

Article

Expression Profiles of Microsatellites in Fruit Tissues of *Akebia trifoliata* and Development of Efficient EST-SSR Markers

Wei Chen ^{1,†}, Huai Yang ^{1,†}, Shengfu Zhong ¹, Jun Zhu ¹, Qiuyi Zhang ¹, Zhi Li ¹, Tianheng Ren ¹ , Feiquan Tan ¹, Jinliang Shen ², Qing Li ³ and Peigao Luo ^{1,*} 

¹ State Key Laboratory of Plant Breeding and Genetics, Sichuan Agricultural University, Chengdu 611130, China

² College of Forestry, Sichuan Agricultural University, Chengdu 611130, China

³ Department of Biology and Chemistry, Chongqing Industry and Trade Polytechnic, Chongqing 408000, China

* Correspondence: lpglab@sicau.edu.cn

† These authors contributed equally to this work.

Abstract: *Akebia trifoliata*, a member of the family Lardizabalaceae, has high exploitation potential for multiple economic purposes, so genetic improvements to meet requirements for commercial demand are needed. However, this progress is largely impeded by a lack of effective selection markers. In this study, we obtained 271.49 Gb of clean transcriptomic data from 12 samples (three tissues at four developmental stages) of *A. trifoliata* fruit. We identified 175,604, 194,370, and 207,906 SSRs from the de novo assembled 416,363, 463,756, and 491,680 unigene sequences obtained from the flesh, seed, and rind tissues, respectively. The profile and proportion of SSR motifs expressed in each fruit tissue and developmental stage were remarkably similar, but many trinucleotide repeats had differential expression levels among different tissues or at different developmental stages. In addition, we successfully designed 16,869 functional EST-SSR primers according to the annotated unigenes. Finally, 94 and 72 primer pairs out of 100 randomly selected primer pairs produced clear bands and polymorphic bands, respectively. These results were also used to elucidate the expression profiles of different tissues at various stages. Additionally, we provided a set of effective, polymorphic, and reliable EST-SSR markers sufficient for accelerating the discovery of metabolic and pathway-specific functional genes for genetic improvement and increased commercial productivity.

Keywords: *Akebia trifoliata*; transcriptome; EST-SSR markers; expression profile; validation test



Citation: Chen, W.; Yang, H.; Zhong, S.; Zhu, J.; Zhang, Q.; Li, Z.; Ren, T.; Tan, F.; Shen, J.; Li, Q.; et al. Expression Profiles of Microsatellites in Fruit Tissues of *Akebia trifoliata* and Development of Efficient EST-SSR Markers. *Genes* **2022**, *13*, 1451. <https://doi.org/10.3390/genes13081451>

Academic Editor: Jacqueline Batley

Received: 26 July 2022

Accepted: 11 August 2022

Published: 15 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Akebia trifoliata (Thunb.) Koidz, commonly called augmelon in China, is a domesticated representative of the Lardizabalaceae family used in crop cultivation. The growing area throughout East Asia, and especially in China, is rapidly increasing [1]. *A. trifoliata* can be exploited for multiple purposes, such as traditional Chinese medicine [2], delicious fruit [2,3], edible oil [4], and ornamental horticulture [5]. All commercial exploitation centers on the fruit, although other parts, such as stems and roots, have been used in various fields, such as cosmetics and other industries [5].

For example, the various beautiful shapes and attractive colors of whole *A. trifoliata* fruit possess aesthetic value, and *A. trifoliata* can be used as an ornamental plant [1]. However, the plant also produces many secondary metabolites, such as saponins and flavones, in the whole immature fruit and the rind of mature fruit, which have made it a traditional medicinal plant in East Asia, including China, Korea, and Japan, for at least 2000 years [2]. The flesh of mature fruit is delicious and contains many beneficial nutrients, such as sugars [6], free amino acids [3], vitamins [3], and minerals [7], which gradually make it a new fruit for consumption and health purposes [2,8]. Reports have also suggested that the mature fruit has many seeds that have a high oil content (usually above 45%), and the oil is rich [4] in saturated, monounsaturated, and polyunsaturated

fatty acids, with a proportion close to 1:1:1 [9], which matches the composition of ideal oils recommended by an increasing number of authors [10] for nutritional and health purposes. Therefore, *A. trifoliata* also has enormous potential as a new edible oil crop. The economic value of various tissues of the *A. trifoliata* fruit has received much attention, and the rate of corresponding commercial demand and exploitation is increasing. Unfortunately, there has been little work on improving the genetics controlling fruit traits of commercial interest for increased production.

Various molecular markers, such as restriction fragment length polymorphisms (RFLPs), sequence-tagged sites (STSs), single-sequence repeats (SSRs), and single-nucleotide polymorphisms (SNPs), have been developed as powerful selection tools in many cereal crops to assist breeding [11,12]. In recent years, the application of molecular markers in horticultural crops has further promoted the identification of functional genes and population genetic studies in plants, such as *Paeonia* [13], *Rosa* [14], and melon [15], which has greatly accelerated the progress of genetic improvement and has largely enhanced the effectiveness of selection on new commercial cultivars. Among them, SSRs widely exist in intergenic and genic regions of both prokaryotic and eukaryotic organism genomes [16]. SSRs are usually abundant, codominant, reproducible, highly polymorphic, and genomewide, so they have commonly been used as powerful molecular tools in gene mapping, DNA fingerprinting, genotype or varietal identification, molecular marker-assisted breeding, and population diversity analysis [17–19]. The number of SSRs in intergenic regions is generally greater than that in genic regions in a given eukaryotic species, although the density of SSRs in some genic regions could be higher than the average SSR density in intergenic regions because genic regions compose only a small proportion of the entire genome compared with intergenic regions [20,21]. However, SSR markers in genic regions are more effective than those in intergenic regions for genetic improvement programs and the enhancement of economically interesting crop traits, especially in studies exploiting wild species used to improve cultivars because they were designed on the basis of gene-self sequences rather than on flanking gene sequences.

In addition, the original sequences used to detect SSRs can be divided into genomic SSRs and EST-SSRs (expressed sequence tag-SSRs) [22]. Because EST-SSR markers are derived from expressed sequences and located within genic regions, their polymorphism can be directly coupled with functional differences in the corresponding genes [23]. Practically, the EST-SSR markers derived from some genes involved in a specific biological pathway or metabolic process would play a larger role in molecular-assisted selection for a given objective trait influenced by that pathway than genome-wide markers [24]. For example, EST-SSR markers were found to be involved in disease resistance and pigment biosynthesis processes in *Triticum aestivum* and *Prunus salicina* [25,26]. In addition, the development of EST-SSR markers is relatively less time-consuming and less expensive than that of genomic SSR markers, especially for plant species without available reference genomes. Therefore, EST-SSR markers are more likely to be associated with specific phenotypic traits and are more valuable, more effective, and less expensive for molecular breeding and functional gene mining when exploiting wild plants for the genetic improvement of cultivars.

At present, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis are two especially useful and powerful methods to represent the detailed information of the genes or biological protein functions in well-known signal or metabolism pathways or processes in model species [27,28]. In addition, with the development and progress of DNA and RNA sequencing technology, the fast acquisition of transcriptome data from different experimental treatments and different developmental stages of various organs can be made available to scientists worldwide. It is evident that effectively focusing on a few useful genes and corresponding proteins involved in the signal and metabolism pathways related to important agronomic and economic traits via both GO and KEGG analysis is a powerful strategy to further develop expression sequence-based molecular markers, such as EST-SSR markers [29,30].

In this study, we acquired transcriptome data for three tissues, including the rind, flesh, and seed, of *A. trifoliata* fruit at four different developmental stages using Illumina HiSeq technology. These data were then employed to isolate and identify candidate unigenes containing SSRs. Both motif and expression profiles were outlined, and expression differences in trinucleotide repeats (TNRs) in different tissues at different stages were highlighted. Our objective was to develop and validate a set of highly effective EST-SSR markers, which could be valuable for improving traits of interest in *A. trifoliata* for specific commercial or economic purposes as well as for further exploitation of other wild plants.

2. Materials and Methods

2.1. Plant Materials and Sequencing

The clonal lines of “Shusen 1” were used to produce transcriptome data by Illumina sequencing because the available reference genome for *A. trifoliata* was also produced from this strain [31]. All clonal materials were cultivated in the *A. trifoliata* Germplasm Nursery of Sichuan Agricultural University at the Chongzhou Research Station (30°43' N–103°65' E) with a randomized complete block experimental design. Three healthy tissues, including the flesh, peel, and seed, of fruits at four distinct stages, including young, enlargement, coloring, and maturity, using the previously classified standard [1] were randomly collected from 2-year-old plants for transcriptome sequencing, and there were three biological replicates for each treatment. To construct sequencing libraries, total RNA from all 36 samples was extracted using TRIzol reagent (Invitrogen, Waltham, MA, USA) according to the manufacturer’s instructions. The paired-end libraries with 240 bp inserts were prepared and then sequenced on an Illumina HiSeq 2500.

In addition, a total of 45 *A. trifoliata* accessions were selected from germplasm nurseries according to geographic origin, distribution density, and phenotypic character to verify the polymorphism of EST-SSR markers; these strains included 15 selected super strains [3], 15 collected accessions from the Sichuan Basin as the putative origin of *A. trifoliata*, and 15 accessions from a large geographic scope of the natural distribution of *A. trifoliata*.

2.2. De Novo Assembly and SSR Identification

Raw transcriptome sequences were filtered by Trimmomatic (version 0.30, Jülich, Germany) software according to a previously described method [32]. Next, all high-quality reads were de novo assembled using Trinity (version 2.4.0, San Diego, CA, USA) tools [33], and the de novo assembled unigenes were further used to detect SSRs with the MicroSatellite (MISA) protocol [34]. To ensure the high quality of EST-SSR markers, the default parameters were set as follows: the minimum repeat numbers for single-nucleotide repeats (SNRs), dinucleotide repeats (DNRs), and other repeats, including trinucleotide repeats (TNRs), tetranucleotide repeats (TtNRs), pentanucleotide repeats (PNRs), and hexanucleotide repeats (HNRs), were 10, 6, and 5, respectively. In addition, the criterion for compound SSRs was defined if the interval between two SSR sequences was less than 100 bp.

2.3. Statistical Analysis of TNR Motifs

ANOVA (p -value = 0.05) was performed on the number of TNRs among the three tissues and four developmental stages of *A. trifoliata* fruit using SPSS 20 software. The chi-square test was also used to test whether the number of complementary TNRs of two strands fit a 1:1 ratio.

2.4. Functional Annotation of Unigenes

To identify the potential functions of unigenes carrying SSR loci, the sequences of unigenes were mapped to the GO and KEGG databases through the eggNOG-mapper (version 2.0.0, Heidelberg, Germany) tool (E -value $< 1 \times 10^{-5}$) [35]. The GO database classifies the functional annotations of these unigenes into three categories: molecular function, biological process, and cellular component. The annotation of the KEGG database matched

unigene functions to various biological metabolic pathways, including carbohydrate, lipid, amino acid, and secondary metabolite metabolism pathways. In addition, TBtools was used to assist in the analysis of functional annotations of unigenes [36].

2.5. EST-SSR Primer Design

Unigenes carrying SSRs and annotated to metabolic pathways were consequently used as templates to design EST-SSR markers according to the 100 bp sequences flanking the SSR loci using Primer3 (version 2.4.0, San Diego, CA, USA), as previously reported [37]. The software was run with the following parameters: primer length, melting temperature, GC content, and amplification product length ranging from 18 to 23 bp, 50° to 65°, 40% to 60%, and 100 to 300 bp. Only one primer pair was retained for each SSR locus in each unigene.

2.6. DNA Extraction and Verification of EST-SSR Markers

Young leaves of 45 accessions of *A. trifoliata* were collected for genomic DNA extraction according to the previously described CTAB protocol [38]. To verify the effectiveness and polymorphism of the primers, 100 pairs of primers for DNRs or TNRs were randomly selected and synthesized (Table S6). They were then amplified by PCR on a PTC-200 thermocycler (MJ Research, Watertown, MA, USA) with a total reaction volume of 25 µL, containing 1.5 mmol/L MgCl₂, 0.2 mmol/L dNTPs, 1 unit of Taq DNA polymerase, 200 nmol/L of each primer, and 80 ng of template DNA. The PCR conditions were set as follows: 94 °C for 2 min; 32 cycles of 94 °C for 45 s, 55–60 °C for 30 s, and 72 °C for 30 s; and 72 °C for 10 min. PCR products were subsequently separated and visualized on a 6% nondenaturing polyacrylamide gel.

3. Results

3.1. De Novo Assembly of Transcriptome Sequencing Data and Identification of SSRs

All RNAs were sequenced by Illumina sequencing technology, and the raw transcriptome data (271.49 Gb of clean data) were deposited in the National Centre for Biotechnology Information (NCBI) with accession numbers SAMN16551931–33, SAMN16551934–36, SAMN16551937–39, and SAMN16551940–42 for the young, enlargement, coloring, and mature stages, respectively, and further de novo assembled. A total of 416,363 (323.45 Mb), 463,756 (350.58 Mb), and 491,680 (377.16 Mb) nonredundant genes (unigenes) were obtained and carried SSR loci from the flesh, seed, and rind tissues, respectively; we found 119,394, 132,582, and 140,247 unigenes in the flesh, seed, and rind tissues, respectively. Further analysis showed that 175,604, 194,370, and 207,906 SSRs with 19,844, 22,367, and 23,888 compound forms were detected in the flesh, seed, and rind tissues, respectively. The SSRs in the expressed sequences of the fruit tissue of *A. trifoliata* were abundant, and the frequency of occurrence was 549.53 per Mb. The assembled unigenes from different fruit tissues had similar GC contents and average lengths (Table 1). In addition, the proportion of unigenes carrying SSRs was also remarkably similar among various tissues (Table 1).

Table 1. De novo assembly results and SSR identification in *A. trifoliata* fruit tissues.

Parameter	Sum (Flesh)	Sum (Seed)	Sum (Rind)
Total number of sequences examined:	416,363	463,756	491,680
Total size of examined sequences (bp):	323,449,301	350,577,642	377,164,182
Total number of identified SSRs:	175,604	194,370	207,906
Number of sequences containing SSRs:	119,394	132,582	140,247
Number of sequences containing more than 1 SSR:	37,747	41,969	45,387
Number of SSRs present in compound formation:	19,844	22,367	23,888
Percent GC (%):	39.37%	39.29%	39.31%
The average length of the sequence (bp):	777.08	756.47	769.8
Mean distribution distance (SSR/Mb):	542.91	554.43	551.24
Proportion of unigenes with SSRs:	28.68%	28.59%	28.52%

3.2. Frequency and Distribution of SSR Motifs in Three Fruit Tissues

Both the motif profile and proportion of SSRs were extremely similar in the three *A. trifoliata* fruit tissues, and the number of SSRs sharply declined as the repeat unit length increased, among which short motifs, including SNRs, DNRs, and TNRs, accounted for more than 98% of the total SSRs, while long motifs, including TtNRs, PNRs, and HNRs, composed a very small proportion (<1.5%) (Figure 1a; Table S1). Further comparison showed that SNRs had the largest proportion with an average of 60.63%, followed by DNRs with an average of 28.39% and TNRs with an average of 9.69% in the three fruit tissues. In addition, different repeat motifs exhibited various proportions within the same repeat unit length. A/T was the most common repeat type of SNR, accounting for an average of 95.57% of SNRs; AG/CT and AT/AT composed an average of 64.17% and 21.50% of DNRs, respectively; among TNRs, the proportions of AAG/CTT, AAT/ATT, and ATC/ATG motifs were 26.03%, 22.26%, and 15.29%, respectively. Finally, we also found that the difference in both the composition and the proportion of repeat motifs in various tissues at distinct stages was small. For example, in 12 samples (three tissues and four developmental stages), the proportion of A/T in SNRs ranged from 95.18% to 95.95% (Figure 1b); the proportions of AG/CT and AT/AT in DNRs ranged from 61.90% to 67.27% and 19.28% to 23.22%, respectively; and those of AAG/CTT, AAT/ATT, and ATC/ATG in TNRs ranged from 24.33% to 27.28%, 20.65% to 23.86%, and 14.53% to 15.82%, respectively (Table S1).

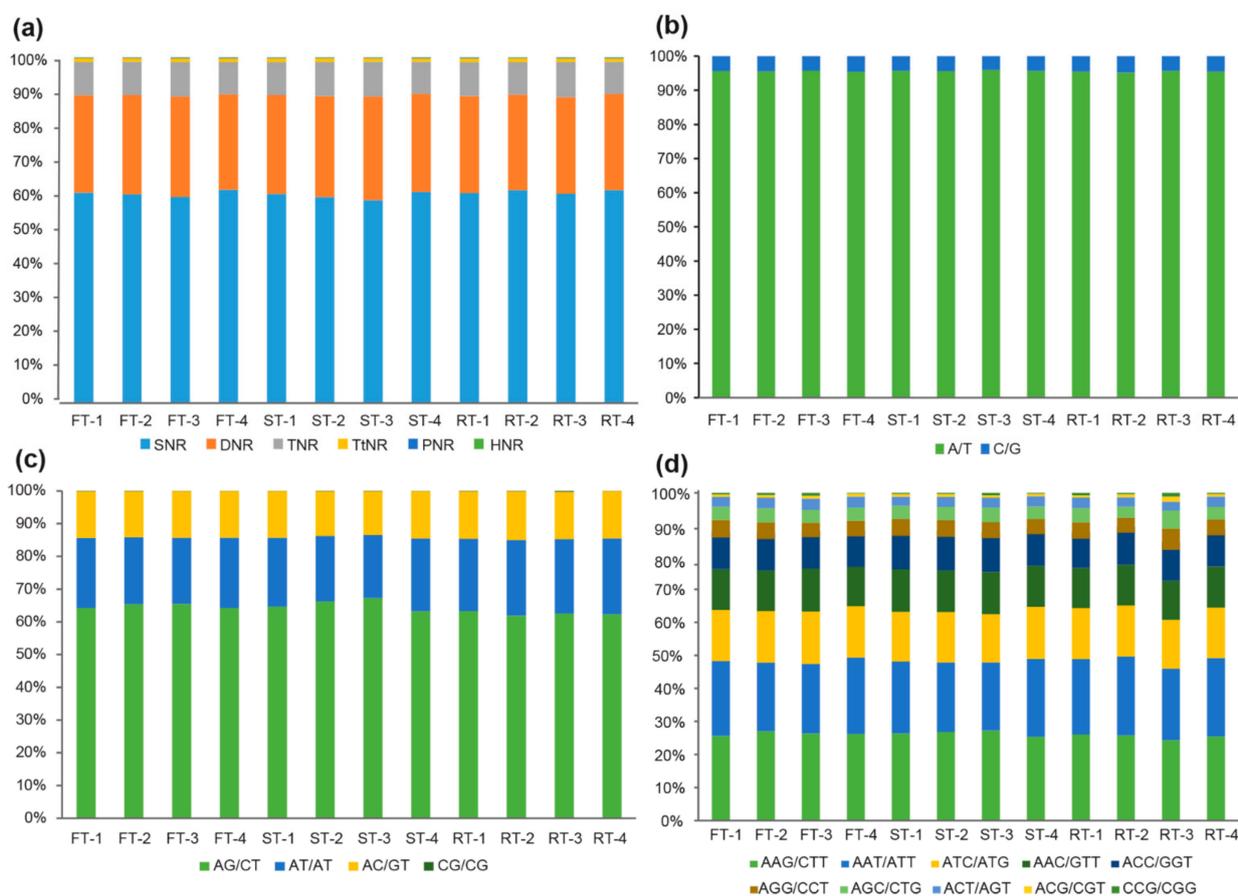


Figure 1. Frequency and distribution of SSR motifs in three fruit tissues of *A. trifoliata* (FT: flesh, ST: seed, RT: rind, -1: young stage, -2: enlargement stage, -3: coloring stage, -4: mature stage). (a) The proportion of different repeat unit lengths, (b) the proportion of different repeat motifs within SNRs, (c) the proportion of different repeat motifs within DNRs, and (d) the proportion of different repeat motifs within TNRs.

3.3. Expression Differences in TNRs in Different Fruit Tissues and Different Developmental Stages

TNRs detected in expressed sequences are a class of microsatellites that may directly alter gene coding and are compacted in coding regions due to the nature of translation and the dependence of translation on triplet codons [39]. In this study, an analysis of variance on 60 TNRs (excluding AAA, TTT, CCC, and GGG) expressed in *A. trifoliata* fruit tissues showed that 18 (30%) TNRs were differentially expressed in various tissues and developmental stages (Table 2). Among them, 15 (83.33%) TNRs had differential expression among different tissues, and the motifs putatively encoded nine different amino acids, including phenylalanine (TTC), leucine (CTT; TTA), tryptophan (TGG), valine (GTT), histidine (CAC; CAT), glutamine (CAA), proline (CCT), serine (AGT, TCA, TCC, TCT), and asparagine (AAC, AAT), while the other three (16.67%) TNRs, ACA, GCC, and GTA, putatively encoded threonine (ACA), valine (GTA), and alanine (GCC), respectively, and exhibited differential expression at various developmental stages (Table 2).

Table 2. Analysis of variance for TNRs in different fruit tissues and developmental stages.

Motif	Amino Acid	Three Fruit Tissues					Developmental Period				
		Sum of Squares	Degree of Freedom	Mean of Squares	F-Value	p-Value	Sum of Squares	Degree of Freedom	Mean of Squares	F-Value	p-Value
AAC	Asn	554.67	2	277.33	6.84	0.028	124.67	3	41.56	1.03	0.446
AAT	Asn	19,478.17	2	9739.08	6.49	0.032	207.58	3	69.19	0.05	0.986
AGT	Ser	148.17	2	74.08	17.21	0.003	45.67	3	15.22	3.54	0.088
CAA	Gln	1083.17	2	541.58	6.87	0.028	385.67	3	128.56	1.63	0.279
CAC	His	1383.5	2	691.75	7.11	0.026	222.92	3	74.31	0.76	0.554
CAT	His	914.67	2	457.33	6.74	0.029	150.92	3	50.31	0.74	0.565
CCT	Pro	184.67	2	92.33	9.03	0.015	29.67	3	9.89	0.97	0.467
CTT	Leu	578.17	2	289.08	5.55	0.043	158.25	3	52.75	1.01	0.45
GTT	Val	501.17	2	250.58	8.28	0.019	82	3	27.33	0.9	0.493
TCA	Ser	2931.17	2	1465.58	12.52	0.007	39.58	3	13.19	0.11	0.949
TCC	Ser	115.17	2	57.58	9.21	0.015	8.25	3	2.75	0.44	0.733
TCT	Ser	3095.17	2	1547.58	8.4	0.018	885.67	3	295.22	1.6	0.285
TGG	Trp	580.67	2	290.33	10.58	0.011	216.33	3	72.11	2.63	0.145
TTA	Leu	5202	2	2601	8.15	0.019	1187.58	3	395.86	1.24	0.375
TTC	Phe	5552.67	2	2776.33	11.8	0.008	174.92	3	58.31	0.25	0.86
ACA	Thr	132.17	2	66.08	1.85	0.236	689.67	3	229.89	6.45	0.026
GCC	Ala	0.17	2	0.08	0.03	0.969	50.92	3	16.97	6.43	0.026
GTA	Val	2	2	1	0.14	0.872	107.33	3	35.78	5.03	0.045

Chi-square tests for two corresponding TNRs on complementary strands showed that among 30 potential combinations of the 60 TNRs, only two combinations (ATA/TAT and CCT/GGA) were expressed with the expected 1:1 ratio in all samples. In contrast, 11 combinations significantly deviated from the ratio in all samples, and 8 combinations also significantly deviated from the ratio in some samples at the $p = 0.05$ level. Statistical tests of the remaining 9 combinations were not executed because the expression level was too low (Figure 2; Table S3). This suggested that the expression preference of complementary motifs varied widely in specific tissues and developmental stages. For instance, the complementary motifs AGA (Arg) and TCT (Ser) had significant expression preferences in both rind and seed tissues at all developmental stages, but the preference in flesh only existed at some developmental stages (Figure 2). The expression preference of the complementary motifs ACA (Thr) and TGT (Cys) existed in all tissues at young stages, but it disappeared at later stages of development (Figure 2).

3.4. Functional Annotation of Unigenes Containing SSRs

To clarify the function of the unigenes carrying SSRs in the fruit tissues of *A. trifoliata*, a total of 392,223 unigene sequences consisting of 119,394 from flesh, 132,582 from seeds, and 140,247 from rind were BLAST searched in the GO and KEGG databases. A total of 91,238 unigenes containing SSRs were annotated to 529,711 different terms by Gene Ontology (GO) analysis, and they were grouped into three large categories: cellular component (97,510, 18.41%), molecular function (99,115, 18.71%), and biological process (333,086, 62.88%) with 3, 18, and 28 subcategories, respectively (Figure 3a). Among the three types of cellular component categories, the GO terms involved in the cellular anatomical entity type were the

most enriched (80,340, 82.39%), followed by the protein-containing complex in the cellular component category (17,160, 17.60%) (Figure 3a). Likewise, among 18 types of molecular function categories, the top 3 types were catalytic activity (39,000, 39.35%), binding (32,636, 32.93%), and transcription regulator activity (8871, 8.95%), and among 28 types of biological process categories, cellular process (71,572, 21.49%), metabolic process (58,038, 17.42%), and biological regulation (36,799, 11.05%) were the three most abundant types (Figure 3a).

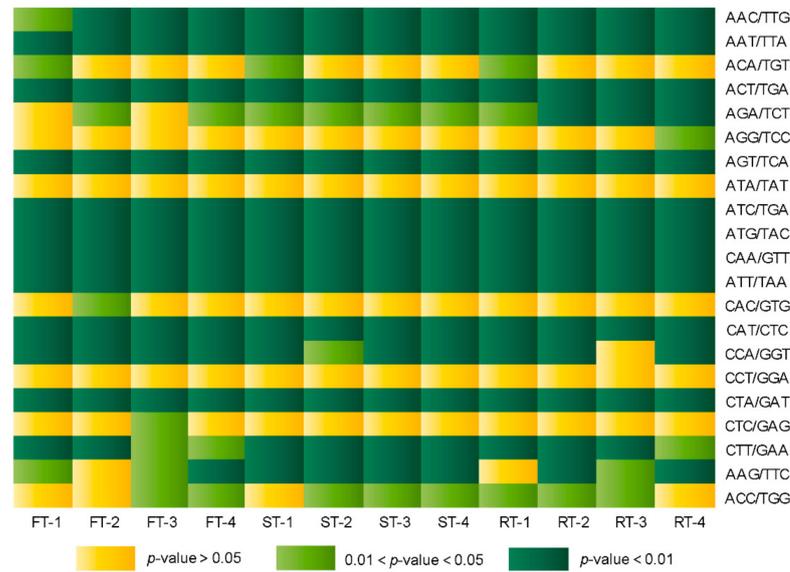


Figure 2. Chi-square test of complementary TNRs.

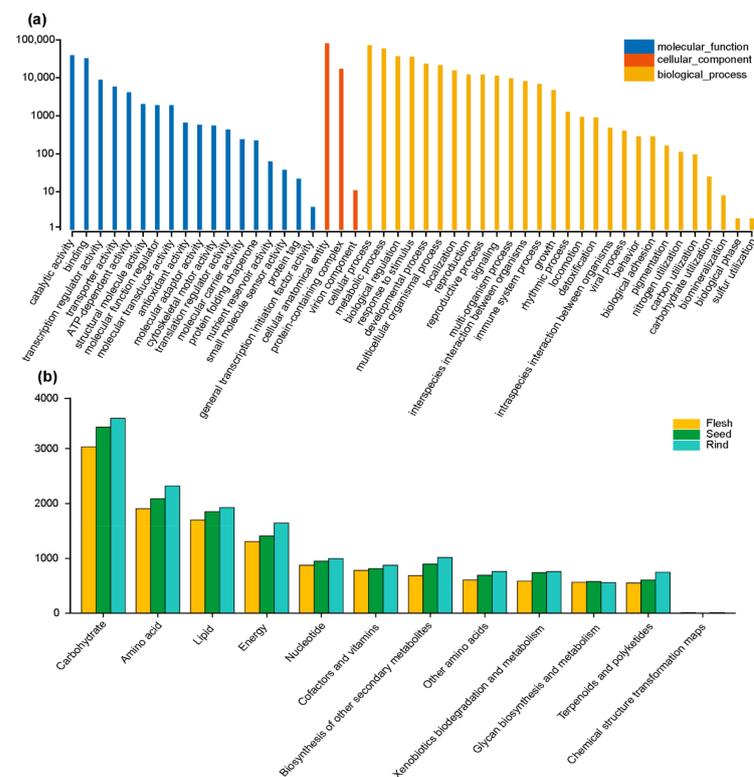


Figure 3. GO and KEGG metabolic pathway annotation of unigenes containing SSRs. (a) Distribution of GO classification of unigenes carrying SSRs; (b) Distribution of KEGG classification of unigenes carrying SSRs.

The results of KEGG analysis showed that 22,687 (5.78%) out of 392,223 unigenes containing SSRs were assigned to 41,757 (flesh: 12,581, seed: 14,009, rind: 15,167) annotated metabolic pathways (Table S4), and they were classified into 12 categories, among which carbohydrate (flesh: 3031, seed: 3392, rind: 3554), amino acid (flesh: 1903, seed: 2092, rind: 2321), and lipid metabolism (flesh: 1697, seed: 1852, rind: 1930) were the three largest categories in terms of number (Figure 3b). In addition, many other valuable metabolic pathways, such as nucleotide, terpenoid, and polyketide metabolism, were found in the fruit tissue of *A. trifoliata* (Table S4).

3.5. Development of Functional Molecular Markers

Among 22,687 unigenes carrying SSRs and annotated into KEGG metabolic pathways, a total of 12,114 (53.40%) unigenes could be used to design the primers for EST-SSRs, while the remaining 10,573 (46.60%) unigenes were not suitable for developing EST-SSR markers due to the short flanking sequences, although they also contained SSRs. A total of 16,869 EST-SSR primer pairs were successfully designed from 12,114 unigenes, in which 3017 (24.91%) unigenes could be used to design more than one EST-SSR primer pair because they carried more than one SSR (Table S5). This batch of functional EST-SSR markers was predictably associated with 12 classes of metabolic pathways based on KEGG functional annotation of unigenes. Among 16,869 primer pairs involved in 23,102 metabolic pathways, 4993 (21.61%) were related to carbohydrate metabolism, 3193 (13.82%) primer pairs were related to lipid metabolism, 2919 (12.64%) primer pairs were related to amino acid metabolism, 1298 (5.62%) primer pairs were related to biosynthesis of other secondary metabolites, and 1171 (5.07%) primer pairs were related to metabolism of terpenoids and polyketides (Table S5). In addition, we noted that 16,869 EST-SSR markers mainly consisted of 10,034 (59.48%) SNRs, 3628 (21.51%) DNRs, and 1659 (9.83%) TNRs, while only 153 TtNRS, 4 PNRs, and 3 HNRs were identified; the other 1388 (8.23%) markers were combined SSR types.

3.6. Validation of EST-SSR Markers

A total of 100 pairs of EST-SSR primers were randomly selected from the primers derived from DNRs and TNRs (Table S6), in which 94 primers successfully amplified clear bands and 72 primers produced polymorphic bands in a population consisting of 45 accessions of *A. trifoliata* (Figure S1). A total of 270 alleles were detected using 72 polymorphic primer pairs, and the average number of alleles and polymorphic information content of each locus was 3.75, ranging from 1 to 9, and 0.396, ranging from 0 to 0.831, respectively.

4. Discussion

4.1. Gene Expression Profile of Various Tissues of *A. trifoliata* Fruit

A. trifoliata is a newly domesticated horticultural crop widely cultivated in Southeast Asia due to its multiple commercial purposes. The great economic value of *A. trifoliata* is primarily embodied in its fruits, and specific fruit tissues have various commercial utilizations. For example, the rind is an important component of medicinal plants because it is rich in secondary metabolites, such as saponins and flavones [2], while the flesh is the main edible part of this delicious fresh fruit due to its rich nutritional content, including vitamins, sugars, minerals, and free amino acids [2]. The seeds are the tissue of high interest for cultivating *A. trifoliata* as an oil crop [4]. Therefore, outlining the gene expression profile in different tissues of *A. trifoliata* fruit is valuable for effectively mining functional genes and improving traits of commercial interest. In the present study, the total RNA of three tissues (rind, flesh, and seed) of *A. trifoliata* fruit at four different developmental stages (young, enlargement, coloring, and maturity) was sequenced by Illumina sequencing technology, and a total of 271.49 Gb of raw transcriptomic data were produced and submitted to NCBI. The available data have been downloaded to study the expression profile of functional genes and repeat sequences by various authors [20,31]. The large amount of expression data provides abundant molecular information for further identifying functional genes involved

in specific metabolic pathways and developing EST markers, which could accelerate genetic improvement by molecular marker-assisted selection.

4.2. Distribution, Composition, and Expression Profile of SSR Loci in *A. trifoliata* Fruit Tissues

Functionally, SSRs were initially regarded as “junk” and “neutral” DNA sequences, so the traditional view was that they represented chromosomal regions with few functional genes [40]. However, with the development of whole-genome sequencing technology, genomic data analysis of many species found that there is high SSR density in the flanking regions of functional genes [41,42], which indicated that SSRs could be contained anywhere in chromosomes. In fact, some studies have confirmed that SSRs widely exist in coding and noncoding regions [43,44]. Expressed sequences also contain a large number of SSRs, and moreover, they can exhibit differential expression levels, similar to functional genes.

We found that SSRs were highly enriched in the expressed sequences of various fruit tissues at various developmental stages in *A. trifoliata*. Comparative analysis of the expression of SSRs in *A. trifoliata* fruit tissues revealed that the expression frequencies of different SSRs were similar in different fruit tissues and developmental stages, the proportion of dominant SSRs was very consistent (Figure 1b–d), and the average density of EST-SSRs was 549.53 per Mb (Table 1). A previous study showed that the average density of g-SSRs was 665.28 per Mb in the *A. trifoliata* genome [20], which suggested that the SSR density of coding regions could generally be lower than that of noncoding regions. In addition, in all three fruit tissues, the proportion of unigenes carrying SSRs was more than 28.00%, which was noticeably higher than that in rice (4.7%) [45], maize (1.5%) [45], wheat (7.41%) [46], and radish (7.93%) [47]. Therefore, it is evident that many genes in the *A. trifoliata* genome contain one or more SSR loci.

In addition, the proportion (98.71%) of short repeat unit SSRs (SNRs, DNRs, and TNRs) in the transcriptome was exceptionally close to that (98.67%) in the *A. trifoliata* genome, and the proportion (60.63%) of SNRs in the transcriptome was obviously lower than that (66.98%) in the genomic data. In contrast, the proportion (9.69%) of TNRs in the transcriptome was obviously higher than that (5.60%) in the genome [20], which further supported the view that most SSRs in the coding region are TNRs [22,41]. Previous studies suggested that SSRs in coding regions could be widely subjected to purifying selection, during which only TNRs could be saved due to the nature of translation and triplet codon slippage of the open reading frame [47–49]. At the same time, the expression abundance of some SSRs is related to the evolution of organisms [50]. For example, AT/TA usually has a lower expression frequency than other motifs of DNRs in plants, especially in dicots [51]. In this study, we also found that there was some variance in the motif composition of the same repeat unit between transcriptomic and genomic data. For example, AT/TA (53.35%) was the dominant motif of DNRs in the genome [20], while AG/CT was the most abundant (64.17%) motif of DNRs in the expressed sequences (Table S1). Similarly, in the genome, AAT/ATT accounted for 45.19% of TNRs [20], while it only accounted for an average of 22.26% of TNRs in expressed sequences (Table S1). Some reports have found that the preference for distinct types of SSRs in functional regions of genes is also wide in plants [42,52]. Therefore, these results indicated that the motif composition of the same repeat unit could be different between genic and intergenic regions.

4.3. The Expression of TNRs Is Spatiotemporally Specific in Fruit Tissues of *A. trifoliata*

We further focused more attention on the differential expression profile of TNRs because they are usually more common than other repeat units in coding regions [53,54]. In the present study, among 60 TNR motifs, 15 motifs exhibited differential expression in specific fruit tissues, while three motifs (ACA, GCC, and GTA) had a differential expression level at selected developmental stages (Table 2), which showed that tissue-specific and developmental stage-specific expression of TNRs is common in *A. trifoliata*, and this differential expression could directly couple with the differential expression of functional genes.

Patterns of DNA double-strand transcription mainly include asymmetric transcription and antisense transcription [55]. A large number of protein-coding genes follow asymmetric transcription rules so that only one coding strand is transcribed. In contrast, the expression of natural antisense transcripts allows complementary strand sequences to be simultaneously transcribed [56], which can regulate gene transcription by forming double-stranded RNA with the mRNA coding [57,58]. In this study, the results of the chi-square test showed that only two complementary motif combinations (ATA/TAT and CCT/GGA) complied with the 1:1 ratio in all 12 transcriptomes of fruit tissue, while many complementary TNR combinations significantly deviated from the 1:1 ratio in all or some samples (Figure 2; Table S3), possibly resulting from the asymmetric expression of double-stranded DNA, which could commonly occur in various fruit tissues at different developmental stages. The complementary TNR expression profile afforded useful information about natural antisense transcription, which further helped us to understand gene expression regulation.

4.4. Development of a Set of Functional EST-SSR Markers and Their Application Prospects

EST-SSR markers are a class of accessible and versatile functional markers because they are directionally derived from expressed sequences associated with functional genes. Although 9394 SSRs were identified and 100 EST-SSR markers were developed according to the transcriptome data of *A. trifoliata*, the transcriptome data were derived from the leaf (which has low economic value) rather than the fruit (which has high economic value) [59]. The markers had less applicable prospects because only the EST-SSRs derived from the transcriptome data of objective organs could be highly associated with corresponding traits of interest [60–62]. In this study, the transcriptome data of three fruit tissues were employed to search for SSR loci in expressed sequences, and we further successfully developed 16,86 EST-SSR primer pairs.

In addition, the 12,114 unigenes deriving 16,869 EST-SSR markers were functionally annotated, among which some were involved in metabolic pathways influencing fruit nutrients, such as carbohydrate (3651), amino acid (2200), and vitamin (1274) metabolism, while others were involved in metabolic pathways influencing oil content and components, such as lipids (2217) and unigenes involved in the biosynthesis of secondary metabolites, such as terpenoids or polyketides (876), glycans (798), and xenobiotics (593) (Table S5). Therefore, when we tried to improve the nutrient characteristics of *A. trifoliata* as a new delicious fresh fruit, 9637 EST-SSR markers (4993, 2919, and 1725 associated with carbohydrates, amino acids, and vitamins, respectively) should be more powerful than the remaining markers (Table S5). Likewise, 3193 EST-SSR markers associated with lipid metabolism and 1298 EST-SSR markers associated with secondary metabolites could be more valuable than the other markers when exploiting *A. trifoliata* as an oil crop and medicinal plant, respectively.

Finally, PCR amplification results showed that 94 out of 100 primer pairs produced clear products, and 72 primer pairs exhibited polymorphism, which suggested that the EST-SSR markers developed in this study were dependable and effective. Although the PIC value (0.396) was slightly lower than that of the genomic SSR markers developed according to the genome sequence [20], partly due to the more conserved genic regions compared with intergenic regions across the whole genome [63], the high polymorphism of the set of EST-SSR markers should meet almost all the needs of molecular markers in the genetic improvement field. It was evident that the newly developed 16,869 EST-SSR markers could have large application prospects in the future.

5. Conclusions

In conclusion, a complete set of transcriptomic data of three *A. trifoliata* fruit tissues at four developmental stages was published in this study. The identification of SSRs in de novo assembled transcriptome data showed that the fruit tissues of *A. trifoliata* possess abundant SSRs, and the expression profile of the SSRs was outlined. The proportion of the same type of SSR motifs was very similar in different fruit tissues and developmental stages, but the expression levels of some TNRs were different in different tissues or developmental stages.

We further developed a total of 16,869 EST-SSRs according to the 12,114 assembled unigene sequences, and their validation was confirmed by PCR amplification. Finally, comparative analysis of expression showed that complementary TNRs widely exhibited tissue- and time-specific expression preferences, which could be related to gene expression regulation by natural antisense transcripts. These 16,869 EST-SSR markers, especially the 1659 (9.83%) TNR EST-SSR markers, could be used as powerful molecular tools for molecular selection, functional gene identification, and gene mining in *A. trifoliata* in the future.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/genes13081451/s1>, Figure S1: EST-SSR marker amplification results were separated by nondenaturing polyacrylamide gel; Table S1: The number and proportion of various types of SSR motifs in the fruit tissues of *A. trifoliata*; Table S2: Analysis of variance for trinucleotide repeat motifs; Table S3: Chi-square test for complementary trinucleotide repeat motifs; Table S4: KEGG annotation of unigenes carrying SSRs; Table S5: Information about EST-SSR primers in the fruit tissue of *A. trifoliata*; Table S6: 100 pairs of EST-SSRs that were synthesized and polymorphically validated.

Author Contributions: Conceptualization, W.C., J.S., Q.L. and P.L.; methodology, H.Y.; software, H.Y.; validation, W.C., H.Y. and Q.Z.; formal analysis, W.C. and J.Z.; resources, P.L.; data curation, S.Z.; writing—original draft preparation, W.C. and H.Y.; writing—review and editing, H.Y. and P.L.; visualization, W.C.; supervision, Z.L., T.R. and F.T.; project administration, P.L.; funding acquisition, P.L. and Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Department of Sichuan Province, grant numbers 2020YFN0057, 2020YFN0091, and 2020YJ0331, and the Bureau of Science and Technology of Fuling of Chongqing, grant number FLKJ-2021ABB1016.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data analyzed during this study are included in the manuscript and supplementary information files, and transcriptomic data of *A. trifoliata* fruit tissues have been deposited in the National Center for Biotechnology Information (NCBI) database (<https://www.ncbi.nlm.nih.gov/sra> (accessed on 29 October 2020)) under accession numbers SAMN16551931–33, SAMN16551934–36, SAMN16551937–39, and SAMN16551940–42.

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

References

1. Yang, H.; Chen, W.; Fu, P.; Zhong, S.; Guan, J.; Luo, P. Developmental stages of *Akebia trifoliata* fruit based on volume. *Hortic. Sci. Technol.* **2021**, *39*, 823–831.
2. Li, L.; Yao, X.; Zhong, C.; Chen, X.; Huang, H. *Akebia*: A potential new fruit crop in China. *HortScience* **2010**, *45*, 4–10. [[CrossRef](#)]
3. Guan, J.; Fu, P.; Wang, X.; Yu, X.; Zhong, S.; Chen, W.; Yang, H.; Chen, C.; Yang, H.; Luo, P. Assessment of the breeding potential of a set of genotypes selected from a natural population of *Akebia trifoliata* (Three-Leaf Akebia). *Hortic. Sci. Technol.* **2022**, *8*, 116. [[CrossRef](#)]
4. Bai, C. Fatty acids analysis in *Decaisnea insignis* and *Akebia trifoliata* seed oil by GC-MS. *Acta Bot. Boreali-Occident. Sin.* **2007**, *5*, 1035–1038.
5. Maciag, D.; Dobrowolska, E.; Sharafan, M.; Ekiert, H.; Tomczyk, M.; Szopa, A. *Akebia quinata* and *Akebia trifoliata*—a review of phytochemical composition, ethnopharmacological approaches and biological studies. *J. Ethnopharmacol.* **2021**, *280*, 114486. [[CrossRef](#)] [[PubMed](#)]
6. Zou, S.; Gao, P.; Jia, T.; Huang, H. Physicochemical characteristics and nutritional composition during fruit ripening of *Akebia trifoliata* (Lardizabalaceae). *Hortic. Sci. Technol.* **2022**, *8*, 326. [[CrossRef](#)]
7. Zhong, W.; Ma, Y. Fruit nutrition quality of a good *Akebia trifoliata* plant. *Guizhou Agric. Sci.* **2015**, *43*, 140–142.
8. Yi, F.; Yong, B.I.; Xiao-Chun, L.I.; Yan-Bo, H.U.; Xiang-Lin, L. Provenance selection and biological characteristics of fruit type *Akebia trifoliata* (Thunb.) Koidz. *Hubei Agric. Sci.* **2021**, *60*, 68–71.
9. Su, S.; Wu, J.; Peng, X.; Li, B.; Li, Z.; Wang, W.; Ni, J.; Xu, X. Genetic and agro-climatic variability in seed fatty acid profiles of *Akebia trifoliata* (Lardizabalaceae) in China. *J. Food Compos. Anal.* **2021**, *102*, 104064. [[CrossRef](#)]
10. Yang, X.; Zhang, Y.; Lin, J.; Pen, A.; Ying, C.; Cao, W.; Mao, L. A lower proportion of dietary saturated/monounsaturated/polyunsaturated fatty acids reduces the expression of adiponectin in rats fed a high-fat diet. *Nutr. Res.* **2012**, *32*, 285–291. [[CrossRef](#)] [[PubMed](#)]

11. Jena, K.K.; Mackill, D.J. Molecular markers and their use in marker-assisted selection in rice. *Crop Sci.* **2008**, *48*, 1266–1276. [[CrossRef](#)]
12. Prasanna, B.M.; Pixley, K.; Warburton, M.L.; Xie, C. Molecular marker-assisted breeding options for maize improvement in Asia. *Mol. Breed.* **2010**, *26*, 339–356. [[CrossRef](#)]
13. Fan, Y.; Wang, Q.; Dong, Z.; Yin, Y.; Teixeira Da Silva, J.A.; Yu, X. Advances in molecular biology of *Paeonia* L. *Planta* **2020**, *251*, 23. [[CrossRef](#)] [[PubMed](#)]
14. Gahlaut, V.; Kumari, P.; Jaiswal, V.; Kumar, S. Genetics, genomics and breeding in *Rosa* species. *J. Hortic. Sci. Biotechnol.* **2021**, *96*, 545–559. [[CrossRef](#)]
15. Maleki, M.; Shojaeiyan, A.; Rashidi Monfared, S. Population structure, morphological and genetic diversity within and among melon (*Cucumis melo* L.) landraces in Iran. *J. Genet. Eng. Biotechnol.* **2018**, *16*, 599–606. [[CrossRef](#)]
16. Gupta, P.K.; Varshney, R.K. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* **2000**, *113*, 163–185. [[CrossRef](#)]
17. Joshi, S.P.; Ranjekar, P.K.; Gupta, V.S. Molecular markers in plant genome analysis. *Curr. Sci. India* **1999**, *77*, 230–240.
18. Ni, J.; Colowit, P.M.; Mackill, D.J. Evaluation of genetic diversity in rice subspecies using microsatellite markers. *Crop Sci.* **2002**, *42*, 601–607. [[CrossRef](#)]
19. Kaur, S.; Panesar, P.S.; Bera, M.B.; Kaur, V. Simple sequence repeat markers in genetic divergence and marker-assisted selection of rice cultivars: A review. *Crit. Rev. Food Sci.* **2015**, *55*, 41–49. [[CrossRef](#)]
20. Zhong, S.; Chen, W.; Yang, H.; Shen, J.; Ren, T.; Li, Z.; Tan, F.; Luo, P. Characterization of microsatellites in the *Akebia trifoliata* genome and their transferability and development of a whole set of effective, polymorphic, and physically mapped simple sequence repeat markers. *Front. Plant Sci.* **2022**, *13*, 860101. [[CrossRef](#)]
21. Zhang, L.; Zuo, K.; Zhang, F.; Cao, Y.; Wang, J.; Zhang, Y.; Sun, X.; Tang, K. Conservation of noncoding microsatellites in plants: Implication for gene regulation. *BMC Genom.* **2006**, *7*, 323. [[CrossRef](#)] [[PubMed](#)]
22. Zeng, S.; Xiao, G.; Guo, J.; Fei, Z.; Xu, Y.; Roe, B.A.; Wang, Y. Development of a EST dataset and characterization of EST-SSRs in a traditional Chinese medicinal plant, *Epimedium sagittatum* (Sieb. Et Zucc.) Maxim. *BMC Genom.* **2010**, *11*, 94. [[CrossRef](#)] [[PubMed](#)]
23. Gong, L.; Deng, Z. EST-SSR markers for gerbera (*Gerbera hybrida*). *Mol. Breed.* **2010**, *26*, 125–132. [[CrossRef](#)]
24. Feng, S.P.; Li, W.G.; Huang, H.S.; Wang, J.Y.; Wu, Y.T. Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). *Mol. Breed.* **2009**, *23*, 85–97. [[CrossRef](#)]
25. González, M.; Salazar, E.; Castillo, J.; Morales, P.; Mura-Jornet, I.; Maldonado, J.; Silva, H.; Carrasco, B. Genetic structure based on EST-SSR: A putative tool for fruit color selection in Japanese plum (*Prunus salicina* L.) breeding programs. *Mol. Breed.* **2016**, *36*, 68. [[CrossRef](#)]
26. Ercan, S.; Ertugrul, F.; Aydin, Y.; Akfirat, F.S.; Hasancebi, S.; Akan, K.; Mert, Z.; Bolat, N.; Yorgancilar, O.; Altinkut-Uncuoglu, A. An EST-SSR marker linked with yellow rust resistance in wheat. *Biol. Plant.* **2010**, *54*, 691–696. [[CrossRef](#)]
27. Consortium, G.O. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **2004**, *32*, D258–D261. [[CrossRef](#)] [[PubMed](#)]
28. Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids. Res.* **2000**, *28*, 27–30. [[CrossRef](#)] [[PubMed](#)]
29. Nahas, L.D.; Alsamman, A.M.; Hamwiah, A.; Al-Husein, N.; Lababidi, G. Characterization of EST-SSR markers in bread wheat EST related to drought tolerance and functional analysis of SSR-containing unigenes. *Highlights BioSci.* **2020**, *3*, 1–12. [[CrossRef](#)]
30. Chen, H.; Wang, L.; Liu, X.; Hu, L.; Wang, S.; Cheng, X. De novo transcriptomic analysis of cowpea (*Vigna unguiculata* L. Walp.) for genic SSR marker development. *BMC Genet.* **2017**, *18*, 65. [[CrossRef](#)]
31. Yu, X.; Zhong, S.; Yang, H.; Chen, C.; Chen, W.; Yang, H.; Guan, J.; Fu, P.; Tan, F.; Ren, T.; et al. Identification and characterization of NBS resistance genes in *Akebia trifoliata*. *Front. Plant Sci.* **2021**, *12*, 758559. [[CrossRef](#)]
32. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
33. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)]
34. Thiel, T.; Michalek, W.; Varshney, R.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)]
35. Huerta-Cepas, J.; Forslund, K.; Coelho, L.P.; Szklarczyk, D.; Jensen, L.J.; von Mering, C.; Bork, P. Fast Genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **2017**, *34*, 2115–2122. [[CrossRef](#)] [[PubMed](#)]
36. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)] [[PubMed](#)]
37. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—New capabilities and interfaces. *Nucleic Acids. Res.* **2012**, *40*, e115. [[CrossRef](#)] [[PubMed](#)]
38. Murray, M.G.; Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids. Res.* **1980**, *8*, 4321–4326. [[CrossRef](#)]
39. Lawson, M.J.; Zhang, L. Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biol.* **2006**, *7*, R14. [[CrossRef](#)] [[PubMed](#)]

40. John, B.; Hewitt, G.M. Patterns and pathways of chromosome evolution within the Orthoptera. *Chromosoma* **1968**, *25*, 40–74. [[CrossRef](#)]
41. Zhang, L.; Yuan, D.; Yu, S.; Li, Z.; Cao, Y.; Miao, Z.; Qian, H.; Tang, K. Preference of simple sequence repeats in coding and noncoding regions of *Arabidopsis thaliana*. *Bioinformatics* **2004**, *20*, 1081–1086. [[CrossRef](#)]
42. Morgante, M.; Hanafey, M.; Powell, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* **2002**, *30*, 194–200. [[CrossRef](#)] [[PubMed](#)]
43. Fox, G.; Preziosi, R.F.; Antwis, R.E.; Benavides-Serrato, M.; Combe, F.J.; Harris, W.E.; Hartley, I.R.; Kitchener, A.C.; de Kort, S.R.; Nekaris, A.; et al. Multi-individual microsatellite identification: A multiple genome approach to microsatellite design (MiMi). *Mol. Ecol. Resour.* **2019**, *19*, 1672–1680. [[CrossRef](#)]
44. Fujimori, S.; Washio, T.; Higo, K.; Ohtomo, Y.; Murakami, K.; Matsubara, K.; Kawai, J.; Carninci, P.; Hayashizaki, Y.; Kikuchi, S.; et al. A novel feature of microsatellites in plants: A distribution gradient along the direction of transcription. *FEBS Lett.* **2003**, *554*, 17–22. [[CrossRef](#)]
45. Kantety, R.V.; La Rota, M.; Matthews, D.E.; Sorrells, M.E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* **2002**, *48*, 501–510. [[CrossRef](#)] [[PubMed](#)]
46. Peng, J.H.; Lapitan, N.L.V. Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. *Funct. Integr. Genom.* **2005**, *5*, 80–96. [[CrossRef](#)] [[PubMed](#)]
47. Zhai, L.; Xu, L.; Wang, Y.; Cheng, H.; Chen, Y.; Gong, Y.; Liu, L. Novel and useful genic-SSR markers from de novo transcriptome sequencing of radish (*Raphanus sativus* L.). *Mol. Breed.* **2014**, *33*, 611–624. [[CrossRef](#)]
48. Chen, H.; Chen, X.; Tian, J.; Yang, Y.; Liu, Z.; Hao, X.; Wang, L.; Wang, S.; Liang, J.; Zhang, L.; et al. Development of gene-based SSR markers in rice Bean (*Vigna umbellata* L.) based on transcriptome data. *PLoS ONE* **2016**, *11*, e151040. [[CrossRef](#)]
49. Ashrafi, H.; Hill, T.; Stoffel, K.; Kozik, A.; Yao, J.; Chin-Wo, S.R.; Van Deynze, A. De novo assembly of the pepper transcriptome (*Capsicum annuum*): A benchmark for in silico discovery of SNPs, SSRs and candidate genes. *BMC Genom.* **2012**, *13*, 571. [[CrossRef](#)] [[PubMed](#)]
50. Qin, Z.; Wang, Y.; Wang, Q.; Li, A.; Hou, F.; Zhang, L. Evolution analysis of simple sequence repeats in plant genome. *PLoS ONE* **2015**, *10*, e144108. [[CrossRef](#)] [[PubMed](#)]
51. Kumpatla, S.P.; Mukhopadhyay, S. Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome* **2005**, *48*, 985–998. [[CrossRef](#)] [[PubMed](#)]
52. Li, Y.; Korol, A.B.; Fahima, T.; Nevo, E. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* **2004**, *21*, 991–1007. [[CrossRef](#)] [[PubMed](#)]
53. Liu, L.; Qin, M.; Yang, L.; Song, Z.; Luo, L.; Bao, H.; Ma, Z.; Zhou, Z.; Xu, J. A genome-wide analysis of simple sequence repeats in *Apis cerana* and its development as polymorphism markers. *Gene* **2017**, *599*, 53–59. [[CrossRef](#)] [[PubMed](#)]
54. Niyitanga, S.; Yao, J.; Ibrahim, A.K.; Afzal, M.Z.; Chen, S.; Zhang, L. Development of genome-wide SSR markers for physical map construction with PCR-based polymorphic SSRs in Jute (*Corchorus* spp.). *Trop. Plant Biol.* **2022**, *15*, 65–77. [[CrossRef](#)]
55. Pelechano, V.; Steinmetz, L.M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **2013**, *14*, 880–893. [[CrossRef](#)] [[PubMed](#)]
56. Rosikiewicz, W.; Makołowska, I. Biological functions of natural antisense transcripts. *Acta Biochim. Pol.* **2016**, *63*, 665–673. [[CrossRef](#)]
57. Munroe, S.H.; Zhu, J. Overlapping transcripts, double-stranded RNA and antisense regulation: A genomic perspective. *Cell. Mol. Life Sci.* **2006**, *63*, 2102–2118. [[CrossRef](#)] [[PubMed](#)]
58. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C.C.; Suzuki, M.; Kawai, J.; et al. Antisense transcription in the mammalian transcriptome. *Science* **2005**, *309*, 1564–1566. [[CrossRef](#)]
59. Niu, J.; Wang, Y.; Shi, Y.; Wang, X.; Sun, Z.; Huang, K.; Gong, C.; Luan, M.; Chen, J. Development of SSR markers via de novo transcriptome assembly in *Akebia trifoliata* (Thunb.) Koidz. *Genome* **2019**, *62*, 817–831. [[CrossRef](#)]
60. Li, C.; Zhu, Y.; Guo, X.; Sun, C.; Luo, H.; Song, J.; Li, Y.; Wang, L.; Qian, J.; Chen, S. Transcriptome analysis reveals ginsenosides biosynthetic genes, microRNAs and simple sequence repeats in *Panax ginseng* C. A. Meyer. *BMC Genom.* **2013**, *14*, 245. [[CrossRef](#)]
61. Wu, J.; Cheng, F.; Cai, C.; Zhong, Y.; Jie, X. Association mapping for floral traits in cultivated *Paeonia rockii* based on SSR markers. *Mol. Genet. Genom.* **2017**, *292*, 187–200. [[CrossRef](#)] [[PubMed](#)]
62. Wan, Y.; Zhang, M.; Hong, A.; Zhang, Y.; Liu, Y. Characteristics of microsatellites mined from transcriptome data and the development of novel markers in *Paeonia lactiflora*. *Genes* **2020**, *11*, 214. [[CrossRef](#)] [[PubMed](#)]
63. Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **2011**, *177*, 309–334. [[CrossRef](#)]