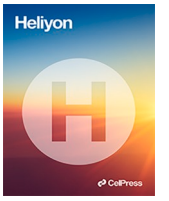
Heliyon 10 (2024) e26720

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Association mapping analysis for cultivated and weedy types of *Perilla* crop collected from South Korea using morphological characteristics and SSR markers

So Jung Jang [a,b,1], Kyu Jin Sa [a,c,1], Zhen Yu Fu [a,1], Ju Kyong Lee [a,b,*]

[a] *Department of Applied Plant Sciences, College of Agriculture and Life Sciences, Kangwon National University, Chuncheon, 24341, South Korea*
[b] *Interdisciplinary Program in Smart Agriculture, Kangwon National University, Chuncheon, 24341, South Korea*
[c] *Department of Crop Science, College of Ecology & Environmental Sciences, Kyungpook National University, Sangju, 37224, South Korea*

ARTICLE INFO

Keywords:
Perilla frutescens
Genetic diversity and relationship
Association mapping analysis
Population structure
Perilla SSR marker
Perilla crop breeding program

ABSTRACT

There are two cultivated and weedy types of *Perilla* crop (TCWTPC), and they are widely distributed and cultivated in East Asia, especially in South Korea and Japan. The objective of this study is to create simple sequence repeat (SSR) markers linked to morphological traits that show differences between accessions of the TCWTPC using recently designed SSR primer sets in *Perilla* crop. Genetic diversity within 52 accessions of the TCWTPC, gathered from South Korea, was assessed using 28 novel *Perilla* SSR primer sets. Based on the assessment, a collection of 28 *Perilla* SSR primer sets were shown to exhibit polymorphism and yielded a total of 142 alleles across the 52 accessions of the TCWTPC. Through inspection of a phylogenetic tree and population structure, the 52 accessions of the TCWTPC were classified into three major groups. Although most accessions of the TCWTPC were relatively clearly distinguished, SSR markers failed to distinguish several accessions belonging to the two weedy types of the *Perilla* crop. By using an association mapping analysis (AMA) of the 28 *Perilla* SSR markers and seven morphological characteristics in the 52 TCWTPC accessions, we detected that three of the *Perilla* SSR markers (KNUPF134, KNUPF137, KNUPF149) were associated with plant and seed characteristics. The novel SSR primer sets developed in *Perilla* crop should be useful in AMA for assessing genetic diversity and relationships between and within TCWTPC accessions, and this information will be helpful for genetic mapping in breeding programs for *Perilla* crop.

1. Introduction

Perilla frutescens Britt. is a self-fertilizing species of Labiatae. The species includes two cultivated types (or two varieties) [cultivated var. *frutescens* (CF) and cultivated var. *crispa* (CC)] based on their uses in East Asia. For a long time in East Asia, CF has been used as an oil or vegetable crop, and CC has been used as an herbal medicine or vegetable crop [1–6]. Today, the two cultivated and weedy types of *Perilla* crop (TCWTPC) have the most extensive distribution and cultivation in South Korea and Japan [4–7]. CF is widely cultivated in South Korea, serving both as a seed oil crop and a leafy vegetable crop. In particular in South Korea, as the consumption of fresh

* Corresponding author. Department of Applied Plant Sciences, College of Agriculture and Life Sciences, Kangwon National University, Chuncheon, 24341, South Korea.
E-mail address: jukyonglee@kangwon.ac.kr (J.K. Lee).
[1] S.J. Jang, K.J. Sa, and Z.Y. Fu contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e26720
Received 2 July 2023; Received in revised form 25 January 2024; Accepted 19 February 2024
Available online 28 February 2024

leaves and seed oil has been increasing, the cultivation area of CF has also greatly increased [8,9]. Meanwhile CC is not actively cultivated in South Korea due to a decline in its utilization as an herbal medicine, and it is frequently encountered in the form of a relict [4,5]. On the other hand, CC is widely cultivated in Japan, and its fresh leaves are used as vegetables and pickles, while CF is rarely cultivated in Japan and is only occasionally cultivated in some mountainous areas of the central and northern regions as a relict form [3,6].

Perilla crop has attracted attention worldwide in recent years as a cash crop for uses including as food oil, vegetable, health functional food, and herbal medicine plant [8,10]. In particular, the seeds of CF are a good source of polyunsaturated fatty acids, encompassing both linoleic acid and linolenic acid [8,9,11]. Furthermore, the Perilla leaves of the TCWTPC have a difference in color, green for CF (Fig. 1A) and purple for CC (Fig. 1B). As such, the TCWTPC are well distinguished based on their morphological and physiological characteristics (as shown in Fig. 1). CF has mostly large and soft seeds (more than 2 mm), non-wrinkly green leaves, and its own specific aroma (by *Perilla* ketone ingredients). In contrast, CC is characterized by diminutive and firm seeds (measuring less than 2 mm), purple or green wrinkly or non-wrinkly leaves, and its own specific aroma (by *Perilla* aldehyde ingredients). The original ancestry of the TCWTPC remains unidentified in East Asia, including Korea. However, two weedy varieties, namely weedy var. *frutescens* (WF) and weedy var. *crispa* (WC), are frequently encountered in South Korea, often being observed in areas like wastelands, around farmers' houses, and within farmers' fields [4,9,12–14]. In addition, many plant taxonomists have conducted research to differentiate between and within accessions of the TCWTPC. As mentioned earlier, while morphological characteristics such as leaf color and seed size effectively distinguish the two cultivated types (CF, CC) of *Perilla* crop, the same criteria do not distinctly differentiate between the two weedy types (WF, WC) of *Perilla* crop. The TCWTPC have the same chromosome number of $2n = 40$ and can be crossed by artificial crossing, and also intermediate hybrid types exist in the natural habitat [15–20]. Therefore, there are still many difficulties in identifying between and within accessions of the TCWTPC by morphological characteristics.

In addition to morphological characteristics, various molecular marker technologies have been used to understand the genetic variation between cultivated and wild accessions of many crop species. In the case of the TCWTPC as well, many genetic diversity (GD) studies have been conducted among accessions of the TCWTPC using various DNA-based molecular markers such as random amplified polymorphic DNA (RAPD) [6,21], amplified fragment length polymorphism (AFLP) [5,12], and *Perilla* SSRs [20,22]. In particular, in *Perilla* crop, *Perilla* SSRs or microsatellite markers are recognized as the preferred molecular markers for GD and genetic relationship

(A)

(B)

Fig. 1. The plant characteristics of two varieties of *Perilla* crop. (A) var. *frutescens*, (B) var. *crispa*.

studies between and within accessions of the TCWTPC. This preference is attributed to their codominance, high reproducibility, DNA polymorphism, and ease of use with polymerase chain reaction (PCR)-marker technology [13,23–25]. However, the amount of developed *Perilla* SSR primer sets in the TCWTPC remains insufficient for comprehensive genetic diversity and relationships studies between and within accessions of the TCWTPC. Owing to the advantages of *Perilla* SSR markers, studies are in progress on the development of molecular markers related to useful traits in the TCWTPC. Recently, Fu et al. (2022) developed 30 new *Perilla* SSR primer sets (mainly dinucleotide SSR type) from *Perilla* crop using RNA sequencing and used 17 of these *Perilla* SSR primer sets for GD analysis and phylogenetic relationship studies among accessions of the TCWTPC collected in South Korea. Furthermore, many researchers in our laboratory team recently reported that they had developed new SSR primer sets in *Perilla* crop [13,22,26], and they have used these successfully for studies on GD, phylogenetic relationships, and association mapping analysis (AMA) among TCWTPC accessions.

Hence, the objective of this study is to develop molecular markers related to morphological traits that show differences between and within accessions of the TCWTPC using new *Perilla* SSR primer sets recently developed by Fu et al. [13]. Expected outcomes from this study include valuable markers for future differentiation of TCWTPC accessions and the identification of molecular markers for advantageous morphological traits in breeding programs for *Perilla* crop, including applications like marker-assisted selection (MAS).

2. Materials and methods

2.1. Plant materials and DNA extraction

This study used 52 accessions of the TCWTPC (21 CF, 14 WF, 17 WC) that were selected from previous studies by Fu et al. [13,14] (Supplementary data, Table S1). For SSR analysis, *the* total DNA of the 52 accessions of the TCWTPC was extracted from young leaves at the seedling stage of each *Perilla* accession based on the report of Fu et al. [13].

2.2. Morphological characteristic survey and SSR analysis

To evaluate the morphological variation between and within accessions of *the TCWTPC*, this study used morphological data previously reported by Fu et al. [14] of the 52 *Perilla* accessions selected from South Korea, which were used for genetic variation analysis in previous studies by Fu et al. [13,14]. Fu et al. [14] investigated one quantitative and 10 qualitative morphological characteristics for the 52 accessions of the TCWTPC (*S*upplementary data, Table S2). Also, this study used 28 novel *Perilla* SSR primer sets that were recently developed by Fu et al. [13] (*S*upplementary data, Table S3). In accordance with a method recently published by Fu et al. [13],

Table 1
Characteristics of the 28 *Perilla* SSR primer sets including allele size range, allele number, MAF, GD and PIC among 52 accessions of the TCWTPC collected from South Korea.

SSR loci	Allele size range (bp)	No. of alleles	MAF	GD	PIC
KNUPF132	210–240	4	0.750	0.408	0.374
KNUPF133	160–180	4	0.769	0.382	0.350
KNUPF134	155–180	13	0.308	0.851	0.838
KNUPF135	200–220	2	0.846	0.260	0.226
KNUPF136	190–220	3	0.808	0.329	0.305
KNUPF137	190–230	8	0.442	0.728	0.694
KNUPF138	220–260	3	0.519	0.548	0.449
KNUPF139	130–150	6	0.365	0.735	0.691
KNUPF140	135–175	3	0.904	0.178	0.170
KNUPF141	140–170	5	0.635	0.543	0.498
KNUPF142	300–315	3	0.923	0.144	0.138
KNUPF143	150–180	5	0.577	0.587	0.532
KNUPF144	170–190	6	0.635	0.558	0.525
KNUPF145	200–220	5	0.769	0.377	0.340
KNUPF146	180–190	4	0.654	0.505	0.447
KNUPF147	200–230	5	0.442	0.695	0.646
KNUPF148	165–190	13	0.192	0.874	0.861
KNUPF149	190–210	3	0.673	0.489	0.435
KNUPF150	170–190	4	0.692	0.482	0.443
KNUPF151	175–185	5	0.423	0.704	0.655
KNUPF152	110–145	5	0.462	0.601	0.520
KNUPF154	130–145	4	0.673	0.463	0.387
KNUPF155	160–190	4	0.538	0.558	0.469
KNUPF156	180–220	3	0.750	0.384	0.324
KNUPF157	160–180	7	0.365	0.787	0.761
KNUPF158	180–240	5	0.712	0.467	0.440
KNUPF159	150–175	5	0.558	0.605	0.549
KNUPF160	200–225	5	0.481	0.666	0.614
Average		5.07	0.602	0.532	0.489

*MAF: Major Allele Frequency, GD: Gene Diversity, PIC: Polymorphism Information Content.

we performed SSR amplification and DNA electrophoresis experiments on the 52 accessions of the TCWTPC.

2.3. Data analysis

Polymorphic information content (PIC), number of alleles (NA), GD, and major allele frequency (MAF) were calculated for the 52 accessions of the TCWTPC and the 28 novel *Perilla* SSR primer sets, utilizing Power Marker version 3.25 [27]. Genetic similarity (GS) was calculated using the Dice similarity index [28]. The phylogenetic relationships among the 52 TCWTPC accessions were examined based on the UPGMA method, employing SAHN clustering within the NTSYS-pc V2.1 application [29]. Population structure analysis (PSA) for the 52 TCWTPC accessions was conducted using the Bayesian model-based clustering approach implemented in STRUCTURE v2.3 software [30]. The optimum *k* value for the entire population was determined based on the simulation method of delta *K* (ΔK) value [31] using the web-based software program STRUCTURE HARVESTER (https://taylor0.biology.ucla.edu/structureHarvester/). Finally, AMA was carried out with TASSEL 3.0 software [32], employing a general linear model (Q GLM) to assess marker-trait associations (MTAs). The Q GLM method utilized in this study is consistent with the approach detailed in a prior report by Park et al. [9].

3. Results

3.1. SSR variation of TCWTPC accessions from South Korea

Genetic variation at the 28 *Perilla* SSR primer sets in the 52 accessions (CF, WF, WC) of the TCWTPC collected from South Korea was measured with regard to NA, MAF, GD, and PIC (Table 1). The total of 28 *Perilla* SSR primer sets detected a total of 142 alleles among the 52 accessions of the TCWTPC. The NA values for each *Perilla* SSR primer set ranged from 2 (KNUPF135) to 13 (KNUPF134 and KNUPF148) with an average of 5.07. The MAF values for each *Perilla* SSR primer set ranged from 0.192 (KNUPF 148) to 0.923 (KNUPF 142) with an average of 0.602. The range of GD values for each *Perilla* SSR primer set varied from 0.144 (KNUPF 142) to 0.874 (KNUPF148) with a mean of 0.532. The range of PIC values for each *Perilla* SSR primer set varied from 0.138 (KNUPF 142) to 0.861 (KNUPF148) with a mean of 0.489 (Table 1).

3.2. Population structure analysis (PSA) and phylogenetic relationships between TCWTPC accessions and their weedy types collected from South Korea using Perilla SSR markers

According to the PSA results for the 52 accessions of the TCWTPC using the STRUCTURE software with ΔK based on the optimal grouping method, the ΔK value reached the highest peak value when K = 2 (Fig. 2). Therefore, based on membership probability >0.75, the optimal grouping number was 2. The 52 accessions of the TCWTPC were separated into two main groups (Groups I, II) and a mixed group with K = 2 (Fig. 3). Group I consisted of 21 CF accessions and three WF accessions. Group II consisted of nine WF accessions and 17 WC accessions. The mixed group included only two WF accessions (Fig. 3).

Meanwhile, the UPGMA dendrogram results for the 52 accessions of the TCWTPC (Fig. 4) show that the 52 accessions of the TCWTPC were classified into three groups with GS of 41.2%. Group I contained 21 CF accessions and five WF accessions. Group II

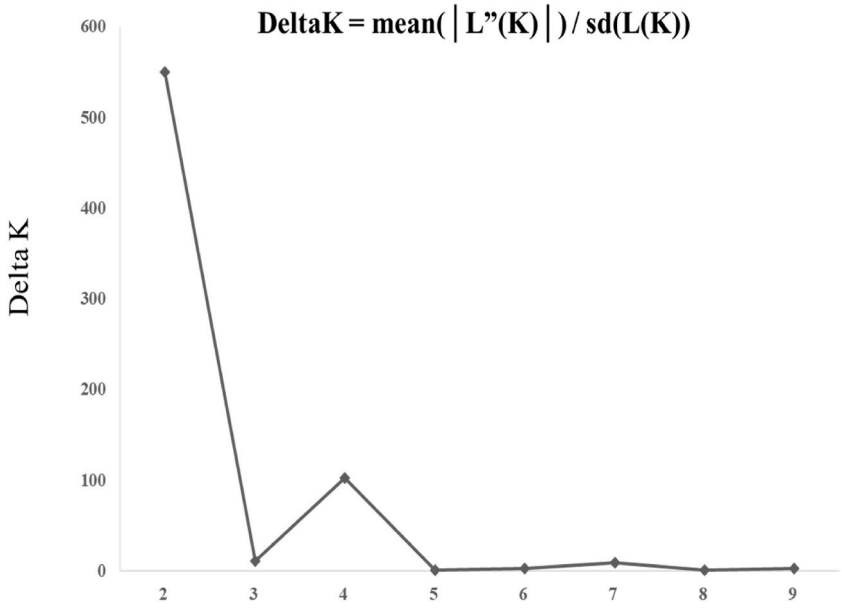


Fig. 2. Magnitude of *ΔK* as a function of *K*; the peak value of *ΔK* was at *K* = 2.

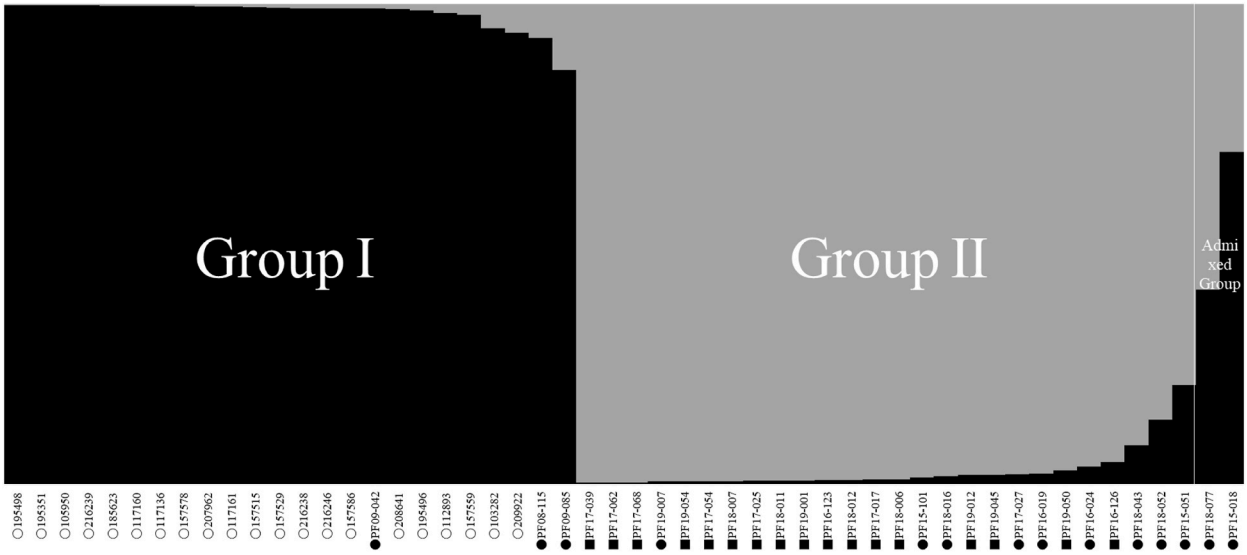


Fig. 3. Population structure of the 52 accessions of the TCWTPC collected from South Korea based on 28 *Perilla* SSR primer sets for $K = 2$. ○: accessions of CF, ●: accessions of WF, ■: accessions of WC.

contained 19 accessions, which were two WF accessions and 17 WC accessions. Group III included seven WF accessions. Among all the TCWTPC accessions, seven WF accessions were located in the out group (Fig. 4).

3.3. AMA of SSR markers and morphological characteristics among accessions of TCWTPC collected from South Korea

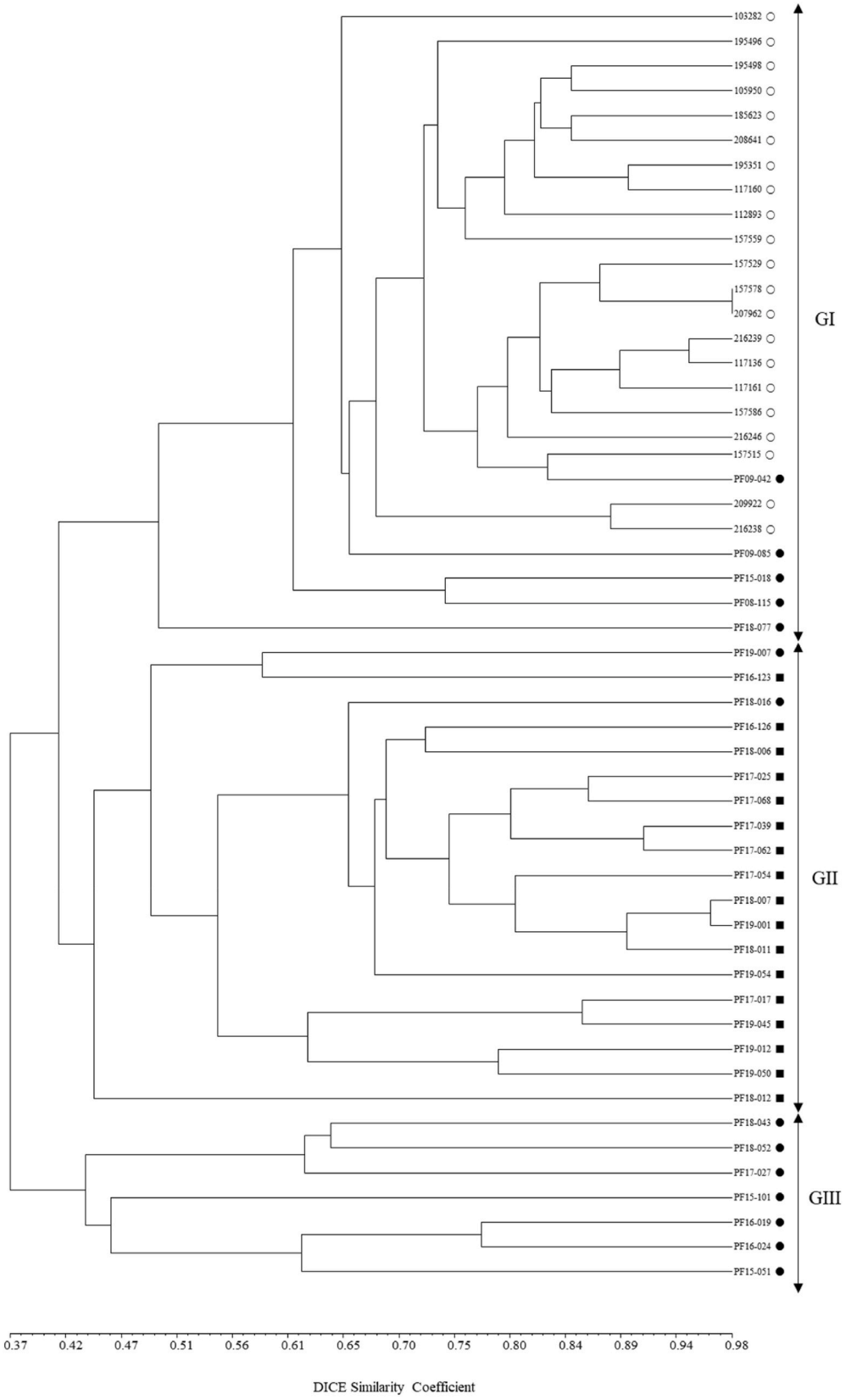
According to the survey results (Supplementary data, Table S4) of the one quantitative and 10 qualitative characteristics of the 52 accessions of the TCWTPC by Fu et al. [14], morphological differences between CF and WC were shown by most of the qualitative characteristics, namely color of leaf adaxial side (QL1), color of leaf abaxial side (QL2), stem color (QL4), flower color (QL5), plant fragrance (QL6), seed size (QL9), and seed hardness (QL10). Meanwhile, between CF and WF, the only morphological differences were the seed characteristics *w*eight per 100 seeds (QN1), QL9, and QL10; and between WF and WC, the only morphological differences were shown by QL1, QL2, and QL5 (Supplementary data, Table S4). In addition, among the 11 morphological characteristics, most of the leaf and seed characteristics showed statistically significant correlations among the 52 accessions of the TCWTPC (Supplementary data, Table S5). Therefore, in our study, we performed AMA using 28 *Perilla* SSR markers and the seven phenotypic characteristics (QN1, QL1, QL2, QL4, QL5, QL9, and QL10) that clearly showed morphological differences between and within accessions of the TCWTPC (Supplementary data, Tables S4 and S5). In the AMA of the 28 novel *Perilla* SSR markers and seven phenotypic characteristics, we identified significant MTAs through the TASSEL statistical program. From the results, we detected 132 MTAs containing 24 novel *Perilla* SSR markers associated with the seven phenotypic characteristics using a Q GLM at a significance level of $P \leq 0.05$ (Supplementary data, Table S6).

However, to avoid false positives, in our study we used Q + K MLM, and three *Perilla* SSR markers were associated with QL2, QL4, QL9, and QL10 ($P \leq 0.05$ or $P \leq 0.01$) (Table 2). Finally we confirmed five overlapping significant MTAs between Q GLM and Q + K MLM at a significance level of $P \leq 0.05$ or $P \leq 0.01$. Among the five significant MTAs, KNUPF134 was associated with QL4, KNUPF149 was associated with QL10, and KNUPF137 was associated with QL2, QL9, and QL10 (Table 2).

4. Discussion

4.1. Genetic diversity and morphological variation among TCWTPC accessions collected in South Korea

In *Perilla* crop, differences in morphological characteristics between and within accessions of the TCWTPC provide taxonomically important clues to understand the process of evolution or domestication of the TCWTPC [3,22,33]. As mentioned in the Introduction, in East Asia, *Perilla* crop has traditionally been categorized into two cultivated types (CF and CC) based on their respective uses, including as a seed oil or a vegetable (or herb medicinal) crop (Makino et al., 1961; [4,6]). In addition, the TCWTPC showed distinct morphological differences in leaf and seed characteristics (Fig. 1A and B): CF is characterized by large, soft seeds and green leaves, whereas CC is distinguished by small, hard seeds and leaves that can be either green or purple [4]. In East Asia, although wild species of *Perilla* crop are still unknown, Nitta and Ohnishi [21] and Lee and Ohnishi [4] reported the presence of weedy types (WF and WC) in *Perilla* crop for the first time. However, in the case of var. *frutescens*, Lee and Ohnishi [4,5] reported that there were no morphological differences and no AFLP markers between the cultivated var. *crispa* (which is mainly cultivated in Japan) and the weedy var. *crispa* (which grows wild in South Korea). For this reason, Lee and Ohnishi [4,5] suggested that CC, which is mainly cultivated and used in



(caption on next page)

Fig. 4. UPGMA dendrogram of the 52 accessions of the TCWTPC collected from South Korea based on 28 *Perilla* primer sets. ○: accessions of CF, ●: accessions of WF, ■: accessions of WC.

Table 2
Information on significant *MTAs* markers using the GLM and MLM method for 52 accessions of the TCWTPC.

Trait	Marker	GLM	MLM
QL2	KNUPF137	**	*
QL4	KNUPF134	**	*
QL9	KNUPF137	**	**
QL10	KNUPF137	**	**
	KNUPF149	*	*

*P ≤ 0.05, **P ≤ 0.01.

Japan, has not yet differentiated sufficiently from wild *Perilla* species or WC. Therefore, to understand the differentiation process of the TCWTPC, research using morphological characteristics or molecular markers is considered necessary.

In a preceding investigation conducted by Fu et al. [14], an examination of morphological characteristics was carried out among 52 TCWTPC accessions gathered in South Korea. Additionally, our research team has reported on the development of novel *Perilla* SSR primer sets within the *Perilla* crop [13,20,22,26]. In this study, we attempted to generate molecular markers associated with phenotypic characteristics based on AMA using *Perilla* SSR primer sets for morphological traits showing morphological differences between and within accessions of the TCWTPC. In particular, the available genetic information, such as genetic maps and quantitative trait loci (QTL), is insufficient for *Perilla* crop breeding studies. Therefore, the construction of allelic groups corresponding to specific phenotypic traits is necessary for studies in AMA. AMA is easier and less expensive than QTL mapping analysis using F_2, NIL, and RIL populations and has been suggested as a preferred method for identifying genetic loci associated with the inheritance of quantitative or qualitative traits [34,35].

In order to develop phenotypic trait-associated markers that show morphological differences between the TCWTPC, DNA fingerprinting was performed in this study using novel *Perilla* SSR loci for 52 TCWTPC accessions. The results revealed that a total of 142 alleles were detected across 28 *Perilla* SSR loci in 52 accessions of the TCWTPC collected from South Korea (Table 1). Among the 28 *Perilla* SSR loci, which were recently published by Fu et al. [13] and used for analysis in this study, 11 *Perilla* SSR loci were being used

Table 3
Genetic variation obtained from each *Perilla* SSR locus among 52 accessions of the TCWTPC collected from South Korea.

SSR Loci	Cultivated var. *frutescens* (n = 21)				Weedy var. *frutescens* (n = 14)				Weedy var. crispa (n = 17)			
	Allele No	MAF	GD	PIC	Allele No	MAF	GD	PIC	Allele No	MAF	GD	PIC
KNUPF132	2	0.905	0.172	0.157	3	0.857	0.255	0.240	3	0.471	0.554	0.452
KNUPF133	3	0.714	0.444	0.398	3	0.643	0.500	0.427	2	0.941	0.111	0.105
KNUPF134	8	0.238	0.821	0.797	8	0.286	0.837	0.818	3	0.647	0.512	0.453
KNUPF135	1	1.000	0.000	0.000	2	0.571	0.490	0.370	2	0.882	0.208	0.186
KNUPF136	1	1.000	0.000	0.000	2	0.786	0.337	0.280	3	0.588	0.554	0.482
KNUPF137	5	0.429	0.680	0.625	7	0.429	0.745	0.716	2	0.941	0.111	0.105
KNUPF138	1	1.000	0.000	0.000	3	0.429	0.612	0.530	2	0.941	0.111	0.105
KNUPF139	4	0.714	0.463	0.434	4	0.571	0.602	0.553	2	0.647	0.457	0.352
KNUPF140	1	1.000	0.000	0.000	2	0.786	0.337	0.280	2	0.882	0.208	0.186
KNUPF141	4	0.524	0.630	0.574	3	0.500	0.561	0.465	2	0.882	0.208	0.186
KNUPF142	2	0.905	0.172	0.157	3	0.857	0.255	0.240	1	1.000	0.000	0.000
KNUPF143	1	1.000	0.000	0.000	5	0.571	0.622	0.587	4	0.706	0.471	0.439
KNUPF144	3	0.905	0.177	0.169	5	0.500	0.673	0.632	3	0.706	0.443	0.384
KNUPF145	4	0.857	0.259	0.248	3	0.500	0.561	0.465	2	0.882	0.208	0.186
KNUPF146	3	0.571	0.553	0.473	3	0.571	0.582	0.517	2	0.824	0.291	0.248
KNUPF147	4	0.524	0.621	0.560	5	0.286	0.765	0.726	2	0.941	0.111	0.105
KNUPF148	6	0.381	0.757	0.723	10	0.143	0.888	0.877	6	0.353	0.775	0.744
KNUPF149	2	0.905	0.172	0.157	3	0.571	0.582	0.517	3	0.471	0.623	0.546
KNUPF150	2	0.952	0.091	0.087	4	0.500	0.643	0.585	4	0.529	0.616	0.555
KNUPF151	4	0.762	0.395	0.365	3	0.429	0.643	0.567	5	0.471	0.699	0.660
KNUPF152	4	0.857	0.259	0.248	3	0.643	0.520	0.464	3	0.765	0.381	0.340
KNUPF154	1	1.000	0.000	0.000	3	0.786	0.357	0.325	3	0.765	0.381	0.340
KNUPF155	2	0.952	0.091	0.087	4	0.429	0.622	0.547	3	0.824	0.304	0.281
KNUPF156	2	0.952	0.091	0.087	2	0.857	0.245	0.215	2	0.588	0.484	0.367
KNUPF157	4	0.476	0.662	0.607	4	0.643	0.541	0.502	4	0.471	0.678	0.628
KNUPF158	1	1.000	0.000	0.000	3	0.429	0.612	0.530	4	0.588	0.588	0.540
KNUPF159	2	0.952	0.091	0.087	5	0.571	0.612	0.571	2	0.647	0.457	0.352
KNUPF160	3	0.810	0.327	0.303	4	0.429	0.663	0.600	3	0.706	0.457	0.411
Mean	2.857	0.796	0.283	0.262	3.893	0.556	0.559	0.505	2.821	0.716	0.393	0.348

MAF: Major Allele Frequency, GD: Gene Diversity, PIC: Polymorphism Information Content.

for analysis for the first time. Therefore, the results of this study obtained using the novel 11 *Perilla* SSR primer sets provide useful genetic information on GD, phylogenetic relationships, PSA, and AMA of the 52 accessions of the TCWTPC collected in South Korea.

In this study, the average GD values of CF, WF, and WC accessions from South Korea were 0.283, 0.559, and 0.393, respectively (Table 3). Although the GD results between the TCWTPC in this study were analyzed using only 52 *Perilla* accessions, the results are similar to previous analyses conducted by Lee et al. [12] using AFLP markers and by Fu et al. [13] using *Perilla* SSR markers. That is, in these analyses also the GD observed in the WF accessions in South Korea is significantly greater than that observed in the CF and WC accessions.

In South Korea, the two weedy types (WF and WC) of *Perilla* crop are frequently found in habitats such as streams, roadsides, mountains, and around farmers' houses and fields [4,5,12,22]. Furthermore, in South Korea CF is the most cultivated and used compared with CC, which is mainly cultivated and used in Japan and is currently not cultivated at all in South Korea [4–6]. In the results of the morphological characteristics survey, the accessions of the TCWTPC showed clear differences in the leaf and seed-related characteristics QN1, QL1, QL2, QL4, QL5, QL9, and QL10 *(*Supplementary data, Tables S4 and S5). Therefore, these morphological characteristics are deemed valuable for distinguishing between accessions of the TCWTPC.

4.2. PSA, genetic relationships, and AMA between accessions of TCWTPC collected in South Korea

To elucidate the taxonomic and phylogenetic relationships between and within accessions of the TCWTPC collected from South Korea, the genetic relationships among the 52 accessions of the TCWTPC were analyzed using PSA. The 52 accessions of the TCWTPC were separated into two main groups at K = 2 (Fig. 3). Excluding exceptional accessions, the majority of TCWTPC accessions were distinctly differentiated in the PSA. That is, Group I comprised accessions of CF, and Group II included accessions of WF and WC. These results are similar to those of the phylogenetic relationship analysis using UPGMA, as shown in Fig. 4, except for outstanding accessions. However, among the accessions of WF in the PSA and phylogenetic relationship analysis, there were some accessions that were located between CF and WC. In a previous study using RAPD markers by Nitta and Ohnishi [21], it was reported that accessions of WF and WC had probably derived from hybrids between the TCWTPC.

Furthermore, in the results of PSA (Fig. 3) and a phylogenetic tree (Fig. 4), some accessions of WF were located close to CF. In our study, three accessions (PF08-115, PF09-042, PF09-085) of WF were considered as probable escape type from CF. This phenomenon was reported by Lee and Ohnishi [5] and Sa et al. [22]. In the case of WF, several accessions were closely related with accessions of WC, and also seven accessions of WF were at the outermost part of the phylogenetic tree (Fig. 4). Therefore, it is very difficult to understand the taxonomic positions of these weedy samples between and within accessions of the TCWTPC in the phylogenetic tree, and it is possible that the accessions of WF originate from several different routes, as previously noted by Lee and Ohnishi [5] and Sa et al. [22, 24].

Several studies have been conducted to differentiate between and within accessions of the TCWTPC. Ha et al. [25] and Kim et al. [20] specifically reported on AMA for identifying SSR markers linked to seed characteristics (seed germination rate, seed hardness, seed size) and leaf characteristics (leaf color and shape) in the TCWTPC and found several *Perilla* SSR markers linked to seed and leaf related traits. In our study, in order to conduct AMA, we used genotypes of the novel 28 *Perilla* SSR markers and phenotypic characteristics for seven morphological traits related to seed and leaf traits in the 52 accessions of the TCWTPC. Also, we used the Q + K MLM method to avoid false positives for the 132 MTAs related to the seven phenotypic characteristics shown by the Q GLM method *(S*upplementary data, Table S6). According to the AMA, we confirmed five overlapping significant MTA markers between Q GLM and Q + K MLM, with a significance level of $P \leq 0.05$ or $P \leq 0.01$, as follows: KNUPF134 was related to QL4; KNUPF137 was related to three characteristics, namely QL2, QL9, and QL10; and KNUPF149 was associated with QL10 (Table 2). According to Fu et al. [14], combinations between QL2 and QL9 (0.590**), between QL2 and QL10 (−0.602**), and between QL9 and QL10 (−0.961**) exhibited stronger positive and negative correlation coefficients compared with other combinations [14]. Hence, the *Perilla* SSR marker KNUPF137 linked to leaf and seed characteristics is thought to be useful for identifying accessions of the TCWTPC.

In previous studies, many researchers reported *Perilla* SSR primer sets related to leaf and seed traits in accessions of the TCWTPC [20,25]. In the current study, by using the novel *Perilla* SSR primer sets developed by Fu et al. [13], we performed AMA for seven phenotypic characteristics related to leaf and seed traits among 52 accessions of the *TCWTPC* using the 28 novel *Perilla* SSR primer sets. However, in *Perilla* species, the available *Perilla* SSR primer sets remain inadequate for comprehensive studies on GD, PIC, phylogenetic relationships, PSA, and AMA in the accessions of the TCWTPC.

Therefore, the genetic information of the 11 novel *Perilla* SSR primer sets used for the first time in this study will offer valuable insights for genetic analysis of the TCWTPC in the future. Moreover, three *Perilla* SSR markers (KNUPF134, KNUPF137, KNUPF149) related to leaf and seed characteristics are regarded as highly valuable selection markers for phenotypic traits linked to leaf and seed characteristics in the TCWTPC. Therefore, the novel *Perilla* SSR primer sets outlined in this study will be useful for PSA and AMA in assessing GD, PIC, and phylogenetic and genetic relationships between and within accessions of the TCWTPC and will be useful for genetic mapping and cultivar development in *Perilla* crop breeding programs such as MAS.

5. Conclusion

The objective of this study was to develop molecular markers linked to morphological traits that show differences between accessions of the TCWTPC using novel *Perilla* SSR primer sets. In this research study, a total of 28 *Perilla* SSR primer sets exhibited polymorphism, resulting in a total allele count of 142 among the 52 accessions of the TCWTPC. The accessions of CF, WF, and WC collected in South Korea showed mean GD values of 0.283, 0.559, and 0.393, respectively. Therefore, in South Korea, *the* accessions of

WF showed the highest genetic variation, while *the* accessions of CF showed the lowest genetic variation. During the analysis of population structure and the construction of a phylogenetic tree, the 52 accessions of the TCWTPC were classified into three major groups. While the majority of TCWTPC accessions were relatively clearly distinguished, certain accessions belonging to the two weedy types (WF and WC) of *Perilla* crop were not clearly distinguished by *Perilla* SSR markers. By using AMA of 28 *Perilla* SSR primer sets and seven morphological characteristics in the 52 accessions of the TCWTPC, we detected three *Perilla* SSR markers (KNUPF134, KNUPF137, KNUPF149) associated with leaf and seed characteristics, and these are expected to be highly valuable as selection markers for traits related to plant and seed characteristics between accessions of the TCWTPC. Also, they are considered as valuable molecular markers for distinguishing between and within accessions of the TCWTPC. Hence, the novel *Perilla* SSR primer sets used in this study should be useful for PSA and AMA in assessing GD, PIC, and phylogenetic relationships between and within accessions of the TCWTPC as well as being useful for genetic mapping and cultivar development in *Perilla* crop breeding programs such as MAS.

Additional information

No additional data and information is available for this paper.

Funding

This study was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2022R1F1A1063300) and the Cooperative Research Program for Agriculture Science & Technology Development (project no. PJ01422701), Rural Development Administration, Republic of Korea.

Data availability statement

The authors declare the date is included in the article and no additional data available.

CRediT authorship contribution statement

So Jung Jang: Formal analysis. **Kyu Jin Sa:** Writing – review & editing, Writing – original draft. **Zhen Yu Fu:** Investigation. **Ju Kyong Lee:** Writing – review & editing, Conceptualization.

Declaration of competing interest

We declare that this manuscript is original, has not been published and is not under consideration for publication elsewhere and there are no conflicts of interest to disclose.

Acknowledgments

We are thankful to all the institutions and people who helped with this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e26720.

References

[1] T. Makino, Makino's New Illustrated FLora of Japan, Hokuryukan, Tokyo, 1961 (in Japanese).
[2] H.L. Li, The vegetables of ancient China, Econ. Bot. 23 (1969) 235–260. http://www.jstor.org/stable/4253062.
[3] M. Nitta, Origin of *Perilla* Crops and Their Weedy Type, Ph.D. Thesis, Kyoto University, Kyoto, Japan, 2001, p. 78.
[4] J.K. Lee, O. Ohnishi, Geographical differentiation of morphological characters among *Perilla* crops and their weedy types in East Asia, Breeding science 51 (2001) 247–255, https://doi.org/10.1270/jsbbs.51.247.
[5] J.K. Lee, O. Ohnishi, Genetic relationships among cultivated types of *Perilla frutescens* and their weedy types in East Asia revealed by AFLP markers, Genet. Resour. Crop Evol. 50 (2003) 65–74, https://doi.org/10.1023/A:1022951002271.
[6] M. Nitta, J.K. Lee, O. Ohnishi, Asian Perilla crops and their weedy forms: their cultivation, utilization and genetic relationships, Econ. Bot. 57 (2003) 245–253. https://www.jstor.org/stable/4256682.
[7] M. Nitta, J.K. Lee, C.W. Kang, M. Katsuta, S. Yasumoto, D. Liu, T. Nagamine, O. Ohnishi, The distribution of *Perilla* species, Genet. Resour. Crop Evol. 52 (2005) 797–804.
[8] H. Park, K.J. Sa, D.Y. Hyun, S. Lee, J.K. Lee, Identifying SSR markers related to seed fatty acid content in *Perilla* crop (*Perilla frutescens* L.), Plants 10 (2021) 1404, https://doi.org/10.3390/plants10071404.
[9] H. Park, K.J. Sa, S. Lee, J.K. Lee, Genetic variation of seed oil characteristics in native Korean germplasm of *Perilla* crop (*Perilla frutescens* L.) using SSR markers, Genes & Genomics 44 (2022) 1159–1170, https://doi.org/10.1007/s13258-022-01289-y.
[10] S.M. Talbott, K. Hughes, *Perilla* Seed Oil (*Perilla frutescens*). The Health Professional's Guide to Dietary Supplements, Lippincott Williams & Wil. ins, Baltimore MD, 2006, pp. 258–260.

[11] M. Asif, Health effects of omega-3,6,9 fatty acids: *Perilla frutescens* is a good example of plant oils, Oriental Pharmacy and Experimental Medicine 11 (2011) 51–59, https://doi.org/10.1007/s13596-011-0002-x.
[12] J.K. Lee, M. Nitta, N.S. Kim, C.H. Park, K.M. Yoon, Y.B. Shin, O. Ohnishi, Genetic diversity of *Perilla* and related weedy types in Korea determined by AFLP analyses, Crop Sci. 42 (2002) 2161–2166, https://doi.org/10.2135/cropsci2002.2161.
[13] Z.Y. Fu, K.J. Sa, H. Park, S.J. Jang, Y.J. Kim, J.K. Lee, Utilization of novel *Perilla* SSR markers to assess the genetic diversity of native *Perilla* germplasm accessions collected from South Korea, Plants 11 (2022) 2974, https://doi.org/10.3390/plants11212974.
[14] Z.Y. Fu, K.J. Sa, S.K. Lee, J.K. Lee, Morphological variation of cultivated and weedy types of *Perilla* crop collected from South Korea, Korean journal of medicinal crop science 31 (2023) 1–11, https://doi.org/10.7740/kjcs.2013.58.4.408.
[15] I. Nagai, On "shiso" and "egoma", Agric. Hortic. 10 (1935) 2265–2273.
[16] Y. Yamane, Cytogenetic studies in *Perilla* and *Coleus*. I. Chromosome numbers, Jpn. J. Genet. 25 (1950) 220 (in Japanese).
[17] G. Honda, Y. Koezuka, M. Tabata, Genetic studies of fruit color and hardness in *Perilla frutescens*, Jpn. J. Breed. 40 (1990) 469–474, https://doi.org/10.1270/jsbbs1951.40.469.
[18] G. Honda, A. Yuba, T. Kojima, M. Tabata, Chemotaxonomic and cytogenetic studies on *Perilla frutescens* var. *citiodora* ('Lemon Egoma'), Natural medicines 48 (1994) 185–190.
[19] S.E. Lim, K.J. Sa, J.K. Lee, Bulk segregant analysis identifies SSR markers associated with leaf-and seed related traits in *Perilla* crop (*Perilla frutescens* L.), Genes & Genomics 43 (2021) 323–332, https://doi.org/10.1007/s13258-021-01056-5.
[20] J.Y. Kim, K.J. Sa, Y.J. Ha, J.K. Lee, Genetic variation and association mapping in F_2 population of *Perilla* crop (*Perilla frutescens* L.) using new developed *Perilla* SSR markers, Euphytica 217 (2021) 135, https://doi.org/10.1007/s10681-021-02867-z.
[21] M. Nitta, O. Ohnishi, Genetic relationships among two *Perilla* crop, shiso and egoma, and the weedy type revealed by RAPD markers, Genes Genet 74 (1999) 43–48, https://doi.org/10.1266/ggs.74.43.
[22] K.J. Sa, I.K. Choi, K.C. Park, J.K. Lee, Genetic diversity and population structure among accessions of *Perilla frutescens* (L.) Britton in East Asia using new developed microsatellite markers, Genes & genomics 40 (2018) 1319–1329, https://doi.org/10.1007/s13258-018-0727-8.
[23] Y.J. Park, J.K. Lee, N.S. Kim, Simple sequence repeat polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops, Molecules 14 (2009) 4546–4569, https://doi.org/10.3390/molecules14114546.
[24] K.J. Sa, S.H. Choi, M. Ueno, K.C. Park, Y.J. Park, K.H. Ma, J.K. Lee, Identification of genetic variations of cultivated and weedy types of *Perilla* species in Korea and Japan using morphological and SSR markers, Genes & Genomics 35 (2013) 649–659, https://doi.org/10.1007/s13258-013-0117-1.
[25] Y.J. Ha, K.J. Sa, J.K. Lee, Identifying SSR markers associated with seed characteristics in *Perilla* (L.), Physiol. Mol. Biol. Plants 27 (2021) 93–105, https://doi.org/10.1007/s12298-021-00933-3.
[26] K.J. Sa, S.E. Lim, I.K. Choi, K.C. Park, J.K. Lee, Development and characterization of new microsatellite markers for *Perilla frutescens* (L.) Britton, Am. J. Plant Sci. 10 (2019) 1623–1630, https://doi.org/10.4236/ajps.2019.109115.
[27] K. Liu, S. Muse, PowerMarker: an integrated analysis environment for genetic marker analysis, Bioinformatics 21 (2005) 2128–2129, https://doi.org/10.1093/bioinformatics/bti282.
[28] L.R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (1945) 297–302.
[29] F.J. Rohlf, NTSYS-Pc: Numerical Taxonomy and Multivariate Analysis System Version 2.1 Exeter Software, Applied Biostatics Inc., New York, 2000.
[30] J.K. Pritchard, W. Wen, Documentation for STRUCTURE Software: Version 2, 2003. http://www.pritc.h.bsd.uchicago.Edu.
[31] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol. 14 (2005) 2611–2620, https://doi.org/10.1111/j.1365-294X.2005.02553.x.
[32] P.J. Bradbury, Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S. Buckler, TASSEL: software for association mapping of complex traits in diverse samples, Bioinformatics 23 (2007) 2633–2635, https://doi.org/10.1093/bioinformatics/btm308.
[33] S.J. Ma, J.K. Lee, Morphological variation of two cultivated types of *Perilla* crop from different areas of China, Horticultural Science and Technology 35 (2017) 510–522, https://doi.org/10.12972/kjhst.20170054.
[34] S.A. Quarrie, V. Lazić-Jančić, D. Kovačević, A. Steed, S. Pekić, Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize, J. Exp. Bot. 50 (1999) 1299–1306, https://doi.org/10.1093/jxb/50.337.1299.
[35] P. Wenzl, H. Raman, J. Wang, M. Zhou, E. Huttner, A. Kilian, A DArT platform for quantitative bulked segregant analysis, BMC Genom. 8 (2007) 1–10, https://doi.org/10.1186/1471-2164-8-196.