Phytotaxa 626 (3): 170–180
https://www.mapress.com/pt/

Article

ISSN 1179-3155 (print edition)
PHYTOTAXA
ISSN 1179-3163 (online edition)

https://doi.org/10.11646/phytotaxa.626.3.3

Development of SSR markers identification system for *Carex* L. based on RAD sequencing

RONGFENG CUI[1,6*,§], BO WEI[2,7*], FANG LIANG[1,8], ZIJING LI[1,9], AIXIANG DONG[1,10], YAN ZHOU[1,11], JI NAIZHE[1,12], DONGYUN LIU[3,13], QING WANG[4,14], FANG FU[5,15], LIPING SUN[1,16], HELAN QIN[1,17], CHAO YUAN[1,18], LINA SONG[1,19], TAO WANG[1,20], QIAN WANG[1,21] & JIN SHAN[1,22]

[1]*Beijing Key Laboratory of Greening Plants Breeding, Beijing Institute of Architecture, 100102, Beijing, P.R. China.*
[2]*National Key Laboratory of Wheat Improvement, Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory of Advanced Agricultural Sciences in Weifang, 261325, Shandong, P.R. China.*
[3]*College of Landscape Architecture and Tourism, Hebei Agricultural University, 071000, Baoding, Hebei, P.R. China.*
[4]*Beijing Radiation Center, Beijing Academy of Science and Technology, 102300, Beijing, P.R. China.*
[5]*Science and technology exploitation and examination base of Mentougou District, 102300, Beijing, P.R. China.*
[6] *cuirongfeng@163.com; https://orcid.org/0000-0003-0632-5474*
[7] *bo.wei@pku-iaas.edu.cn; https://orcid.org/0009-0004-6879-6065*
[8] *540926864@qq.com; https://orcid.org/0009-0006-0872-5278*
[9] *lizijing8341@126.com; https://orcid.org/0009-0002-4641-0702*
[10] *gooddax@sina.com; https://orcid.org/0009-0000-1786-7506*
[11] *zhouy661@163.com; https://orcid.org/0000-0002-7353-122*
[12] *jinaizhe1016@qq.com; https://orcid.org/0000-0002-7437-3913*
[13] *liudongyun0505@163.com; https://orcid.org/0000-0003-2925-3549*
[14] *wangqing80@126.com; https://orcid.org/0000-0002-3118-6368*
[15] *foolfun78@163.com; https://orcid.org/0009-0005-8431-5804*
[16] *pingsun1984@126.com; https://orcid.org/0009-0005-6023-9560*
[17] *qinhelan71@yahoo.com.cn; https://orcid.org/0009-0002-0079-3824*
[18] *ychy512@126.com; https://orcid.org/0009-0002-9680-9711*
[19] *272975643@qq.com; https://orcid.org/0009-0008-4994-4789*
[20] *44125583@qq.com; https://orcid.org/0009-0002-8387-3669*
[21] *wangqian-200@163.com; https://orcid.org/0009-0000-6932-882X*
[22] *3068424610@qq.com; https://orcid.org/0009-0005-2379-0986*
** Corresponding author*

Abstract

In this study, we designed and evaluated an efficient set of simple sequence repeat (SSR) markers from restriction site-associated DNA (RAD) sequencing data of four widely utilized *Carex* L. accession across Northern China, especially in Beijing. Based on their genomic sequencing data, we developed 400 SSR markers and evaluated their amplification specificities among which a total of 17 SSR markers were identified as a core set of molecular markers for efficient assessment of diverse *Carex* L. accessions. Using this molecular identification system, we classified 26 *Carex* L. accessions into 3 genetically distinct groups. The establishment of a molecular identification system based on a core set of 17 SSR markers provides an essential basis for evaluating genetic relationships among *Carex* L. accessions along with a wealth of genetic resources for marker-assisted selection and breeding of *Carex* L..

Key words: *Carex* L., Restriction site-associated DNA sequencing, SNP, SSR, Molecular identification

Introduction

The genus *Carex* L. contains at least 2,000 grass species belonging to the *Cyperaceae* family and are generally considered as sedges (Jiménez-Mejías *et al.* 2016; Martín-Bravo *et al.* 2019; Reznicek 1990). These grass species are mainly distributed in temperate, as well as cold regions, and contribute significantly in turf management, forage production, and ecological preservation (Martín-Bravo *et al.* 2019). However, only few studies have focused on genetic marker development for molecular breeding of *Carex* L. accessions. Previous studies were mainly limited to the physiological

investigation (Teng *et al.* 2019), geographical distribution (Jiménez-Mejías *et al.* 2016; Martín-Bravo *et al.* 2019), and stress-resistance evaluation (Li *et al.* 2018). Consequently, studies on the germplasm evaluation and molecular markers based marker-assisted breeding of this ecologically important genus lag far behind. Thus, *Carex* L. breeding urgently needs a theoretical basis at the molecular level and further exploration of its genetic resources.

Development of stable molecular markers based on genomic sequences are essential for the analyses of genetic divergence, quantitative trait loci (QTLs) mapping, and marker-assisted breeding (Vidak *et al.* 2021). Studies involving *Carex* L. have utilized polymorphisms of nuclear ribosomal DNA (nrDNA) internal transcribed spacers (ITS), the nrDNA external transcribed spacer (ETS), and the chloroplast DNA (cpDNA) maturase K (matK) as molecular markers in phylogenetic analyses (Hendrichs *et al.* 2004b; Hendrichs *et al.* 2004a; Jiménez-Mejías *et al.* 2016; Bruce *et al.* 2012; Martín-Bravo *et al.* 2019). Moreover, an inter-simple sequence repeats-polymerase chain reaction (ISSR-PCR) system is developed and optimized in *Carex* L. (Ning *et al.* 2014a; Ning *et al.* 2014b). However, the number of markers are relatively lower for establishing a genome-wide screening system and/or molecular breeding (Ning *et al.* 2014a; Ning *et al.* 2014b). Restriction site-associated DNA (RAD) sequencing is a useful strategy for genome sequence assembly and molecular markers development, especially for species with a shortage of genomic sequences (Wang *et al.* 2012). To date, RAD sequencing has been successfully used for various plant species, such as *Pinctada fucata* (Li and He 2014), *Pseudotaxus chienii*, *Trifolium pretense*, and *Vitis vinifera* (Wang *et al.* 2012). These studies highlight the power of RAD sequencing in genomic sequence production.

Carex L. is a perennial grass with wide distribution in China that occurs mainly beneath tree crowns, as it is highly shade tolerant (Xue *et al.* 2005; Shi 2007). As afforestation in China accelerates, *Carex* L. is expected to be planted more widely (Ma *et al.* 2001). However, to date, the genetic resources of *Carex* L. have not been properly exploited, hampering the progress of *Carex* L. breeding. In this study, we aimed to establish a molecular identification system based on genome-wide SSR markers for *Carex* L. We sequenced four widely used *Carex* L. accessions using a RAD sequencing strategy. Utilizing a set of 17 selected primers, we designed and evaluated SSRs by amplifying 26 *Carex* L. accessions mainly distributed in North China and having greater potential ornamental applications. Our results provide a novel molecular identification platform for genetic diversity analysis, molecular phylogenetic analysis, germplasm selection, and finally the molecular breeding of *Carex* L. accessions.

Material and Methods

Plant materials

Twenty-six *Carex* L. accessions were mainly distributed in North China, which were collected by the Beijing Institute of Landscape Architecture and used to evaluate the SSR markers (Table S1). Of these, four promising accessions (each with three biological replications)—Laoyv-2 (C11), Qinglv-2 (C12), Pizhen-2 (C13), and Jiao-2 (C15)—which were the most widely used in reproduction, market application, landscaping, and environmental engineering in North China (especially in Beijing), were used as plant materials for RAD sequencing (Figure 1). The SSR markers designed after RAD sequencing were evaluated on 26 *Carex* L. accessions for the development of molecular identification system.

RAD sequencing

A restriction site-associated DNA sequencing strategy was used for the genome sequencing. The RAD tags were obtained by digesting the total 1 ug genomic DNAs of the *Carex* L. accessions using the restriction enzyme *EcoR* I for 30 minutes at 37°C. The digested DNA fragments were purified and screened to ensure that fragments had a length of 300–500 bp, and then ligated with Illumina compatible dual indexed-adaptors (NEBNext Multiplex adaptors, Japan) according to the manufacturer instructions. The library was sequenced using an Illumina Hiseq4000 platform (BGI, Shenzhen, China) with Pair-End 150 bp mode. Three biological replicates were used for each accession in RAD sequencing.

RAD data analysis

Fastp software was used to remove low-quality raw reads (contain adapter-base, or valid read length < 18 bp, or reads average quality value < 20) (Chen *et al.* 2018). Finally, a total of 16.31 Gb clean data were kept for further analysis.

SNP sites were found mainly as following procedures: firstly, stacks command each sample clean reads were conservatively *de novo* assembled and clustered to generate putative loci by "ustacks" command in Stack software (version 2.5.4) (Rochette and Catchen 2017). Then, each locus was examined and filtered to generate cluster locus

by “cstacks” command (with the number of mismatches allowed between any two alleles of the population, N=3). “cstacks” command was used to match each population against the locus, to get variant sites for each individual.

The SSR sites in the locus were predicted by MISA software (with default parameters) (Beier *et al.* 2017). The target locus was evaluated and primers were designed using PRIMER3 software (primer length 18–28 nucleotides, PCR product size 200–300 bp) (Kõressaar *et al.* 2018). The sequences of each original primer set were listed in Table S2.

DNA preparation and PCR detection

Genomic DNAs of all of the plant materials were extracted from young leaves of *Carex* L. seedlings according to a modified cetyltrimethylammonium bromide (CTAB) method (Murray and Thompson 1980). The PCR reactions were performed in a 20 µl system comprising 0.3 µl (20 µM) of each primer, 0.4 µl (25 mM) of dNTPs, 10 µl 2X GC Buffer II, 0.2 µl LA Taq (Takara, Japan), 1.0 µl (50 ng/µl) of template DNA, and 4.8 µl ddH_2O. The PCR program consisted of an initial denaturation of 5 min at 94℃; 35 cycles of 30 s at 94℃, 30 s at 54℃, and 2 min at 72℃; and a final extension of 10 min at 72℃. The PCR products were separated by electrophoresis in 1.5% agarose gels and 8% polyacrylamide gels using ethidium bromide (Tiangen, China). The gel images were obtained using a UV spectrometer (BioRad, USA).

Data analysis

Quantity one 1-D software (version29.0, Bio-Rad) was used to illustrate the schematic pictures of polyacrylamide gel electrophoresis results. According to the size of SSR amplifications, products with same size were labeled as 1, and otherwise as 0. The genetic patterns of SSR markers for all accessions were compiled in a data matrix and further analyzed using NTSYS software to calculate the similarity coefficient (Dice coefficient). UPGMA clustering analysis was performed to calculate the genetic coefficients. Popgene32 software was used to analyze the genetic diversity of the tested plant materials.

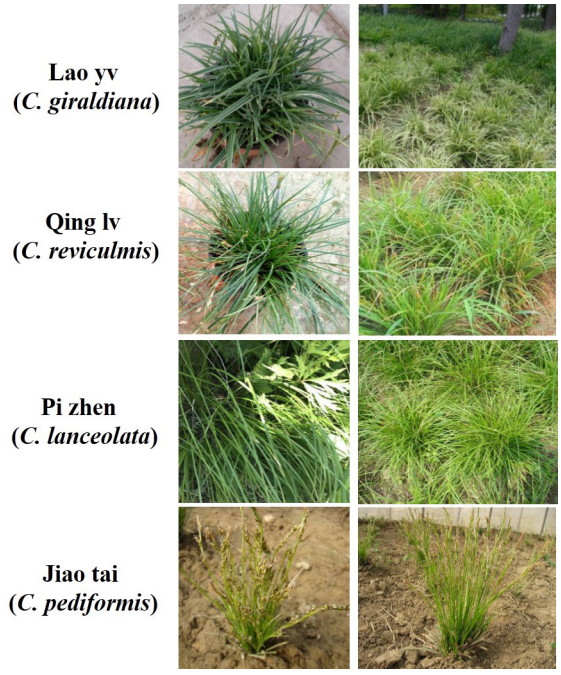

FIGURE 1. The *Carex* L. accessions used for DNA sequencing in this research.

Results

Sequence analysis of Carex L. by RAD sequencing

Through restriction site-associated DNA sequencing (RAD-seq), a total of 16.31 Gb of sequence data was obtained (SRA accession number: PRJNA656443), with an average of 1.34 Gb per sample. The proportion of base quality Q30 reached 92.62%. The average number of clean reads across the 12 samples was 4,388,585 (Table 1).

TABLE 1. RAD sequencing statistics.

Sample ID	Clean Reads	Clean Base (bp)	GC (%)	Q30(%)	TagNumber	Accession Number in NCBI
Jiaotai-1	4,399,020	1,308,286,295	43.94	93.49	308,059	SAMN15785106
Jiaotai-2	4,041,271	1,201,911,090	43.95	93.43	300,628	SAMN15785107
Jiaotai-3	4,180,933	1,239,084,143	43.97	93.46	305,670	SAMN15785108
Laoyv-1	4,533,212	1,338,944,443	36.54	92.75	182,994	SAMN15785109
Laoyv-2	4,161,007	1,229,088,285	36.42	92.78	174,267	SAMN15785110
Laoyv-3	4,735,957	1,849,387,980	35.5	90.83	209,505	SAMN15785111
Pizhen-1	3,978,999	1,179,160,585	36.55	92.89	300,901	SAMN15785112
Pizhen-2	4,232,541	1,245,735,165	36.87	92.87	314,835	SAMN15785113
Pizhen-3	4,795,877	1,411,540,613	36.64	92.64	324,683	SAMN15785114
Qinglv-1	4,376,454	1,292,480,174	37.50	92.81	276,261	SAMN15785115
Qinglv-2	4,495,448	1,327,856,714	37.46	92.92	216,780	SAMN15785116
Qinglv-3	4,732,310	1,407,311,143	37.27	92.87	222,893	SAMN15785117

SNP statistics

Upon further analysis of polymorphic tags, high-quality SNP loci were obtained. For each *Carex* L. accession, the number of SNPs were comparable among three biological replications. Among four accessions, Jiaotai-2 (average 2,756.33±165.28), Pizhen-2 (average 2,806.00±171.58), and Qinglv-2 (average 2,637.33±463.80) had comparable numbers of SNPs; however, Laoyv-2 had considerably fewer SNPs (760.00±78.54) (Table 2). Interestingly, we found that the *Carex* L. genome contains a large heterozygous chromosome region based on the SNP analyses. In Jiaotai-2, the heterozygous SNPs accounted for 97% of the total SNPs, whereas in Pizhen-2 and Qinglv-2, the percentage of heterozygous SNPs were 75% and 60%, respectively. Similarly, Laoyv-2 had the lowest rate of heterozygous SNPs (33.15%) (Table 2).

Establishment of an SSR identification system for Carex L.

To establish the molecular identification system, we first designed SSR markers based on the clusters present in at least five sequencing samples (Figure 1). Then, we selected 400 pairs of SSR markers to detect whether they could produce the expected amplified fragments from *Carex* L. accessions (Table S3). Three accessions, Laoyv-1 (C1), Cyan-1 (C2), and Jiaotai-1 (C4), were randomly selected as small sample of the 26 accessions to screen the SSR markers (Figure S1). The results showed that out of the 400 SSR markers, 18 pairs produced clear bands with the expected size and these were further verified by DNA sequencing (Figure S1).

Then, we detected the polymorphism of 26 *Carex* L. accessions using the 18 SSR markers by polyacrylamide gel electrophoresis (Supplementary file 1). Several bands were amplified by individual primers of individual samples, and some of the primers could not amplify the products. Then, based on the polyacrylamide gel electrophoresis, the allele patterns were transformed into schematic pictures (Supplementary file 2). Further, a 0-1 data matrix was obtained by reading the bands of each pair of primers (Table S3). According to the "0" and "1" matrix data, the polymorphisms of SSR markers were analyzed by Popgene32 software. Allelic loci analyses indicated that there were 193 alleles among 26 accessions based on 18 SSR markers, and the average number of alleles was 10.72. Marker P40 was eliminated due to low amplification specificity. Therefore, a total of 173 alleles (NA) were obtained by 17 pairs of primers, with an average of 10.18 for each pair of primers (Table 3). The numbers of amplified alleles were quite different among different primers, with the highest value being 16 (P35) and the lowest being 5 (P261). Furthermore, P35 possessed the highest PIC (0.9211) and distinguished 25 *Carex* L. accessions, except of C13 and C14 (Table S4). Therefore, the 17 polymorphic SSR primer pairs were considered as a molecular identification system for the *Carex* L. accessions, and P35 was the most efficient primer pair because it possessed higest polymorphism information content (Table 4).

TABLE 2. The statistical analysis of SNP among the four *Carex* L. accessions.

ID Number	SNP Number	Transition Number	Transversion Number	Ti/Tv	Heterozygosity SNP Number	Homozygosity SNP Number	Heterozygosity SNP ratio	Homozygosity SNP ratio
Jiaotai-1	2525.00	1369.00	1156.00	1.18	2440.00	85.00	0.96	0.03
Jiaotai-2	2843.00	1570.00	1273.00	1.23	2749.00	94.00	0.97	0.03
Jiaotai-3	2901.00	1592.00	1309.00	1.22	2797.00	104.00	0.96	0.04
Average	2756.33±165.28	1510.33±100.34	1246±65.31	1.21±0.02	2662.00±158.20	94.33±7.76	0.97±0.00	0.03±0.00
Laoyv-1	701.00	435.00	266.00	1.64	248.00	453.00	0.35	0.65
Laoyv-2	708.00	435.00	273.00	1.59	244.00	464.00	0.34	0.66
Laoyv-3	871.00	519.00	352.00	1.47	258.00	613.00	0.30	0.70
Average	760.00±78.54	463±39.6	297.00±39.00	1.57±0.07	250.00±5.89	510.00±72.97	0.33±0.03	0.67±0.03
Pizhen-1	2622.00	1563.00	1059.00	1.48	2015.00	607.00	0.77	0.23
Pizhen-2	2761.00	1662.00	1099.00	1.51	2060.00	701.00	0.75	0.25
Pizhen-3	3035.00	1781.00	1254.00	1.42	2203.00	832.00	0.73	0.27
Average	2806.00±171.58	1668.67±89.12	1137.33±84.1	1.47±0.04	2092.67±80.15	713.33±92.27	0.75±0.02	0.25±0.02
Qinglv-1	3130.00	1829.00	1301.00	1.41	2082.00	1048.00	0.67	0.33
Qinglv-2	2016.00	1219.00	797.00	1.53	1012.00	1004.00	0.50	0.49
Qinglv-3	2766.00	1634.00	1132.00	1.44	1743.00	1023.00	0.63	0.37
Average	2637.33±463.8	1560.67±254.37	1076.67±209.44	1.46±0.05	1612.33±446.49	1025.00±18.02	0.60±0.07	0.40±0.07

TABLE 3. SSR primer sets developed in this study

Code	SSR type	Forward primer (up, 5'-3') Reverse primer (down, 5'-3')	Annealing Temperature	Product size (bp)
P35	(AG)9	TCAGATTCTAATTGATTCATTGCTTCA	56	119
		TCGTCGAAGAGAACAAGGTACC		
P55	(TCT)9	ACCACTTCAAATCTCTCCTCCAC	54	82
		TGGTACCAACTAAGCTCAACCA		
P73	(GA)9	CAGACCCGAGGTGAAAGAGG	55	89
		ACTCCGACAACGTTTACGGT		
P98	(TA)9	CAGCGCATAATATATATATTGTCCTGT	56	121
		TCTCAAACAAAATGATGATGGTGA		
P100	(CT)9	ACCAATCAGATCCATGATCTCCA	55	107
		TCCCATTTTTGTTGTTTTCTTCTTTT		
P116	(AG)9	AGGCGAAAAGAAAACCCCAA	53	93
		GGGTTTTGAGAGCTCAGAGAGA		
P119	(GA)9	TCACGGAAGGAAGGATTTTTATTTT	54	118
		GGATTGCTTGATTAACAGAGCCA		
P123	(TA)9	TGTTCCACTTGCCAATTGTCC	55	106
		TCTCTTTTGGTCCAATTGATTATGAA		
P128	(CT)9	ACATCTGCTGAAATCTCCATCCA	56	102
		GGAACCCTGTAATTAATCTTCATCCT		
P154	(TA)9	GCATAACTTCCCAACACGCT	56	98
		TCGATAGGCATGCATGGTGA		
P261	(AT)6	TCTTGAAAAGTTGAAATGCTTGACA	55	123
		ACCCCACATGAAGAAGGGAT		
P280	(TGTA)8	TCACATTTTATTGCAAAATACTCGCA	54	130
		ACGAAACCAACAACTTCTTTTGT		
P305	(AT)6	CCCGTTTTGCCCATCAGGTA	55	124
		CATCAAACAGAGGATATGAACATAGT		
P318	(TC)7	TCGGTTGATACCAAGTGTTCA	55	123
		TGCATCATCATCTCTCTCTCTTGG		
P337	(TA)8	TCTAGCTAGAATTATTGCGAAAGAAA	56	119
		ACCCTCCATCTTCAGTTGTGA		
P346	(AG)6	TCCGTCAAAACCCTAGTGCTC	54	119
		TCCGCCTCTTTCGCTTTCAT		
P356	(GA)6	TCAGTGAACTAGAGATAACAAATGGA	55	117
		TGAACATCAAAGCAAGCCCT		

TABLE 4. Evaluation of the 17 pairs of SSR developed in this study

Locus	Sample Size	na*	ne*	I*	PIC*
P35	52	16	13.52	2.674	0.9211
P55	40	12	7.1429	2.192	0.8459
P73	46	11	9.8879	2.3422	0.89
P98	38	6	2.8314	1.3666	0.6161
P100	44	8	6.4533	1.937	0.8256
P116	50	12	8.1169	2.2443	0.8644
P119	46	7	3.6357	1.5382	0.6887
P123	50	10	5.9524	1.996	0.8134
P128	44	7	2.8471	1.3833	0.6152
P154	46	13	9.5315	2.3757	0.8858
P261	48	5	2.0719	1.0409	0.4829
P280	40	14	9.7561	2.4601	0.8892
P305	48	12	7.7838	2.2473	0.8593
P318	38	9	6.8113	2.0433	0.8367
P337	52	11	7.9529	2.1907	0.8612
P346	48	11	5.76	2.0044	0.8062
P356	42	9	3.9552	1.7459	0.7257
Mean	45	10.1765	6.7065	1.9872	0.7898

Note: * na, Observed number of alleles; * ne, Effective number of alleles; * I, Shannon's Information index; * PIC, Polymorphic information content.

Kinship analysis of Carex L. using the SSR identification system

According to the genotypic matrix data, the genetic similarity coefficient and genetic distance matrix of 26 tested accessions were calculated (Table S4 and Table 5). The values of the genetic similarity coefficient (GSC) among the different *Carex* L. accessions ranged from 0.00 to 1.00, indicating prevalence of abundant genetic divergence among them (Table 5). Larger the GSC value, closer the relationship, and vice versa. Among all accessions, Pizhen-2 (C13) and Aicong-2 (C14) showed the largest GSC (1.0), whereas GSCs of other comparisons were relatively smaller, which indicated that Pizhen-2 and Aicong-2 possessed the closest relationship, whereas others had relatively distant relationships (Table 5). In addition, the GSC values of several other comparisons were zero indicating their most distant relationship and wider genetic diversities (Table 5). According to the GSC matrix, the SSR detection data of 26 tested accessions were analyzed by the UPGMA method to obtain a tree cluster diagram. The results revealed that 26 samples were roughly divided into three groups with an average genetic similarity coefficient of 0.1796 (Figure 2). Of the four widely used accessions, Laoyv-2 (C11) was in the group I, Qinglv-2 (C12) and Jiao-2 (C15) were in group II, Pizhen-2 (C13) in group III, indicating that each group of *Carex* L. accessions was fully utilized in garden landscapes.

Discussion

Carex L. has a broad prospect in the application of environmental greening and beautification (Zhang *et al.* 2010; Liang *et al.* 2012). Therefore, it is important to study the *Carex* L. classification and the relationship between different populations and environments, as well as to restore the fragility of the ecological environment of our country (Ma *et al.* 2001). It was essential to utilize nucleotide polymorphisms to study the *Carex* L. classification and determine the genetic relationships and evolution of *Carex* L. accessions.

Is the significantly high percentage of the heterozygous genome related to abiotic resistance?

The vitality of *Carex* L. is tenacious; many species in this genus are resistant to shade, barren land, and cold environments (Shi 2007). In this study, based on SNP analysis, we found that Jiao-2 had the highest heterozygosity (96.58%), followed by Pizhen-2 (74.68%), Qinglv-2 (59.91%), and Laoyv-2 (33.15%). The dramatic differences in

the heterozygous SNP percentages were possibly related to different degrees of asexual propagation among different *Carex* L. accessions. In addition, it is speculated that *Carex* L. accessions with high heterozygosity may have high heterosis, which may also be related to the wider adaptability of this genus to the environment and poor tolerance and resistance to diseases and insects. The heterozygous genome of *Carex* L. can be maintained by asexual propagation. *Carex* L. Jiaotai is the most primitive, with a low garden utilization rate, so its heterozygosity is the highest. In comparison, Laoyv is widely-used and gradually purified artificially; consequently, its genome is gradually becoming homozygous. The high heterozygosity of *Carex* L. is beneficial for protection of the diversity in genetic resources and have great significance for development of *Carex* L. varieties suitable for different ecological regions.

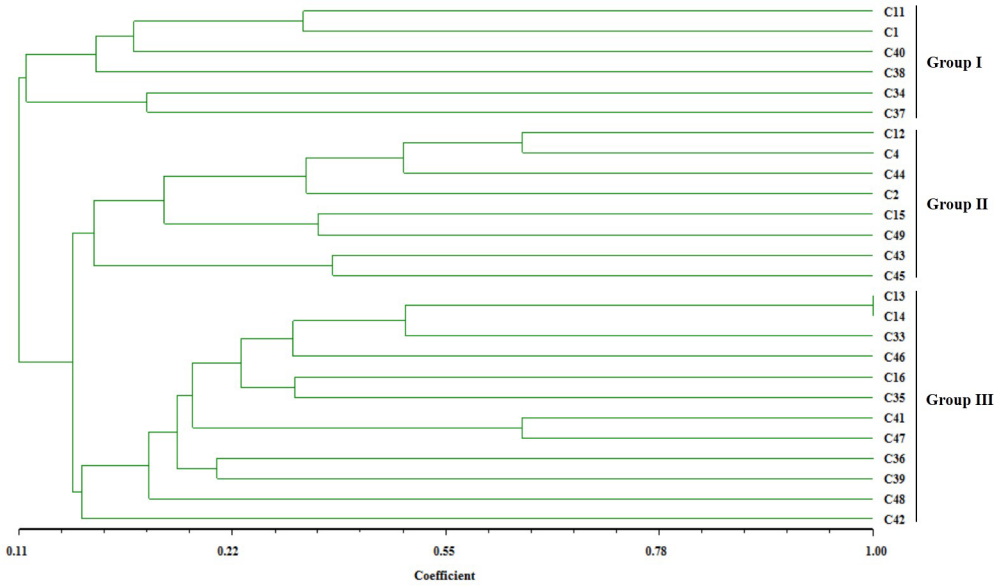

FIGURE 2. Cluster analysis of the twenty-six *Carex* L. accessions.

Establishment of the SSR identification for in Carex L.

There are several kinds of excellent lawn ground cover plants with early green return, good color, long growth duration, and good aesthetic beautification effects. The inherent reproductive strategies, strong vegetative reproduction ability, special physiological integration, and strong vitality enables *Carex* L. to grow in extremely fragile ecological environments, such as alpine grasslands , alpine ice margins, and many wetland systems (Yang *et al.* 2000). Therefore, it plays an extremely important role in maintaining the fragile ecological environment. However, there are few studies on the establishment of a *Carex* L. germplasm resource bank, gene bank, and the evaluation and classification of *Carex* L. species by molecular means. In this study, four representative *Carex* L. accessions were sequenced and analyzed for the first time by RAD sequencing, and a large number of genome sequences were obtained. Based on the genome sequences, a core SSR identification system was developed and evaluated to classify *Carex* L. accessions. SSRs are one of the important genetic markers for the QTL mapping, assessment of genetic diversity and population structure, hybridity testing, and marker-assisted plant breeding (Vidak *et al.* 2021). In wheat, the genome-wide SSR markers were utilized to construct genetic linkage maps (Somers *et al.* 2004; Han *et al.* 2015), identify the exotic introgressed fragments (Gu *et al.* 2015) and QTL mapping (Cui *et al.* 2011). Meanwhile, SSRs were also used in other crop species such as *Brassica*, rice and barley (Mohammadi *et al.* 2020; Daware *et al.* 2016; Lowe *et al.* 2004) as well as in cash crop species such as grape (Adam-Blondon *et al.* 2004), coconut (Caro *et al.* 2022), and *Lilium* species (Biswas *et al.* 2020). Here, this research bridge gaps in molecular identification system of *Carex* L. and lays a scientific foundation for plant researches for innovative molecular resources based improvement of *Carex* L. accessions. As more molecular markers would be available in near future, the genetic relationships would be established among diverse *Carex* L. accessions and their relative traits including shade-tolerance, green-returning, leaf color, and growth duration, which could accelerate efforts for molecular breeding based improvements of *Carex* L..

TABLE 5. Similarity coefficient matrix of 26 samples based on SSR Markers.

	C11	C12	C13	C14	C15	C16	C33	C34	C35	C36	C37	C38	C39	C40	C41	C42	C43	C44	C45	C46	C47	C48	C49	C1	C2	C4
C11	****	0.09	0.12	0.12	0.15	0.29	0.08	0.22	0.12	0.17	0.03	0.27	0.23	0.19	0.18	0.02	0.02	0.09	0.06	0.14	0.19	0.18	0.06	0.40	0.13	0.04
C12	2.36	****	0.14	0.14	0.22	0.11	0.15	0.09	0.13	0.04	0.03	0.02	0.15	0.08	0.24	0.13	0.20	0.40	0.15	0.27	0.37	0.07	0.29	0.10	0.27	0.63
C13	2.12	1.94	****	1.00	0.19	0.32	0.51	0.16	0.38	0.29	0.03	0.11	0.33	0.09	0.32	0.27	0.09	0.12	0.09	0.43	0.29	0.18	0.38	0.19	0.19	0.20
C14	2.12	1.94	0.00	****	0.19	0.32	0.51	0.16	0.38	0.29	0.03	0.11	0.33	0.09	0.32	0.27	0.09	0.12	0.09	0.43	0.29	0.18	0.38	0.19	0.19	0.20
C15	1.87	1.53	1.67	1.67	****	0.00	0.06	0.10	0.06	0.08	0.00	0.00	0.02	0.14	0.06	0.08	0.06	0.21	0.06	0.17	0.16	0.04	0.42	0.14	0.28	0.19
C16	1.22	2.24	1.14	1.14	0.00	****	0.33	0.21	0.40	0.24	0.12	0.19	0.32	0.09	0.31	0.17	0.11	0.21	0.12	0.27	0.17	0.27	0.13	0.24	0.13	0.23
C33	2.50	1.92	0.67	0.67	2.75	1.12	****	0.16	0.41	0.23	0.09	0.14	0.23	0.03	0.22	0.25	0.20	0.16	0.20	0.31	0.21	0.18	0.28	0.19	0.15	0.18
C34	1.50	2.36	1.83	1.83	2.34	1.56	1.81	****	0.27	0.23	0.24	0.08	0.17	0.14	0.16	0.20	0.16	0.17	0.21	0.24	0.21	0.18	0.17	0.23	0.09	0.08
C35	2.09	2.08	0.97	0.97	2.75	0.93	0.89	1.32	****	0.34	0.18	0.09	0.25	0.00	0.37	0.22	0.17	0.23	0.20	0.31	0.28	0.28	0.21	0.17	0.15	0.18
C36	1.75	3.24	1.23	1.23	2.53	1.43	1.45	1.47	1.08	****	0.00	0.19	0.31	0.12	0.18	0.13	0.10	0.13	0.09	0.31	0.24	0.28	0.12	0.10	0.21	0.10
C37	3.62	3.60	3.54	3.54	0.00	2.11	2.42	1.43	1.73	0.00	****	0.12	0.05	0.04	0.00	0.00	0.03	0.08	0.12	0.06	0.00	0.05	0.00	0.06	0.00	0.00
C38	1.32	3.87	2.20	2.20	0.00	1.68	1.99	2.50	2.40	1.65	2.13	****	0.10	0.12	0.11	0.11	0.00	0.02	0.00	0.09	0.00	0.00	0.04	0.17	0.10	0.02
C39	1.49	1.87	1.12	1.12	3.93	1.14	1.47	1.77	1.38	1.16	2.91	2.26	****	0.08	0.30	0.20	0.26	0.11	0.21	0.29	0.18	0.20	0.12	0.10	0.07	0.14
C40	1.64	2.47	2.40	2.40	1.94	2.36	3.48	1.97	0.00	2.16	3.22	2.09	2.47	****	0.00	0.04	0.06	0.03	0.12	0.12	0.09	0.11	0.09	0.26	0.03	0.03
C41	1.73	1.43	1.14	1.14	2.79	1.17	1.53	1.85	0.99	1.69	0.00	2.22	1.20	0.00	****	0.18	0.15	0.26	0.17	0.19	0.63	0.25	0.14	0.18	0.07	0.23
C42	3.70	2.07	1.31	1.31	2.56	1.78	1.40	1.62	1.51	2.05	0.00	2.21	1.60	3.29	1.69	****	0.11	0.02	0.06	0.11	0.03	0.07	0.15	0.05	0.00	0.00
C43	3.93	1.61	2.46	2.46	2.79	2.20	1.63	1.85	1.75	2.28	3.56	0.00	1.35	2.83	1.93	2.25	****	0.41	0.44	0.28	0.16	0.23	0.22	0.10	0.12	0.26
C44	2.38	0.93	2.12	2.12	1.55	1.56	1.81	1.79	1.49	2.00	2.52	3.89	2.18	3.58	1.37	3.70	0.89	****	0.21	0.22	0.33	0.15	0.25	0.13	0.35	0.62
C45	2.79	1.92	2.42	2.42	2.75	2.15	1.59	1.58	1.59	2.46	2.13	0.00	1.57	2.09	1.75	2.90	0.83	1.58	****	0.36	0.19	0.20	0.26	0.19	0.07	0.20
C46	1.96	1.33	0.83	0.83	1.79	1.30	1.17	1.42	1.17	1.16	2.85	2.42	1.25	2.12	1.65	2.23	1.29	1.51	1.03	****	0.48	0.33	0.35	0.15	0.24	0.33
C47	1.65	0.99	1.23	1.23	1.83	1.75	1.55	1.55	1.28	1.43	0.00	0.00	1.73	2.45	0.46	3.66	1.81	1.12	1.65	0.73	****	0.26	0.26	0.20	0.16	0.33
C48	1.70	2.60	1.73	1.73	3.28	1.32	1.71	1.70	1.27	1.26	2.95	0.00	1.59	2.22	1.38	2.61	1.46	1.93	1.60	1.09	1.33	****	0.17	0.23	0.07	0.14
C49	2.85	1.22	0.98	0.98	0.87	2.03	1.28	1.75	1.55	2.12	0.00	3.16	2.14	2.45	1.95	1.86	1.49	1.39	1.36	1.04	1.35	1.77	****	0.28	0.30	0.33
C1	0.91	2.32	1.67	1.67	1.97	1.43	1.65	1.47	1.77	2.30	2.89	1.77	2.32	1.35	1.69	2.96	2.28	2.00	1.65	1.92	1.61	1.48	1.27	****	0.14	0.06
C2	2.02	1.31	1.65	1.65	1.29	2.08	1.92	2.43	1.92	1.58	0.00	2.32	2.70	3.41	2.66	0.00	2.15	1.04	2.61	1.43	1.83	2.73	1.21	1.98	****	0.60
C4	3.22	0.46	1.63	1.63	1.67	1.48	1.73	2.52	1.73	2.26	0.00	3.81	1.94	3.50	1.45	0.00	1.37	0.47	1.61	1.12	1.10	1.98	1.10	2.77	0.51	****

Acknowldgement

This work was supported by grants from the Beijing Natural Science Foundation Project (6172015) and National Natural Science Fund Projects (32072061 and 31100865). We thanks Dr. Qasim Raza for revising the grammar and expression in the article.

References

Adam-Blondon, A.F., Roux, C., Claux, D., Butterlin, G., Merdinoglu, D. & This, P. (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. *Theoretical and Applied Genetics* 109: 1017–1027. https://doi.org/10.1007/s00122-004-1704-y

Beier, S., Thiel, T., Münch, T., Scholz, U. & Mascher, M. (2017) MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33: 2583–2585. https://doi.org/10.1093/bioinformatics/btx198

Biswas, M.K., Bagchi, M., Nath, U.K., Biswas, D., Natarajan, S., Jesse, D.M.I., Park, J.I. & Nou, I.S. (2020) Transcriptome wide SSR discovery cross-taxa transferability and development of marker database for studying genetic diversity population structure of *Lilium* species. *Scientific reports* 10: 18621. https://doi.org/10.1038/s41598-020-75553-0

Bruce, A.F., Habibollah, G., Robert, F.C.N. & Julian, R.S. (2012) Phylogeny of *Carex* subg. *Vignea* (Cyperaceae) based on amplified fragment length polymorphism and nrDNA data. *Systematic Botany* 37: 913–925. https://doi.org/10.1600/036364412X656464

Caro, R.E.S., Cagayan, J., Gardoce, R.R., Manohar, A.N.C., Canama-Salinas, A.O., Rivera, R.L., Lantican, D.V., Galvez, H.F. & Reaño, C.E. (2022) Mining and validation of novel simple sequence repeat (SSR) markers derived from coconut (*Cocos nucifera* L.) genome assembly. *Journal of Genetic Engineering & Biotechnology* 20: 71. https://doi.org/10.1186/s43141-022-00354-z

Chen, S., Zhou, Y., Chen, Y. & Gu, J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34: i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Cui, F., Li, J., Ding, A., Zhao, C., Wang, L., Wang, X., Li, S., Bao, Y., Li, X., Feng, D., Kong, L. & Wang, H. (2011) Conditional QTL mapping for plant height with respect to the length of the spike and internode in two mapping populations of wheat. *TAG Theoretical and Applied Genetics* 122: 1517–1536. https://doi.org/10.1007/s00122-011-1551-6

Daware, A., Das, S., Srivastava, R., Badoni, S., Singh, A.K., Agarwal, P., Parida, S.K. & Tyagi, A.K. (2016) An efficient strategy combining SSR markers- and advanced QTL-seq-driven QTL mapping unravels candidate genes regulating grain weight in rice. *Frontiers in plant science* 7: 1535. https://doi.org/10.3389/fpls.2016.01535

Gu, L., Wei, B., Fan, R., Jia, X., Wang, X. & Zhang, X. (2015) Development, identification and utilization of introgression lines using Chinese endemic and synthetic wheat as donors. *Journal of Integrative Plant Biology* 57: 688–697. https://doi.org/10.1111/jipb.12324

Han, B., Wang, C., Tang, Z., Ren, Y., Li, Y., Zhang, D., Dong, Y. & Zhao, X. (2015) Genome-wide analysis of microsatellite markers based on sequenced database in Chinese spring wheat (*Triticum aestivum* L.). *PloS one* 10:e0141540. https://doi.org/10.1371/journal.pone.0141540

Hendrichs, M., Michalski, S., Begerow, D., Oberwinkler, F. & Hellwig, F. (2004a) Phylogenetic relationships in *Carex*, subgenus *Vignea* (Cyperaceae), based on ITS sequences. *Plant Systematics and Evolution* 246: 109–125. https://doi.org/10.1007/s00606-004-0127-1

Hendrichs, M., Oberwinkler, F., Begerow, D. & Bauer, R. (2004b) *Carex*, subgenus *Carex* (Cyperaceae)—A phylogenetic approach using ITS sequences. *Plant Systematics and Evolution* 246: 89–107. https://doi.org/10.1007/s00606-004-0128-0

Jiménez-Mejías, P., Hahn, M., Lueders, K., Starr, J.R., Brown, B.H., Chouinard, B.N., Chung, K.-S., Escudero, M., Ford, B.A., Ford, K.A., Gebauer, S., Gehrke, B., Hoffmann, M.H., Jin, X.-F., Jung, J., Kim, S., Luceño, M., Maguilla, E., Martín-Bravo, S., Míguez, M., Molina, A., Naczi, R.F.C., Pender, J.E., Reznicek, A.A., Villaverde, T., Waterway, M.J., Wilson, K.L., Yang, J.-C., Zhang, S., Hipp, A.L. & Roalson, E.H. (2016) Megaphylogenetic specimen-level approaches to the *Carex* (Cyperaceae) phylogeny using ITS, ETS,

and matK sequences: implications for classification. *Systematic Botany* 41: 500–518.
https://doi.org/10.1600/036364416X692497

Kõressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R. & Remm, M. (2018) Primer3_masker: integrating masking of template sequence with primer design software. *Bioinformatics* 34: 1937–1938.
https://doi.org/10.1093/bioinformatics/bty036

Li, Y. & He, M. (2014) Genetic mapping and QTL analysis of growth-related traits in *Pinctada fucata* using restriction-site associated DNA sequencing. *PloS one* 9: e111707.
https://doi.org/10.1371/journal.pone.0111707

Liang, F., Dong, A. & Ma, Y. (2012) Resource survey of wild *Carex* plants and evaluation of their ornamental characteristics in Beijing. *Pratacultural Science* 6: 710–716.

Lowe, A.J., Moule, C., Trick, M. & Edwards, K.J. (2004) Efficient large-scale development of microsatellites for marker and mapping applications in Brassica crop species. *Theoretical and Applied Genetics* 108: 1103–1112.
https://doi.org/10.1007/s00122-003-1522-7

Ma, W., Han, L. & Luo, J. (2001) A new lawn plant resource: genus *Carex* L. *Pratacultural Science* 18: 43–56.

Martín-Bravo, S., Jiménez-Mejías, P., Villaverde, T., Escudero, M., Hahn, M., Spalink, D., Roalson, E.H., Hipp, A.L., Group, t.G.C., Benítez-Benítez, C., Bruederle, L.P., Fitzek, E., Ford, B.A., Ford, K.A., Garner, M., Gebauer, S., Hoffmann, M.H., Jin, X.-F., Larridon, I., Léveillé-Bourret, É., Lu, Y.-F., Luceño, M., Maguilla, E., Márquez-Corro, J.I., Míguez, M., Naczi, R., Reznicek, A.A. & Starr, J.R. (2019) A tale of worldwide success: Behind the scenes of *Carex* (Cyperaceae) biogeography and diversification. *Journal of Systematics and Evolution* 57: 695–718.
https://doi.org/10.1111/jse.12549

Mohammadi, S.A., Abdollahi Sisi, N. & Sadeghzadeh, B. (2020) The influence of breeding history, origin and growth type on population structure of barley as revealed by SSR markers. *Scientific reports* 10: 19165.
https://doi.org/10.1038/s41598-020-75339-4

Murray, M.G. & Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic acids research* 8: 4321–4326.
https://doi.org/10.1093/nar/8.19.4321

Ning, H., Wang, W., Zheng, C., Li, Z., Zhu, C. & Zhang, Q. (2014a) Genetic diversity analysis of sedges (*Carex* spp.) in Shandong, China based on inter-simple sequence repeat. *Biochemical Systematics and Ecology* 56: 158–164.
https://doi.org/10.1016/j.bse.2014.05.014

Ning, H., Zheng, C., Wang, W., Zhu, C., Li, Z. & Zhang, Y. (2014b) Establishment and optimization of lSSR-PCR system in Carex. *Molecular Plant Breeding* 12: 349–355.

Reznicek, A.A. (1990) Evolution in sedges (*Carex cyperaceae*). *Canadian Journal of Botany* 68: 1409–1432.
https://doi.org/10.1139/b90-180

Rochette, N.C. & Catchen, J.M. (2017) Deriving genotypes from RAD-seq short-read data using Stacks. *Nature protocols* 12: 2640–2659.
https://doi.org/10.1038/nprot.2017.123

Shi, J. (2007) Study on ecological adaptability of *Carex giraldiana* in Beijing area. *Pratacultural Science* 24: 98–102.

Somers, D.J., Isaac, P. & Edwards, K. (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 109: 1105–1114.
https://doi.org/10.1007/s00122-004-1740-7

Vidak, M., Šatović, Z., Liber, Z., Grdiša, M., Gunjača, J., Kilian, A. & Carović-Stanko, K. (2021) Assessment of the origin and diversity of croatian common bean germplasm using phaseolin type, SSR and SNP markers and morphological traits. *Plants* 10: 665.
https://doi.org/10.3390/plants10040665

Wang, N., Fang, L., Xin, H., Wang, L. & Li, S. (2012) Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. *BMC plant biology* 12: 148.
https://doi.org/10.1186/1471-2229-12-148

Xue, H., Sha, W. & Ni, H. (2005) General situation of studies on *Carex* L. *Journal of qiqihar University* 20: 789.

Yang, L.J., Li, X.L., Shi, D.J. & Sa, W.J. (2000) Study on the bio-diversity of alpine plant communities in the higher altitude area of south Qinghai. *Grassland and Turf* 89: 32–35.

Zhang, C., Zhu, X., Cai, K. & Yu, Y. (2010) Evaluation of shade tolerance of *Carex* species available for garden-environment planting. *Journal of Beijing Forestry University* 32: 207–212.