hi22d06 in ‘Akane’ and ‘Akibae’ |
|---------------------------------------------------------------|
|                  | allele1 | freq.1 | allele2 | freq.2 | allele3 | freq.3 | allele4 | freq.4 | genotype |
|-------------------|---------|--------|---------|--------|---------|--------|---------|--------|-----------|
| Akane             | 126     | 0.94   | 123     | 0.02   | 125     | 0.02   | 168     | 0.01   | 126/126   |
| Akibae            | 126     | 0.50   | 129     | 0.46   | 123     | 0.01   | 128     | 0.01   | 126/129   |

Figure 1. Allele frequencies, genotypes and digital fragments of Hi22d06 in ‘Akane’ and ‘Akibae’ obtained by SSR-GBS. The first alleles and second alleles were indicated in red and blue, respectively.

Figure 2. Comparisons of alleles detected by PCR-CE and SSR-GBS in apple (a), pear (b) and tea (c). The alleles’ length (bp) detected in capillary electrophoresis was compared against digital allele length in SSR-GBS.

2.3. Automated Genotyping in SSR-GBS and Repeatability

To develop an accurate automated genotyping system, we set the value of a new variable, α, for each marker (Table 4). The value of α is used as a criterion whether a genotype is homozygous or heterozygous. If the first (most common) allele frequency is more than α, the genotype is assumed to be homozygous. If the first allele frequency is less than or equal to α, the genotype is called heterozygous for the first and second alleles. The range of α was largest (estimated as 0.44) for TsuGNH207 and TsuGNH250, and it was smallest for Mdo.chr1.18 (estimated as 0.01). The α for Hi15h12 showed low values
This marker produced some slippage reads, resulting in alleles from PCR-CE that were 1 bp larger than the target alleles in SSR-GBS. The $\alpha$ values for TsuGNH179 and TM533 were high (0.89–0.97 and 0.87–0.97, respectively). For these markers, the frequency of the second allele was sometimes less than that of the slippage and stutter band from the first allele, probably due to sequence variation in the primer binding site. Because of this fluctuation, we set the optimum $\alpha$ value for each marker based on the average of the minimum and maximum values, which ranged from 0.43 to 0.98 and averaged 0.75.

### Table 4. Optimum range of variable $\alpha$ and correlation of first-allele frequency between the two replicates for each marker.

The range of $\alpha$ for each marker was determined based on the criterion that the genotypes identified by SSR-GBS and PCR-CE were exactly the same for each of the varieties.

| Marker          | Minimum Value of $\alpha$ | Optimum Value of $\alpha$ | Maximum Value of $\alpha$ | Correlation of First-Allele Frequency between Two Replicates |
|-----------------|---------------------------|---------------------------|---------------------------|-------------------------------------------------------------|
| Hi08h08 Apple   | 0.45                      | 0.57                      | 0.69                      | 0.98                                                        |
| Hi09F01 Apple   | 0.69                      | 0.80                      | 0.90                      | 0.98                                                        |
| Hi15h12 Apple   | 0.34                      | 0.43                      | 0.52                      | 0.98                                                        |
| Hi22d06 Apple   | 0.57                      | 0.74                      | 0.90                      | 1.00                                                        |
| Mdo.chr1.18 Apple | 0.60                  | 0.61                      | 0.61                      | 0.98                                                        |
| NZmsCN943067 Apple | 0.57                  | 0.66                      | 0.74                      | 0.99                                                        |
| NZmsEB116209 Apple | 0.55                  | 0.70                      | 0.84                      | 0.99                                                        |
| NZmsEB146613 Apple | 0.58                  | 0.73                      | 0.87                      | 1.00                                                        |
| SAmCN944528 Apple | 0.64                  | 0.78                      | 0.91                      | 0.99                                                        |
| SAmEB132187 Apple | 0.76                  | 0.82                      | 0.88                      | 1.00                                                        |
| TsuGNH124 Pear  | 0.52                      | 0.60                      | 0.67                      | 0.97                                                        |
| TsuGNH161 Pear  | 0.53                      | 0.73                      | 0.93                      | 1.00                                                        |
| TsuGNH179 Pear  | 0.89                      | 0.93                      | 0.97                      | 0.99                                                        |
| TsuGNH194 Pear  | 0.53                      | 0.75                      | 0.96                      | 1.00                                                        |
| TsuGNH204 Pear  | 0.43                      | 0.62                      | 0.80                      | 0.99                                                        |
| TsuGNH207 Pear  | 0.53                      | 0.75                      | 0.97                      | 1.00                                                        |
| TsuGNH208 Pear  | 0.87                      | 0.89                      | 0.91                      | 0.99                                                        |
| TsuGNH250 Pear  | 0.53                      | 0.75                      | 0.97                      | 1.00                                                        |
| CsFM1097 Tea    | 0.55                      | 0.71                      | 0.87                      | 0.99                                                        |
| CsFM1206 Tea    | 0.63                      | 0.77                      | 0.90                      | 1.00                                                        |
| CsFM1566 Tea    | 0.71                      | 0.83                      | 0.95                      | 0.99                                                        |
| MSE0348 Tea     | 0.73                      | 0.85                      | 0.97                      | 0.99                                                        |
| TM107 Tea       | 0.72                      | 0.85                      | 0.98                      | 0.98                                                        |
| TM348 Tea       | 0.61                      | 0.77                      | 0.93                      | 0.99                                                        |
| TM485 Tea       | 0.72                      | 0.84                      | 0.96                      | 1.00                                                        |
| TM553 Tea       | 0.87                      | 0.92                      | 0.97                      | 1.00                                                        |
| TM626 Tea       | 0.73                      | 0.85                      | 0.97                      | 1.00                                                        |

To estimate the repeatability of the analyses, we conducted replicated analyses with a slight change in the multiplex PCR set. In the first replicate, we used all of the primers for each species in a single multiplex PCR, whereas in the second replicate, the primers for each species were split into two sets. Otherwise, the steps were the same. The correlations of the first allele frequency among 2 replicated analyses were quite high (0.98–1.00) regardless of the marker, suggesting that this method has high repeatability when the same procedure and conditions are used.

### 3. Discussion

We developed databases suitable for both PCR-CE and SSR-GBS containing 43 varieties with 10 SSRs for apple, 29 varieties with 9 SSRs for pear, and 44 varieties with 8 SSRs for tea. We were able to distinguish all the varieties that covered most of the varieties planted in Japan; even the number of the markers were around ten. The databases were available to protect breeders’ rights of those varieties. It can also help researchers discriminate varieties in various ways depending on sample size and number of markers. The cost of
SSR-GBS is lower than that of PCR-CE when the number of varieties tested is large [33]. The allele scores for PCR-CE and SSR-GBS differed somewhat (Table 3, Figure 2) because the allele sizes fluctuated depending on the type of fluorescent dyes and amplified nucleotide composition. By using the varieties in this study as controls, researchers can adjust the scores of PCR-CE to those of SSR-GBS and vice versa.

SSR-GBS has mainly been used in studies of population genetics, and few studies related to variety discrimination have been conducted. Here, we were able to select SSRs that were suitable for variety discrimination using SSR-GBS. Out of 37 SSRs tested in this study, 33 were successfully amplified (all except for TsuGNH164, TsuGNH184, CsFM1595 and TM336), but only 27 were suitable for SSR-GBS. The rate of successful PCR amplification was similar to those in other studies using SSR-GBS [25,34,35]. Out of the 27 SSRs suitable for automated genotyping, 19 had a range of $\alpha$ values of more than 0.2, suggesting that those markers had high reliability for SSR-GBS. Although the digital allele size determined by SSR-GBS is unambiguous, marker allele frequencies may fluctuate because of differences in enzymes, thermal cyclers and other laboratory equipment. The markers that have a greater range of $\alpha$ have equal amounts of amplification of the first and second alleles when heterozygous and have small numbers of nonspecific PCR products. Thus, those markers could be used as anchoring markers to connect databases. One of the merits of SSR-GBS is that it can detect length homoplasy (the occurrence of nonhomologous fragments of the same size) because it reveals SNPs [25]. However, we did not use SNPs to create the databases because the objective of this study was to develop databases available for both NGS and capillary electrophoresis that do not focus on SNPs but rather on the sizes of the PCR products.

Applying multiplex PCR and/or mixing several PCR products is necessary for NGS because the NGS platform needs sequence diversity in order to produce high-quality reads. On the other hand, applying different enzymes suitable for multiplex PCR at the same annealing temperature produced the differences between the results of SSR-GBS and PCR-CE. In this study, some allele dropouts were observed in SSR-GBS, as has been observed in other studies [33,36]. Each primer used in this study has an optimal annealing temperature, but it was not applied in these experiments, because the cost of genotyping and labor for the experimental procedure would increase if we subdivided the marker set for multiplex PCR. It is possible that some primers did not amplify alleles that had somewhat different sequences in the primer binding site. In fact, allele dropout was observed in varieties that were genetically distant from the main varieties in the database, i.e., JM varieties, which were developed from crossing between *M. prunifolia* Borkh and *M. pumila* Miller in apple; ‘Oushuu’, one of whose ancestors is Chinese pear; and ‘Cha Chukanbohon No. 3’, which is *C. sinensis* var. *assamica* introduced from India (Figure S1).

Although the accuracy of SSR-GBS, which suffers from allele dropout, is lower than PCR-CE, there are merits of SSR-GBS such as ease of automation and digital allele length, which does not vary among laboratories. We arranged the pipeline for SSR-GBS using the pipeline of Tibihika et al. [25] as a reference. Our pipeline is useful for complete automation of genotyping after the user sets the value of $\alpha$ for each marker. By using fastq reads, primer set sequences and $\alpha$, the genotype score is easily obtained. The pipeline is simple and is composed of four files with the following functions: (1) combining the pairs of fastq files, (2) demultiplexing based on markers, (3) counting alleles and (4) determining genotypes. This pipeline works in a basic Linux system and is useful not only for variety discrimination but also for marker-assisted selection (MAS). Because the important molecular markers for fruit breeding programs were developed using SSRs and indels, the SSR-GBS pipeline can be applied to practical MAS with a slight change. Specifically, by examining four major alleles, the extent of slippage and stutter bands can be estimated because those increase the frequencies of third and fourth alleles. Even though the third and fourth alleles are not necessary for the determination of genotypes in diploid species, output of that information helps to reduce genotyping errors due to artifacts and helps to confirm whether the result of automatic genotyping is accurate.
In conclusion, the databases of apple, pear and tea developed here can provide basic information for PCR-CE and SSR-GBS. These databases and the methods developed in this study are useful for variety discrimination and protecting breeders’ rights. We also identified markers that have high reliability and are suitable for the automation of SSR-GBS. These highly reliable markers can help to connect databases and facilitate automated genotyping using a simple pipeline.

4. Materials and Methods

4.1. Plant Materials

We used a total of 116 varieties consisting of 43 varieties of apple, 29 varieties of pear and 44 varieties of tea (Table 2). The varieties of apple and pear were preserved at the Institute of Fruit Tree and Tea Science, NARO, in Morioka and Tsukuba, respectively. The tea varieties were preserved at the Institute of Fruit Tree and Tea Science, NARO, Makurazaki, Green Tea Laboratory; Saitama Prefectural Agriculture and Forestry Research Center; and Tea Branch Facility, Miyazaki Agricultural Research Institute. Genomic DNA was extracted from young leaves with a DNeasy Plant Mini Kit (Qiagen, Germany). To understand an overview of genetic relationships among varieties, UPGMA dendrograms were created using R package “pvclust” [37]. The confidence levels for branches of the dendrograms were determined by calculating approximately unbiased (AU) \( p \)-values using multiscale bootstrap resampling based on 10,000 replications.

4.2. SSR Genotyping for PCR-CE

The SSR markers used in this study are listed in Table S1 and Table 2. For apple, 12 SSRs that were mainly tri- and tetranucleotide SSRs and that showed clear fragments in PCR-CE in prior analyses were used for genotyping [22,38–41]. The 10 SSRs for pear were similar to those developed in the study by Yamamoto et al. [42], which were composed of tetra- and pentanucleotide repeats. For tea, 15 SSRs previously developed and composed of tetra- and pentanucleotide repeats were used in this experiment [43,44]. BLAST+ was used to estimate the positions of the SSRs in each reference genome [29]. The recent reference genomes were mainly composed of pseudomolecules that covered most of the genes. Chromosome positions of markers were mapped easily to design and to select optimal marker sets in genetic studies. By using reference genomes, researchers can also conduct high through-put genotyping [17,18]. PCR amplification was performed in a 10-µL solution containing 5 µL of 2× Green GoTaq G2 Hot Start Master Mix (0.4 mM each dNTP, Taq DNA polymerase and 4 mM MgCl₂, pH 8.5; Promega, Madison, WI, USA), 20 pmol of each forward primer labeled with a fluorescent dye (5-FAM, 5-HEX, 5-VIC or 5-NES) and unlabeled reverse primer, and 10 ng of genomic DNA. Amplification was performed in 35 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min. PCR products were separated and detected with a 3130xl Genetic Analyzer or a SeqStudio Genetic Analyzer (Life Technologies, Carlsbad, CA, USA). The size of each amplified band was determined manually by comparison with a set of internal-standard DNA fragments (400HD ROX, Life Technologies) in GeneMapper software v. 5.0 (Life Technologies).

4.3. SSR Genotyping Using SSR-GBS

SSR-GBS was performed based on the methods described in previous reports with slight changes in sequence data analysis [25,34]. When constructing each library for SSR-GBS, two-step PCR was applied to add the Illumina adapter sequences that are required for Illumina flow-cell binding (Figure 3). The first PCR was performed using specific primers extended with the part of the Illumina adapters P5 (ACACTCTTTCCCTACACGACGCTCTTCCGATC) and P7 (GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC) as the forward and reverse primers, respectively. In the first replicate, the previously reported primers, consisting of 12-, 10- and 15-primer sets for apple, pear and tea, respectively, were used in a single multiplex reaction. For the second replicate of this experiment, we divided each primer set into 2 subsets, i.e., 6 primers × 2 for apple, 5 primers × 2 for
Fig. 4. Total protein was extracted from leaves of pear (left) and apple (right) using the CTAB method and separated by 12% SDS-PAGE. Molecular markers were 100-bp DNA ladder and 50-bp DNA ladder, respectively. A, B: apple; C: pear; M: molecular marker. The numbers on the right are the molecular weights (kDa).
Figure 3. Summary of SSR-GBS procedure used in this study. Green indicates the SSR being amplified, with SSR primer annealing sequences (red) on each end. A different set of indexes (orange) is added to each sample to allow the sequences from that sample to be distinguished after multiplexing. The sequences of P5 flow cell binding, P5 sequence primer, P7 flow cell binding and P7 sequence primer were AATGATACGGCGACCACCGAGATCTACAC, ACACTCTTTCCCTACACGACG, CAAGCAGAAGACGGCATACGAGAT and GTGACTGGAGTTCAGACGTGT, respectively.

4.4. Automated Genotyping in SSR-GBS and Repeatability

In previous studies of SSR-GBS, allele calling was performed based on whether the frequency of the first allele exceeded 0.9 or on manual evaluation [27]. However, each marker has different features such as stutter bands and base variations in the primer binding site, making it difficult to set a single value for all marker sets. Here we set the
variable $\alpha$, which was determined separately for each marker. If the frequency of the first (most common) allele exceeded $\alpha$, we called the genotype a homozygote for the first allele; otherwise, we called it a heterozygote for the first and second alleles (described in the script '4.genotyping.sh' provided in File S4). The range of $\alpha$ for each marker was determined based on the criterion that the genotypes identified by SSR-GBS and PCR-CE were exactly the same for each of the varieties. In other words, the $\alpha$ values were set so that the genotyping results obtained by SSR-GBS in an automated system would match those previously obtained using PCR-CE, in which genotyping was done by eye. As an example of file structure, the files used to designate the apple cultivars and primers in this pipeline are provided in Files S5 and S6, respectively.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/plants10122796/s1, Table S1: SSR markers used in the present study, Table S2: Genotypes determined by PCR-CE for 43 apple varieties, Table S3: Genotypes determined by PCR-CE for 29 pear varieties, Table S4: Genotypes determined by PCR-CE for 44 tea varieties, Table S5: Genotypes determined by SSR-GBS for 43 apple varieties, Table S6: Genotypes determined by SSR-GBS for 29 pear varieties, Table S7: Genotypes determined by SSR-GBS for 44 tea varieties, Figure S1: UPGMA dendrograms showing clustering based on simple allele sharing distance among (A) 44 apple, (B) 29 pear and (C) 44 tea varieties. The figures in red indicate an approximately unbiased (AU) $p$ value that represents more accurate values of bootstrap probability. File S1: 1.combine.sh. Bash script for combining the pairs of fastq files, File S2: 2.demultiplex.sh. Bash script for demultiplexing based on primer sequences, File S3: 3.CountSSR.sh. Bash script for counting alleles detected in each variety, File S4: 4.genotyping.sh. Bash script for determining the genotypes automatically based on variable $\alpha$, File S5: list_geno. List of apple varieties or IDs used in this study, File S6: list_primer. List of apple primers and variable $\alpha$ used in this study.

Author Contributions: Conceptualization, S.N. and Y.S.; methodology, S.N. and H.K.-K.; software, S.N. and H.K.-K.; validation, S.N., Y.T. and S.M.; formal analysis, F.T. and M.K.; resources, S.M., Y.S., M.K. and F.T.; data curation, S.M., M.K. and F.T.; writing—original draft preparation, S.N.; writing—review and editing, S.N.; visualization, S.N.; project administration, Y.S.; funding acquisition, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The genotype data and the scripts used in this study are available in Supplementary files.

Acknowledgments: We thank Fumiko Hosaka for the technical support in genotyping and organizing experimental data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Luby, J.J.; Shaw, D.V. Does marker-assisted selection make dollars and sense in a fruit breeding program? HortScience 2001, 36, 872–879. [CrossRef]
2. Janick, J. History of Fruit Breeding. Fruit Veg. Cereal Sci. Biotechnol. 2011, 5, 1–7.
3. Saito, T. Advances in Japanese pear breeding in Japan. Breed. Sci. 2016, 66, 46–59. [CrossRef] [PubMed]
4. Yamada, M.; Sato, A. Advances in table grape breeding in Japan. Breed. Sci. 2016, 66, 34–45. [CrossRef]
5. Endo, T.; Fujii, H.; Yoshioka, T.; Omura, M.; Shimada, T. TaqMan-MGB SNP genotyping assay to identify 48 citrus cultivars distributed in the Japanese market. Breed. Sci. 2020, 70, 363–372. [CrossRef]
6. Scribner, K.T.; Pearce, J.M. Microsatellites: Evolutionary and methodological background and empirical applications at individual, population, and phylogenetic levels. In Molecular Methods in Ecology; Blackwell Science Limited: London, UK, 2000; pp. 235–273.
7. Sunnucks, P. Efficient genetic markers for population biology. Trends Ecol. Evol. 2000, 15, 199–203. [CrossRef]
8. Ellegren, H. Microsatellites: Simple sequences with complex evolution. Nat. Rev. Genet. 2004, 5, 435–445. [CrossRef] [PubMed]
9. Selkoe, K.A.; Toonen, R.J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. Ecol. Lett. 2006, 9, 615–629. [CrossRef] [PubMed]
10. Clayton, T.M.; Whitaker, J.P.; Sparkes, R.; Gill, P. Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Sci. Int.* **1998**, *91*, 55–70. [CrossRef]

11. Cipriani, G.; Spadotto, A.; Juran, I.; Di Gaspero, G.; Crespaz, M.; Meneghetti, S.; Frare, E.; Vignani, R.; Cresti, M.; Morgante, M. The SSR-based molecular profile of 1005 grapevines (*Vitis vinifera L.*) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor. Appl. Genet.* **2010**, *121*, 1569–1583. [CrossRef]

12. Moriya, S.; Iwanami, H.; Okada, K.; Yamamoto, T.; Abe, K. A practical method for apple cultivar identification and parent-offspring analysis using simple sequence repeat markers. *Euphytica* **2011**, *177*, 135–150. [CrossRef]

13. Sawamura, Y.; Takada, N.; Yamamoto, T.; Saito, T.; Kimura, T.; Kotobuki, K. Identification of parent-offspring relationships in 55 Japanese pear cultivars using S-RNase allele and SSR markers. *J. Jpn. Soc. Hortic. Sci.* **2008**, *77*, 364–373. [CrossRef]

14. Kimura, T.; Shi, Y.Z.; Shoda, M.; Kotobuki, K.; Matsuura, N.; Hayashi, T.; Ban, Y.; Yamamoto, T. Identification of Asian pear varieties by SSR analysis. *Breed. Sci.* **2002**, *52*, 115–121. [CrossRef]

15. Leigh, F.; Lea, V.; Law, J.; Wolters, P.; Powell, W.; Donini, P. Assessment of EST-and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. *Euphytica* **2003**, *133*, 359–366. [CrossRef]

16. Hosseinizadeh-Colagar, A.; Haghhighatnia, M.J.; Amiri, Z.; Mohadjerani, M.; Tafrihi, M. Microsatellite (SSR) amplification by PCR. *Plants* **2021**, *10*, 532. [CrossRef]

17. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS ONE* **2011**, *6*, e19379. [CrossRef] [PubMed]

18. 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*, 145–153. [CrossRef] [PubMed]

19. Du, F.K.; Xu, F.; Qu, H.; Feng, S.; Tang, J.; Wu, R. Exploiting the transcriptome of Euphrates Poplar, *Populus euphratica* (Salicaceae) to develop and characterize new EST-SSR markers and construct an EST-SSR database. *PloS ONE* **2013**, *8*, e61337. [CrossRef]

20. Dubey, H.; Rawal, H.C.; Rohlina, M.; Luma, U.; Kumar, P.M.; Bandyopadhyay, T.; Gogoi, M.; Singh, N.K.; Mondal, T.K. TeaMiD: A comprehensive database of simple sequence repeat markers of tea. *Database 2020*, 2020. [CrossRef] [PubMed]

21. Nishio, S.; Takada, N.; Terakami, S.; Kato, H.; Inoue, H.; Takeuchi, Y.; Saito, T. Estimation of effective pollen dispersal distance for cross-pollination in chestnut orchards by microsatellite-based paternity analyses. *Sci. Hort. (Amsterdam)* **2019**, *250*, 89–93. [CrossRef]

22. Silfverberg-Dilworth, E.; Matasci, C.L.; Van de Weg, W.E.; Van Kauwen, M.P.W.; Walser, M.; Koudde, L.P.; Soglio, V.; Gianfranceschi, L.; Durel, C.E.; Costa, F. Microsatellite markers spanning the apple (*Malus x domestica* Borkh.) genome. *Tree Genet. Genomes* **2006**, *2*, 202–224. [CrossRef]

23. Matsuki, Y.; Isagi, Y.; Suyama, Y. Identification of multiple microsatellite genotypes and DNA sequences from a single pollen grain. *Mol. Ecol. Notes* **2007**, *7*, 194–198. [CrossRef]

24. Suyama, Y. Procedure for single-pollen genotyping. In *Single-Pollen Genotyping*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 7–15.

25. Tihihik, P.D.; Curto, M.; Dornstauder-Schrammel, E.; Winter, S.; Alemayehu, E.; Waidbacher, R.; Meimberg, H. Identification of microsatellite genotyping by sequencing (SSR-GBS) to measure genetic diversity of the East African *Oreochromis niloticus*. *Conserv. Genet.* **2019**, *20*, 357–372. [CrossRef]

26. Šarhanová, P.; Pfanzelt, S.; Brandl, R.; Himmelbach, A.; Blattner, F.R. SSR-seq: Genotyping of microsatellites using next-generation sequencing reveals higher level of polymorphism as compared to traditional fragment size scoring. *Ecol. Evol.* **2018**, *8*, 10817–10833. [CrossRef]

27. Curto, M.; Winter, S.; Seiter, A.; Schmid, L.; Scheicher, K.; Barthel, L.M.F.; Plass, J.; Meimberg, H. A SSR-GBS marker system on investigation of European Hedgehog species and their hybrid zone dynamics. *Ecol. Evol.* **2019**, *9*, 2814–2832. [CrossRef] [PubMed]

28. Zhan, L.; Paterson, I.G.; Fraser, B.A.; Watson, B.; Bradbury, I.R.; Nadukkalam Ravindran, P.; Reznick, D.; Beiko, R.G.; Bentzen, P. MEGASAT: Automated inference of microsatellite genotypes from sequence data. *Mol. Ecol. Resour.* **2017**, *17*, 247–256. [CrossRef]

29. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Keeler, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 1–9. [CrossRef] [PubMed]

30. Daccord, N.; Celton, J.M.; Linsmith, G.; Becker, C.; Choines, N.; Schijlen, E.; Van de Geest, H.; Bianco, L.; Michelelli, D.; Velasco, R. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.* **2017**, *49*, 1099–1106. [CrossRef]

31. Shirasawa, K.; Itai, A.; Isobe, S. Chromosome-scale genome assembly of Japanese pear (*Pyrus pyrifolia*) variety ‘Nijisseiki.’ *DNA Res.* **2021**, *28*. [CrossRef]

32. Wei, C.; Yang, H.; Wang, S.; Zhao, J.; Liu, C.; Gao, L.; Xia, E.; Lu, Y.; Tai, Y.; She, G. Draft genome sequence of Camellia sinensis var. sinensis provides insights into the evolution of the tea genome and tea quality. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4151–E4158. [CrossRef] [PubMed]

33. Darby, B.J.; Erickson, S.F.; Hervey, S.D.; Ellis-Felege, S.N. Digital fragment analysis of short tandem repeats by high-throughput amplicon sequencing. *Ecol. Evol.* **2016**, *6*, 4502–4512. [CrossRef] [PubMed]

34. Neophytou, C.; Torutaeva, E.; Winter, S.; Meimberg, H.; Hasenauer, H.; Curto, M. Analysis of microsatellite loci in tree of heaven (*Ailanthus altissima* (Mill.) Swingle) using SSR-GBS. *Tree Genet. Genomes* **2018**, *14*, 82. [CrossRef]
35. Lepais, O.; Chancerel, E.; Boury, C.; Salin, F.; Manicki, A.; Taillebois, L.; Dutech, C.; Assisi, A.; Bacles, C.F.E.; Daverat, F. Fast sequence-based microsatellite genotyping development workflow. _PeerJ_ 2020, 8, e9085. [CrossRef]

36. Bornman, D.M.; Hester, M.E.; Schuetter, J.M.; Kasoji, M.D.; Minard-Smith, A.; Barden, C.A.; Nelson, S.C.; Godbold, G.D.; Baker, C.H.; Yang, B.; et al. Short-read, high-throughput sequencing technology for STR genotyping. _Biotech. Rapid Dispatches_ 2012, 2012, 1–6. [CrossRef]

37. Suzuki, R.; Shimodaira, H. Pvclust: An R package for assessing the uncertainty in hierarchical clustering. _Bioinformatics_ 2006, 22, 1540–1542. [CrossRef] [PubMed]

38. Celton, J.-M.; Tustin, D.S.; Chagné, D.; Gardiner, S.E. Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from Malus ESTs and _Pyrus_ genomic sequences. _Tree Genet. Genomes_ 2009, 5, 93–107. [CrossRef]

39. Kunihisa, M.; Moriya, S.; Abe, K.; Okada, K.; Haji, T.; Hayashi, T.; Kim, H.; Nishitani, C.; Terakami, S.; Yamamoto, T. Identification of QTLs for fruit quality traits in Japanese apples: QTLs for early ripening are tightly related to preharvest fruit drop. _Breed. Sci._ 2014, 64, 240–251. [CrossRef]

40. Liebhard, R.; Gianfranceschi, L.; Koller, B.; Ryder, C.D.; Tarchini, R.; Van de Weg, E.; Gessler, C. Development and characterisation of 140 new microsatellites in apple ( _Malus x domestica_ Borkh.). _Mol. Breed._ 2002, 10, 217–241. [CrossRef]

41. Van Dyk, M.M.; Soeker, M.K.; Labuschagne, I.F.; Rees, D.J.G. Identification of a major QTL for time of initial vegetative budbreak in apple ( _Malus x domestica_ Borkh.). _Tree Genet. Genomes_ 2010, 6, 489–502. [CrossRef]

42. Yamamoto, T.; Terakami, S.; Hosaka, F.; Fujii, H.; Kunihisa, M.; Takada, N.; Saito, T.; Nishitani, C. DNA profiling of Japanese pear cultivars by SSR markers with tetra- and penta-nucleotide motifs. _DNA Polymorph._ 2021, 20, 58–61. (In Japanese)

43. Ma, J.; Zhou, Y.; Ma, C.; Yao, M.; Jin, J.; Wang, X.; Chen, L. Identification and characterization of 74 novel polymorphic EST-SSR markers in the tea plant, _Camellia sinensis_ (Theaceae). _Am. J. Bot._ 2010, 97, e153–e156. [CrossRef]

44. Tan, L.Q.; Wang, L.Y.; Wei, K.; Zhang, C.C.; Wu, L.Y.; Qi, G.N.; Cheng, H.; Zhang, Q.; Cui, Q.M.; Liang, J.B. Floral transcriptome sequencing for SSR marker development and linkage map construction in the tea plant ( _Camellia sinensis_). _PLoS ONE_ 2013, 8, e81611. [CrossRef]

45. Magoˇc, T.; Salzberg, S.L. FLASH: Fast length adjustment of short reads to improve genome assemblies. _Bioinformatics_ 2011, 27, 2957–2963. [CrossRef]