

Article

Development of Polymorphic Genic SSR Markers by Transcriptome Sequencing in the Welsh Onion (*Allium fistulosum* L.)

Liuyi Yang ^{1,2,†}, Changlong Wen ^{1,†}, Hong Zhao ¹, Qianchun Liu ¹, Jingjing Yang ¹, Lecheng Liu ^{2,*} and Yongqin Wang ^{1,*}

¹ Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences/Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), Ministry of Agriculture, Beijing 100097, China; E-Mails: yly0716@outlook.com (L.Y.); wenchanglong@nercv.org (C.W.); zhaohong@nercv.org (H.Z.); lqc072150@126.com (Q.L.); yangjingjing@nercv.org (J.Y.)

² College of Horticulture and Gardening, Yangtze University, Jingzhou 434025, China

[†] These authors contributed equally to this work.

* Authors to whom correspondence should be addressed; E-Mails: lchliu18@yangtzeu.edu.cn (L.L.); wangyongqin@nercv.org (Y.W.); Tel.: +86-71-6806-6264 (L.L.); +86-10-5150-3125 (Y.W.); Fax: +86-71-6806-0812 (L.L.); +86-10-5150-5002 (Y.W.).

Academic Editor: Takayoshi Kobayashi

Received: 14 September 2015 / Accepted: 23 October 2015 / Published: 5 November 2015

Abstract: Transcriptome analysis is an efficient way to explore molecular markers in plant species, for which genome sequences have not been published. To address the limited number of markers published for the Welsh onion, this study found 6486 loci of genic simple sequence repeats (SSR), which consisted of 1–5 bp repeat motifs, based on next-generation sequencing (NGS) technology and the RNA-Seq approach. The most abundant motif was mononucleotide (52.33%), followed by trinucleotide (31.96%), and dinucleotide (14.57%). A total of 2525 primer pairs were successfully designed, and 91 out of 311 tested primers were polymorphisms. Overall, 38 genic SSR markers were randomly selected to further validate the degree of genetic diversity, and 22 genic SSR markers (57.89%) showed high levels of polymorphism. The average polymorphism information content (PIC) value and the number of alleles (Na) were 0.63 and 5.27, respectively, and the unweighted pair-group method with arithmetic average (UPGMA) cluster analysis grouped

the 22 *Allium* accessions into three groups with Nei's similarity coefficients ranging from 0.37 to 0.99. This result suggested that these genic SSR markers could be used to develop a higher resolution genetic map and/or to analyze the phylogenetic relationships among *Allium* plants in the near future.

Keywords: simple sequence repeats (SSR); genic marker; transcriptome; genetic diversity; *Allium fistulosum* L.

1. Introduction

The Welsh onion (*Allium fistulosum* L. $2n = 16$) is a cultivar of Alliaceae *Allium* sp., which are some of the most commercially-important biennial and perennial herbs, and its plump stalk, including the leaf sheath and tender leaf, they are consumed as vegetables and condiments worldwide [1]. The plant is thought to have been derived from Northwestern China, and it is widely cultivated worldwide, particularly in East Asian countries, such as Japan, China, and Korea [2,3]. The Welsh onion is highly nutritious, with bactericidal and anti-inflammatory benefits, and it has been used as an herbal medicine to treat many diseases [4]. Due to its great nutritional and medicinal value, the Welsh onion has become one of the main vegetables exported from China in recent years. Additionally, more and more researchers have been attempting to develop molecular markers for the Welsh onion to construct a genetic map or for use in a marker-assisted selection (MAS) breeding system.

A variety of molecular markers, such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), and sequence characterized amplified regions (SCAR), have been developed and used for molecular marker-assisted selection and genetic diversity analysis in the Welsh onion [5–7]. In addition, microsatellites, or simple sequence repeats (SSRs), have also been applied to the construction of chromosome maps and cultivar identification [8–11]. Compared with other molecular markers, SSR markers are more ideal due to their reproducibility, multi-allelic nature, co-dominant inheritance, relative abundance, and good genome coverage [12]; SSR loci are composed of randomly repeated DNA regions with motif lengths of one to six base pairs (bp) that are spread throughout the genome [13]. In recent years, a large number of genomic SSR markers have been developed in the Welsh onion [14,15], but genic SSR markers have largely been ignored until recently; Tsukazaki *et al.*, described transcriptome shotgun assembly (TSA)-derived markers in Japanese bunching onion [10]. Genic SSRs have distinct advantages compared to genomic SSRs in terms of reducing the time and cost of developing SSR markers, and they are more useful as genetic markers because they are part of the transcriptome and represent the variation in the expressed portion of the genome [16]. Recently, hundreds of genic SSR markers, even SNPs (single nucleotide polymorphisms), have been developed by analyzing large amounts of ESTs in the bulb onion (*A. cepa* L.) cDNA libraries [17–19]. Therefore, developing genic SSR markers from transcriptome sequencing of the Welsh onion is a valuable area of research.

Due to the lack of whole genome information and the limited number of markers published for the Welsh onion, the priority of this study was to report on the development of genic SSR markers using a transcriptome approach, which provides new molecular markers for genetic diversity studies in the

Welsh onion and related *Allium* species. In addition, these genic SSR markers will provide important resources for chromosome mapping and phylogenetic analysis of *Allium* plants, potentially contributing to target gene cloning, as well as the utilization of the MAS breeding.

2. Results and Discussion

2.1. Frequency and Distribution of Genic SSRs in the Welsh Onion Transcriptome

Since the RNA-Seq NGS technology allows for lower cost and higher throughput transcriptome sequencing than traditional Sanger sequencing, it has been widely used in many non-model organisms, including *Tapiscia sinensis*, *Cucurbita moschata*, and *Larix kaempferi* [20–22].

In this study, a total of 6766 SSR loci containing 1–5 bp repeat motifs were found in 6206 unigenes of the Welsh onion. Of these, 5497 (88.6%) contained a single SSR locus; 486 (7.8%) contained more than one SSR, and 275 (4.4%) possessed compound SSRs. Moreover, 6486 SSR loci (except for 280 compound SSRs) located in 5983 unigenes were counted and presented based on the transcriptome data (Table 1).

Table 1. Frequencies of the different SSR repeat motif types observed in the Welsh onion transcriptome.

SSR Motif	Repeat Number									Total	Percentage (%)
	5	6	7	8	9	10	11	12	13–24		
Mononucleotide	0	0	0	0	0	2060	548	297	489	3394	52.33
Dinucleotide	0	478	212	119	57	32	37	10	0	945	14.57
Trinucleotide	1328	510	188	47	0	0	0	0	0	2073	31.96
Tetranucleotide	52	14	0	0	0	0	0	0	0	66	1.02
Pentanucleotide	8	0	0	0	0	0	0	0	0	8	0.12
Total	1388	1002	400	166	57	2092	585	307	489	6486	100
Percentage (%)	21.40	15.45	6.17	2.56	0.88	32.25	9.02	4.73	7.54	100	

The repeat motifs among the 6486 SSR loci included a total of 186 types, and the number of repeats varied from five to 24. The frequency and number of each SSR motif and the percentage of the primary repeat motifs were evaluated in Tables 1 and 2. The most abundant motif was mononucleotide (52.33%) followed by trinucleotide (31.96%), dinucleotide (14.57%), tetranucleotide (1.02%), and pentanucleotide (0.12%), and the most common class was $n = 10$ (32.25%), which was mostly composed of mononucleotide repeats. In addition, the most common class of dinucleotide was $n = 6$, and other types of repeat motif were $n = 5$. A/T was the most common motif, representing 98.29% of the mononucleotide repeat motifs and 51.43% of the total repeat motifs. Additionally, the dominant repeat motifs were AG/CT (34.29%), AT/AT (33.65%), and AC/GT (31.11%) among the dinucleotides, and AAG/CTT (26.44%) and AAAT/ATTT (33.33%) among the trinucleotides and tetranucleotides, respectively.

Table 2. Number and frequencies of the main repeat motif types.

Main Repeat Motif	Repeat Number	The Percentage of Each Type of SSR Repeat Motif (%)	The Percentage of Total SSR Repeat Motifs (%)
A/T	3336	98.29	51.43
C/G	58	1.71	0.89
AC/GT	294	31.11	4.53
AG/CT	324	34.29	5.00
AT/AT	318	33.65	4.90
CG/CG	9	0.95	0.14
AAC/GTT	118	5.69	1.82
AAG/CTT	548	26.44	8.45
AAT/ATT	146	7.04	2.25
ACC/GGT	145	6.99	2.24
ACG/CGT	86	4.15	1.33
ACT/AGT	43	2.07	0.66
AGC/GCT	297	14.33	4.58
AGG/CCT	344	16.59	5.30
ATC/GAT	314	15.15	4.84
CCG/CGG	32	1.54	0.49
AAAC/GTTT	7	10.61	0.11
AAAG/CTTT	7	10.61	0.11
AAAT/ATTT	22	33.33	0.34
ACAT/ATGT	19	28.79	0.29

Among the dinucleotide repeat motifs based on the 6468 SSR loci, the frequency of GA as a primary nucleotide (5.00%) was higher than GT (4.53%) in the Welsh onion transcriptome sequences (Table 3), and this result agreed with Tsukazaki *et al.*, (GA 22.1%, GT 15.4%) [10]. Similar results were also obtained for the bulb onion expressed sequence tags (ESTs) (GA 10.1%, GT 9.2%) [17], but GT repeats were much richer than GA repeats in the bulb onion genome [23], which was also the case in Welsh onion [9,15]. According to the above results, the ratio of $(GT)_n:(GA)_n$ may be different in the transcribed and non-transcribed regions of the genome. Additionally, the GA dinucleotide repeat motif can stand for multiple codons resting within the reading frame that can be translated into different amino acids, and GA may be present in the Ala and Leu codons, which have the highest number of proteins [24].

In addition, AT was the most abundant repeat motif among the dinucleotides and was more frequent than GC in the Welsh onion (Table 1). A similar result was reported in Tsukazaki *et al.*, (AT 25.6%, GC 0.5%) [10] and, based on transcriptome sequencing, the same result that the AT repeat motif was more common than GC, has been obtained in many plant species, such as bulb onion, radish, mung bean, pear, Levant cotton, and so on [16,17,25–27]. Although a large number of the AT motif emerged from the transcriptome sequences, it has not typically been used to develop molecular markers because of its self-complementary nature, which leads to the formation of dimers [28].

Table 3. List of 22 *Allium* accessions used in the polymorphism analysis.

No.	Materials	Origin
1	“Yuanzang”	Japan
2	“HanLvDaZhu”	Korea
3	“WuYeQi”1	Hebei, China
4	“WuYeQi”2	Tianjin, China
5	“GaoBaiQingCaiCong”	Hebei, China
6	“ZhongYingTeDa”	Henan, China
7	“SanYu”	Shandong, China
8	“YeFuYiHao”	Shandong, China
9	“ZhangQiuDaCong”	Shandong, China
10	“ErShengZi”	Shandong, China
11	“Zao134”	Shandong, China
12	“ShanXiFenCong”	Shanxi, China
13	“Qi110”	Shandong, China
14	“ZhongHuaJuCong”	Shandong, China
15	“WuYeQi”3	Shanxi, China
16	“ShanXiYunChengKeXingYiHao”	Shanxi, China
17	“HeBeiJinXianDaCong”	Hebei, China
18	“Shou1105”	Shandong, China
19	“LiaoNingChaoYangLinCong”	Liaoning, China
20	“XiYeWanShengCong”	Japan
21	“SongBenYiBenTaiCong”	Japan
22	“7722”	Guangdong, China

2.2. Development and Detection of Genic SSR Markers

Genic SSR markers are regarded as having great potential for genetic diversity analysis and chromosome mapping in crop species because of their specificity and high degree of conservation [29,30]. Due to the steady decrease in cost and the increase in throughput data, NGS technology has become a powerful approach for the high-throughput discovery of genes and the generation of a large amount of sequence data for the identification of molecular markers [31,32]. Therefore, the identification and development of genic SSR markers based on NGS is an efficient and cost-effective strategy.

In this study, a total of 2710 unigenes were used to design primers, and according to the standard, a total of 2512 unigenes (92.69%) were successfully designed, which obtained 2525 primers. A total of 311 (12.32%) primer pairs located in 307 unigenes were randomly selected from 2525 primers (Table S2) to perform the preliminary test, which consisted of 217 primers (69.77%) for the trinucleotide repeats, 50 (16.08%) for the dinucleotide repeats, and 38 (12.22%) and six (1.93%) for the tetranucleotide and pentanucleotide repeats, respectively. Furthermore, the remaining 2214 primers, located in 2205 unigenes, are presented in table S1 for reference.

In the preliminary test, a total of 165 (53.05%) primers generated stable and clear bands, out of which 91 (29.26%) polymorphic primer pairs were reliable (Table S3). Among these genic SSR markers, the majority were trinucleotide repeats (69.23%), followed by dinucleotide repeats (19.78%), and tetranucleotide repeats (10.99%). The result was consistent with the order of SSR motif frequency.

Compared with the bulb onion, the Welsh onion is suitable for genetic studies because it has a 28% smaller genome [33], but some of the genic SSR markers used in many genetic studies of the construction of a chromosome map and cultivar identification in Welsh onion [8,10,11] are derived from the bulb onion. The genic SSR marker in the Welsh onion has been far less developed than that in the bulb onion because less attention is currently being paid to the Welsh onion. Therefore, due to the limited number of SSR markers, we explored more SSR markers based on transcriptome sequencing to enrich the research into molecular marker-assisted selection and genetic diversity in the Welsh onion.

2.3. Validating the Effectiveness of Genic SSR Markers among 22 *Allium* Accessions

A sample of 38 primers (Table S4) was randomly selected from 91 polymorphic primer pairs to validate the effectiveness of genic SSR markers and to differentiate a set of 22 *Allium* accessions (Table 3). Of these 38 primers, six genic SSR markers could not be effectively amplified, and 10 genic SSR markers generated unified or poor polymorphic bands, which could have been due to the use of inappropriate materials or defective SSR markers. A total of 22 genic SSR markers exhibited a high level of polymorphism (Table 4), and 116 alleles were detected in total. The PCR amplification bands generated by the genic SSR markers MCL32, MCL40, and TO249 in 22 *Allium* accessions are displayed in Figure 1.

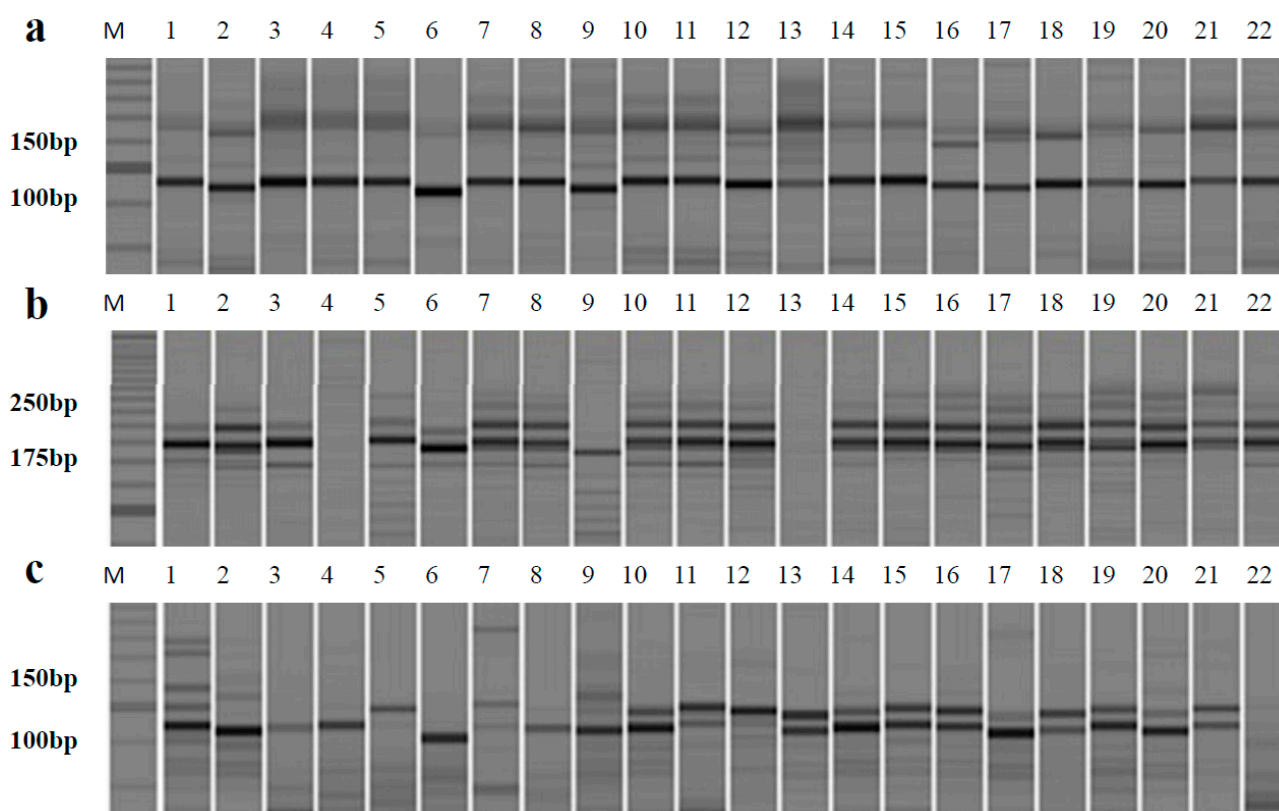


Figure 1. PCR amplification bands generated by the Welsh onion genic SSR markers MCL32 (a), MCL40 (b), and TO249 (c) in 22 *Allium* accessions.

Table 4. Characteristics of the 22 polymorphic SSR markers validated in 22 *Allium* accessions.

No.	Primer	Repeat Motif	Primer Sequence (5'–3')	Size Range (bp)	A ^a	Ng ^b	Na ^c	PIC ^d
1	MCL32	(TG)9	F:gatttcccatgcatcatct R:ggcacaactgctgcatctca	111–117	0.4545	4	4	0.5388
2	MCL37	(AAG)5	F:aaggaatgctacgccagaga R:ctgaattctgctgggtctcc	411–461	0.4333	8	8	0.7195
3	MCL40	(AGA)5	F:cgggaagtagcatcggatcat R:aggctcgtgctttgtgtctt	189–226	0.4250	6	6	0.6882
4	MCL42	(GCA)7	F:cgggaacgaagatggata R:aacgaccaacaacgtccttc	355–497	0.2273	12	9	0.8465
5	MCL47	(GAT)5	F:aggagtggaaagtggatcat R:ggatcgctatccgagtcca	291–298	0.5455	3	3	0.4762
6	WC121	(TCT)5	F:ccgaggatctcgaactgta R:tctcacagtcctcactgc	214–249	0.6429	4	4	0.4935
7	WC122	(AGG)5	F:attcttgaaccgcaccaac R:caagacgagcaacaacctca	299–334	0.3750	7	8	0.6939
8	WC126	(TCC)6	F:tcctgcctcttagatgctt R:caagagggaagggttgta	306–320	0.5000	5	5	0.6381
9	WC183	(ATG)5	F:aacaactgccacacaacca R:cttctggaaggcctcctctt	289–304	0.5909	5	4	0.4900
10	WC193	(CCT)5	F:cccagttaagcagccactc R:tggtcgttcaggaggagctt	266–272	0.5882	3	3	0.4365
11	WC200	(ATC)5	F:acaggatccatggaatgag R:agtgcagactggaatggag	158–167	0.4545	5	4	0.4938
12	WC208	(CCT)5	F:tcattccaccctcactcc R:tcgggaggttcacctacatc	168–194	0.3409	7	6	0.7402
13	WC228	(CAC)6	F:ccaccaccacctcaatatcc R:ctagtcgaggtgcagcatca	336–348	0.5455	5	5	0.5298
14	WC230	(ATC)6	F:ccagagcgattagatgcaca R:tgaattgccaaaagggtctc	310–349	0.3000	10	7	0.7949
15	TO249	(ATTT)5	F:acacggcttttgagacagt R:gagggtccataaccgaaca	108–127	0.3750	5	4	0.6452
16	TO250	(TGTA)6	F:tgctgctattgctgaccatc R:ggaccatggaccaagaaga	429–479	0.3684	6	6	0.7012
17	TCL262	(GA)7	F:ccttgccgttcggttatta R:catttgccgtcaaaagaaca	361–374	0.2857	6	6	0.7452
18	TCL279	(CTG)5	F:ttttccactcctccatcgac R:actcgaaccgtggaaatgac	414–428	0.5000	5	5	0.6257
19	TCL284	(TTC)5	F:cggtcggaaactcagctaac R:cagtgagggaattgagcat	152–174	0.2727	6	6	0.7495
20	TCL285	(GGA)6	F:gcgctctcctatcactgac R:cataacgtatccggcgatct	222–235	0.4000	5	5	0.6756
21	TCL290	(TGC)6	F:gcctcgaccattctggtaaa R:atgccccagatgaacagag	329–340	0.5000	4	4	0.5862
22	TO309	(GGT)5	F:ctcatgcacagaaggcacat R:acgaggtgctcgatactgct	186–218	0.6429	4	4	0.4503
Mean					0.4440	5.6818	5.2727	0.6254

^a A: major allele frequency; ^b Ng: number of genotypes; ^c Na: number of alleles; ^d PIC: polymorphism information content.

The number of alleles (Na) discovered at these loci varied from three to nine with an average of 5.27 alleles per locus. The PIC value illustrated the high degree of polymorphism at each locus, which ranged from 0.44 to 0.85 with a mean of 0.63. According to the previously-described criterion, the three categories were defined as high ($\text{PIC} > 0.5$), moderate ($0.25 < \text{PIC} < 0.5$), and low ($\text{PIC} < 0.25$) [34,35], and a total of 16 (72.73%) primers displayed a high level of PIC. In contrast to other plants, the Na and PIC values for the Welsh onion were lower than those reported in pear buds [36], 9.42 vs. 0.75, respectively. However, the values were higher than those for sweet potato (2.94 and 0.35, respectively) [25], cowpea (3.06 and 0.53, respectively) [37], and mung bean (3 and 0.34, respectively) [16].

2.4. The Results of UPGMA Analysis Using 22 Highly Polymorphic Markers among 22 *Allium* Accessions

The application of genic SSR markers to genetic diversity analysis and fingerprinting has recently been reported in several plant species [25,38–40]. In this study, The UPGMA (unweighted pair-group method with arithmetic average) cluster analysis divided the 22 *Allium* accessions into three groups with Nei's similarity coefficients ranging from 0.37 to 0.99 (Figure 2).

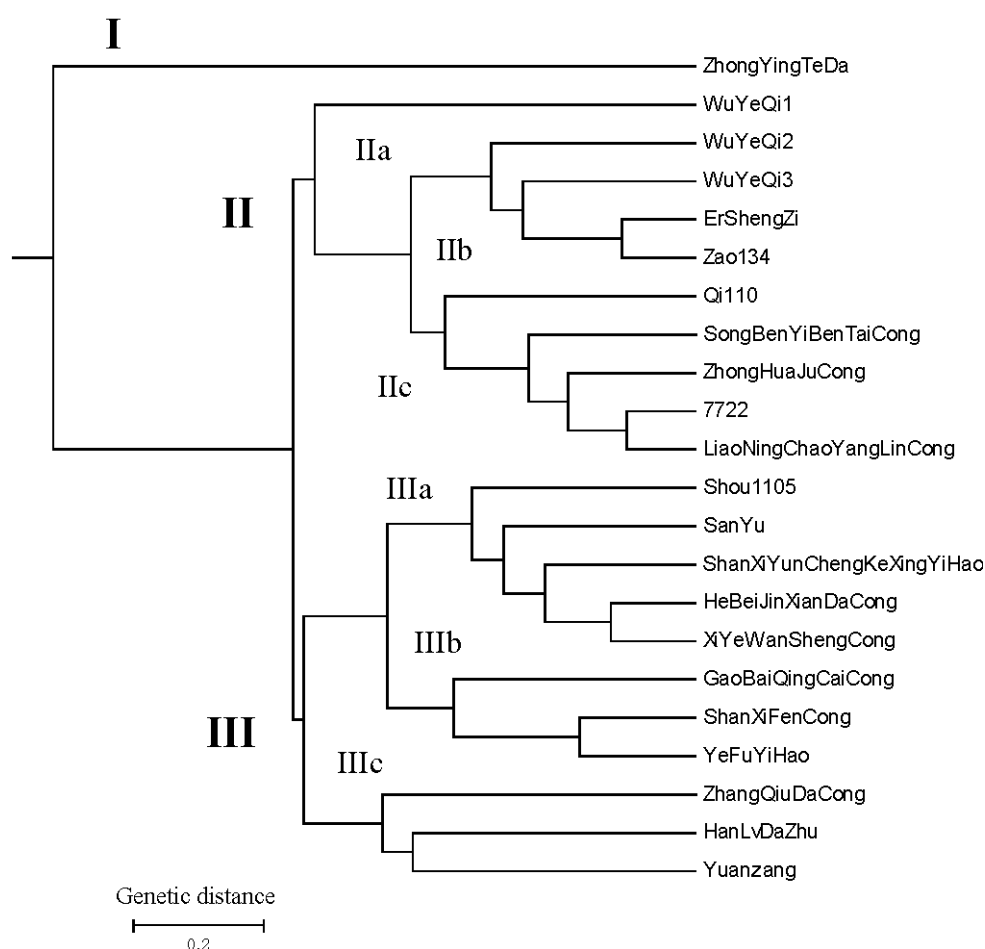


Figure 2. Dendrogram of 22 *Allium* accessions derived from UPGMA analysis of 22 highly polymorphic SSR markers.

Group I only contained “ZhongYingTeDa”, but Group II had three subgroups: Group IIa, containing only “WuYeQi”1; Group IIb, consisting of “WuYeQi”2, “WuYeQi”3, “ErShengZi”, and “Zao134”;

and Group IIc, comprising “Qi110”, “SongBenYiBenTaiCong”, “ZhongHuaJuCong”, “7722”, and “LiaoNingChaoYangLinCong”. In addition, Group III also had three subgroups: Group IIIa, consisting of “Shou1105”, “SanYu”, “ShanXiYunChengKeXingYiHao”, “HeBeiJinXian-DaCong”, and “XiYeWanShengCong”; Group IIIb, comprising “GaoBaiQingCaiCong”, “ShanXiFenCong”, and “YeFu YiHao”; and Group IIIc, containing “ZhangQiuDaCong”, “HanLvDaZhu”, and “Yuanzang”.

In recent years, transcriptome sequencing has proven to be one of the most efficient ways of developing genic SSR markers in many plant species, and many SSR markers based on transcriptome sequencing have been widely utilized in genetic diversity analysis, chromosome mapping and gene-based association studies [19,25,41,42]. Based on the results of our experiments, 22 genic SSR markers divided the 22 *Allium* accessions into three groups and seven subgroups using UPGMA cluster analysis. Although the resulting dendrogram could not sufficiently explain the genetic relationships of the accessions, it demonstrated the effectiveness of the 22 genic SSR markers. A possible reason for the inaccurate dendrogram was that the Welsh onion possesses various local cultivars and a complex genetic background and, furthermore, useful molecular markers from the Welsh onion transcriptome were lacking. Therefore, developing more and better molecular markers from the Welsh onion transcriptome is very important and valuable for the further analysis of *Allium* phylogenetic relationships.

3. Experimental Section

3.1. Plant Materials and DNA Extraction

The 22 *Allium* accessions used to test and verify the polymorphic primer in this study (Table 3) were local cultivars or inbred lines. Genomic DNA was extracted from multiple fresh young leaves of each line using the modified hexadecyl trimethyl ammonium bromide (CTAB) method, and the quality and quantity of the DNA were checked in 1% agarose gels and with a Thermo Scientific NanoDrop™ 2000C Spectrophotometer (Thermo Scientific NanoDrop, South Logan, Utah, USA), respectively.

3.2. Identification of SSR Loci from the Welsh Onion Transcriptome and Primer Design

Based on the 6206 unigenes obtained by the RNA-Seq approach in a previous study (RNA-Seq data have been uploaded to NCBI under accession number SSR1609126 and SSR1609976) [43], 5983 unigenes were searched for 6486 genic simple sequence repeat (SSR) loci, excluding 280 compound SSRs due to statistical complications, in this study using the MicroSATellite identification tool (MISA) [44]. The SSR loci containing repeat units of 1–5 nucleotides were identified, and the minimum SSR length criteria were identified as ten iterations for mononucleotide repeats, six iterations for dinucleotide repeats, and five iterations for the other repeat units.

Primer3 software, version 4.0, was used to design the polymerase chain reaction (PCR) primer in the flanking regions of these SSR loci [45], whose lengths were not less than 50 bp. Excluding mononucleotide and unsatisfactory compound SSRs, a total of 2710 unigenes were used to design primers, and the criteria for screening primer pairs were set to include an optimum primer length of 20 bp, PCR product sizes ranging from 100 to 500 bp, a GC percentage between 40% and 60%, and an annealing temperature ranging from 50 °C to 62 °C. Moreover, the secondary structure, such as hairpin, dimer, false priming, and cross-dimer, cannot emerge from forward and reverse primers.

3.3. PCR Amplification and Screening Polymorphic Primers

PCR amplifications were performed on a C₁₀₀₀TM thermal cycler with a 20 µL final reaction volume containing 20 ng template DNA (4 µL each of 5 ng/µL template DNA), 2 µL of 10 × buffer, 1.6 µL of 2 mM dNTP (containing Mg²⁺), 1 U *Taq* DNA polymerase and 1 µL each of 10 mM forward/reverse primer with the remaining volume composed of Milli-Q water (Milli-Q Advantage A10, Millipore, Bedford MA USA) as a supplement. The PCR amplifications were performed under the following conditions:

94 °C for 2 min, 10 cycles of 94 °C for 30 s, 62 °C–52 °C (reduce by 1 °C in each of the cycles) for 30 s, 72 °C for 30 s followed by 35 cycles of 94 °C for 30 s, 54 °C for 30 s, 72 °C for 30 s, and a final extension of 72 °C for 5 min.

A total of 311 primer pairs were synthesized by the Sangon Biotech Company (Shanghai, China), and the PCR products were checked in MultiNA (Shimadzu Biotech, Kyoto, Japan). In addition, polymorphic primers were obtained by manually reading the bands.

3.4. Validation of Polymorphic Primers

A total of 38 SSR markers were randomly selected from 91 polymorphic primers to validate the genetic diversity in 22 *Allium* accessions with an unweighted pair-group method with arithmetic average (UPGMA) dendrogram. The PCR products were obtained by using the same PCR amplification procedure, and PowerMarker version 3.25 was used to check the polymorphism information content (PIC), the number of alleles (Na), the major allele frequency (A), and the number of genotypes (Ng) [46]. Based on Nei's similarity coefficients, and the confidence level of the branch support evaluated by bootstrap analysis with 1000 replicates, the dendrogram was constructed to estimate the genetic relationship among 22 *Allium* accessions using MEGA 6.06.

4. Conclusions

To explore new molecular markers in the Welsh onion, this study attempted to develop genic SSR markers through transcriptome sequencing and resulted in 2525 mature SSR sequences. Moreover, because of their efficient polymorphism, reproducibility, and transferability, these SSR markers will provide new opportunities for germplasm identification, genetic diversity analysis, phylogenetic relationships studies, chromosome mapping, and even map-based cloning in the Welsh onion and other important related *Allium* species.

Acknowledgments

This work was supported by The National Natural Science Foundation of China (No. 31401873 to Yongqin Wang), the Natural Science Foundation of Beijing (No. 6132016), and the Science and Technology Innovation Capability Construction of the Beijing Academy of Agriculture and Forestry Sciences (No. KJCX20140111). The experiments were performed in the Key Laboratory of the Beijing Vegetable Research Center (BVRC) of the Beijing Academy of Agriculture and Forestry. The authors thank Congjuan Dong, Guanglei Sui and Huyu Liu (working at the BVRC) for their contributions to this study.

Author Contributions

Yongqin Wang, Changlong Wen and Lecheng Liu conceived and designed the experiments. Liuyi Yang performed the experiments and analyzed the data. Liuyi Yang and Changlong Wen wrote the paper. Hong Zhao, Qianchun Liu and Jingjing Yang contributed reagents/materials/analysis tools. All authors read and approved the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Takhtajan, A. *Flowering Plants Origin and Dispersal*; Oliver & Boyd: Edinburgh, UK, 1969; pp. 165–203.
2. Ford-Lloyd, B.V.; Armstrong, S.J. Welsh onion (*Allium wstulosum* L). In *Genetic Improvement of Vegetable Crops*; Kalloo, G., Bergh, B.O., Eds.; Pergamon Press: Oxford, UK, 1993; pp. 51–58.
3. Kumazawa, S.; Katsumata, H. Negi (Japanese bunching onion). In *Sosai-Engei Kakuron (Vegetable Crops)*; Kumazawa, S., Ed.; Yokendo Press: Tokyo, Japan, 1965; pp. 280–289.
4. Jiangsu New Medical College. Dictionary of traditional Chinese Medicine. In *Shanghai Science and Technology*; Shanghai Press: Shanghai, China, 1986; pp. 2316–2317.
5. Gai, S.P.; Meng, X.D.; Xu, L.J. Studies on molecular marker-assisted selection for sterile line and maintainer line in Welsh onion (*Allium fistulosum* L.). *Mol. Plant Breed.* **2004**, *2*, 223–228.
6. Gai, S.P. Development and Identification of DNA Molecular Markers Linked to Cytoplasmic Fertility Loci and Restore Genes and Markers Assisted Selection in Welsh Onion. Ph.D. Thesis, Shandong Agricultural University, Shandong, China, May 2002.
7. Gao, L.M. The Studies of Germ Plasm among Some *Allium* Lines and the Analysis of Main Characters. Master's Thesis, Shandong Agricultural University, Shandong, China, 20 May 2005.
8. Tsukazaki, H.; Yamashita, K.-I.; Yaguchi, S.; Masuzaki, S.; Fukuoka, H.; Yonemaru, J.; Kanamori, H.; Kono, I.; Hang, T.T.M.; Shigyo, M.; *et al.* Construction of SSR-based chromosome map in bunching onion (*Allium fistulosum*). *Theor. Appl. Genet.* **2008**, *117*, 1213–1223.
9. Song, Y.-S.; Suwabe, K.; Wako, T.; Ohara, T.; Nunome, T.; Kojima, A. Development of Microsatellite Markers in Bunching Onion (*Allium fistulosum* L.). *Breed. Sci.* **2004**, *54*, 361–365.
10. Tsukazaki, H.; Yaguchi, S.; Sato, S.; Hirakawa, H.; Katayose, Y.; Kanamori, H.; Kurita, K.; Itoh, T.; Kumagai, M.; Mizuno, S.; *et al.* Development of transcriptome shotgun assembly-derived markers in bunching onion (*Allium fistulosum*). *Mol. Breed.* **2015**, *35*, 55, doi:10.1007/s11032-015-0265-x.
11. Tsukazaki, H.; Yamashita, K.-I.; Kojima, A.; Wako, T. SSR-tagged breeding scheme for allogamous crops: A trial in bunching onion (*Allium fistulosum*). *Euphytica* **2009**, *169*, 327–334.
12. Powell, W.; Machray, G.C.; Provan, J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1996**, *1*, 215–222.
13. Tóth, G.; Gáspári, Z.; Jurka, J. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Res.* **2000**, *10*, 967–981.

14. Tsukazaki, H.; Fukuoka, H.; Song, Y.-S.; Yamashita, K.; Wako, T.; Kojima, A. Considerable heterogeneity in commercial F1 varieties of bunching onion (*Allium wstulosum*) and the proposal of a breeding scheme for conferring genetic traceability using SSR markers. *Breed. Sci.* **2006**, *56*, 321–326.
15. Tsukazaki, H.; Nunome, T.; Fukuoka, H.; Kanamori, H.; Kono, I.; Yamashita, K.; Wako, T.; Kojima, A. Isolation of 1796 SSR clones from SSR-enriched DNA libraries of bunching onion (*Allium wstulosum*). *Euphytica* **2007**, *157*, 83–94.
16. Gupta, S.K.; Bansal, R.; Gopalakrishna, T. Development and characterization of genic SSR markers for mungbean (*Vigna radiata* (L.) Wilczek). *Euphytica* **2014**, *195*, 245–258.
17. Kuhl, J.C.; Cheung, F.; Yuan, Q.; Martin, W.; Zewdie, Y.; MaCallum, J.; Catanach, A.; Rutherford, P.; Sink, K.C.; Jenderk, M.; *et al.* A unique set of 11,008 onion expressed sequence tags reveals expressed sequence and genomic differences between the monocot orders Asparagales and Poales. *Plant Cell* **2004**, *16*, 114–125.
18. Martin, W.J.; McCallum, J.; Shigyo, M. Genetic mapping of expressed sequences in onion and in silico comparisons with rice show scant colinearity. *Mol. Genet. Genom.* **2005**, *274*, 197–204.
19. Duangjit, J.; Bohanec, B.; Chan, A.P.C.; Town, C.D.C.; Havey, M.J.D. Transcriptome sequencing to produce SNP-based genetic maps of onion. *Theor. Appl. Genet.* **2013**, *127*, 535–547.
20. Zhou, X.-J.; Wang, Y.-Y.; Xu, Y.-N.; Yan, R.-S.; Zhao, P.; Liu, W.-Z. *De novo* characterization of flower bud transcriptomes and the development of EST-SSR markers for the endangered tree *Tapiscia sinensis*. *Int. J. Mol. Sci.* **2015**, *16*, 12855–12870.
21. Wu, T.; Luo, S.; Wang, R.; Zhong, Y.; Xu, X.; Lin, Y.E.; He, X.; Sun, B.; Huang, H. The first Illumina-based *de novo* transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Mol. Breed.* **2014**, *34*, 1437–1447.
22. Chen, X.-B.; Xie, Y.-H.; Sun, X.-M. Development and Characterization of polymorphic Genic-SSR markers in *Larix kaempferi*. *Molecules* **2015**, *20*, 6060–6067.
23. Fischer, D.; Bachmann, K. Onion microsatellites for germplasm analysis and their use in assessing intra- and interspecific relatedness within the subgenus *Rhizirideum*. *Theor. Appl. Genet.* **2000**, *101*, 153–164.
24. Yao, L.H.; Zheng, X.Y.; Cai, D.Y.; Gao, Y.; Wang, K.; Cao, Y.F.; Teng, Y.W. Exploitation of *Malus* EST-SSRs and the utility in evaluation of genetic diversity in *Malus* and *Pyrus*. *Genet. Resour. Crop. Evol.* **2010**, *57*, 841–851.
25. Wang, Z.Y.; Li, J.; Luo, Z.X.; Huang, L.F.; Chen, X.L.; Fang, B.P.; Li, Y.J.; Chen, J.Y.; Zhang, X.J. Characterization and development of EST-derived SSR markers in cultivated sweetpotato (*Ipomoea batatas*). *BMC Plant Biol.* **2011**, *11*, 139.
26. Zhai, L.L.; Xu, L.; Wang, Y.; Cheng, H.; Chen, Y.L.; Gong, Y.Q.; Liu, L.W. Novel and useful genic-SSR markers from *de novo* transcriptome sequencing of radish (*Raphanus sativus* L.). *Mol. Breed.* **2014**, *33*, 611–642.
27. Jena, S.N.; Srivastava, A.; Rai, K.M.; Ranjan, A.; Singh, S.K.; Nisar, T.; Srivastava, M.; Bag, S.K.; Mantri, S.; Asif, M.H.; *et al.* Development and characterization of genomic and expressed SSRs for levant cotton (*Gossypium herbaceum* L.). *Theor. Appl. Genet.* **2012**, *124*, 565–576.
28. Wang, Y.W.; Samuels, T.D.; Wu, Y.Q. Development of 1030 genomic SSR markers in switchgrass. *Theor. Appl. Genet.* **2011**, *122*, 677–686.

29. Dutta, S.; Kumawat, G.; Singh, B.P.; Gupta, D.K.; Singh, S.; Dogra, V.; Gaikwad, K.; Sharma, T.R.; Raje, R.S.; Bandhopadhyaya, T.K.; *et al.* Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea (*Cajanus cajan* (L.) Millspaugh). *BMC Plant Biol.* **2011**, *11*, 17, doi:10.1186/1471-2229-11-17.
30. Zhang, H.Y.; Wei, L.B.; Miao, H.M.; Zhang, T.; Wang, C.Y. Development and validation of genic-SSR markers in sesame by RNA-Seq. *BMC Genomics* **2012**, *13*, 316, doi:10.1186/1471-2164-13-316.
31. Silva, P.I.; Martins, A.M.; Gouvea, E.G.; Pessoa-Filho, M.; Ferreira, M.E. Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. *BMC Genomics* **2013**, *14*, 17.
32. Iorizzo, M.; Senalik, D.A.; Grzebelus, D.; Bowman, M.; Cavagnaro, P.F.; Matvienko, M.; Ashrafi, H.; Deynze, A.V.; Simon, P.W. *De novo* assembly and characterization of the carrot transcriptome reveals novel genes, new markers, and genetic diversity. *BMC Genom.* **2011**, *12*, 389, doi:10.1186/1471-2164-12-389.
33. Ricroch, A.; Yockteng, R.; Brown, S.C.; Nadot, S. Evolution of genome size across some cultivated *Allium* species. *Genome* **2005**, *48*, 511–520.
34. Yadav, H.K.; Ranjan, A.; Asif, M.H.; Mantri, S.; Sawant, S.V.; Tuli, R. EST-derived SSR markers in *Jatropha curcas* L.: Development, characterization, polymorphism, and transferability across the species/genera. *Tree Genet. Genomes* **2011**, *7*, 207–219.
35. Yu, L.J.; Luo, Y.F.; Liao, B.; Xie, L.J.; Chen, L.; Xiao, S.; Li, J.T.; Hu, S.N.; Shu, W.S. Comparative transcriptome analysis of transporters, phytohormone and lipid metabolism pathways in response to arsenic stress in rice (*Oryza sativa*). *New Phytol.* **2012**, *195*, 97–112.
36. Yue X.Y.; Liu G.Q.; Zong Y.; Teng Y.W.; Cai D.Y. Development of genic SSR markers from transcriptome sequence of pear buds. *J. Zhejiang Univ. Sci. B* **2014**, *15*, 303–312.
37. Gupta, S.K.; Gopalakrishna, T. Development of unigene-derived SSR markers in cowpea (*Vigna unguiculata*) and their transferability to other *Vigna* species. *Genome* **2010**, *53*, 508–523.
38. Barchi, L.; Lanteri, S.; Portis, E.; Acquadro, A.; Vale, G.; Toppino, L.; Rotino, G.L. Identification of SNP and SSR markers in eggplant using RAD tag sequencing. *BMC Genomics* **2011**, *12*, 304.
39. Triwitayakorn, K.; Chatkulkawin, P.; Kanjanawattanawong, S.; Sraphet, S.; Yoocha, T.; Sangsrakru, D.; Chanprasert, J.; Ngamphiw, C.; Jomchai, N.; Therawattanasuk, K.; *et al.* Transcriptome sequencing of *Hevea brasiliensis* for development of microsatellite markers and construction of a genetic linkage map. *DNA Res.* **2011**, *18*, 471–482.
40. Brandon, S.; Diego, F.; Tierney, B.; Eric, W.; James, P.; Nicholi, V.; Shawn, S.; Juan, Z. Development and validation of 697 novel polymorphic genomic and EST-SSR markers in the American Cranberry (*Vaccinium macrocarpon* Ait.). *Molecules* **2015**, *20*, 2001–2013.
41. Zeng, S.H.; Xiao, G.; Guo, J.; Fei, Z.J.; Xu, Y.Q.; 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 Genomics* **2010**, *11*, 94.
42. Zhang, J.N.; Liang, S.; Duan, J.L.; Wang, J.; Chen, S.L.; Cheng, Z.S.; Zhang, Q.; Liang, X.Q.; Li, Y.R. *De novo* assembly and characterisation of the transcriptome during seed development, and generation of genic-SSR markers in peanut (*Arachis hypogaea* L.). *BMC Genomics* **2012**, *13*, 90, doi:10.1186/1471-2164-13-90.

43. Liu, Q.C.; Wen, C.L.; Zhao, H.; Zhang, L.Y.; Wang, J.; Wang, Y.Q. RNA-Seq Reveals Leaf Cuticular Wax-Related Genes in Welsh Onion. *PLoS ONE* **2014**, *9*, e113290, doi:10.1371/journal.pone.0113290.
44. Dai, H.Y.; Han, G.F.; Yan, Y.J.; Zhang, F.; Liu, Z.C.; Li, X.M.; Li, W.R.; Ma, Y.; Li, H.; Liu, Y.X.; *et al.* Transcript assembly and quantification by RNA-Seq reveals differentially expressed genes between Soft-Endocarp and Hard-Endocarp Hawthorns. *PLoS ONE* **2013**, *8*, e72910, doi:10.1371/journal.pone.0072910.
45. You, F.M.; Huo, N.; Gu, Y.Q.; Luo, M.-C.; Ma, Y.; Hane, D.; Lazo, G.R.; Dvorak, J.; Anderson, O.D. BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics* **2008**, *9*, 253, doi:10.1186/1471-2105-9-253
46. Liu, K.; Muse, S.V. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* **2005**, *21*, 2128–2129.

© 2015 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 license (<http://creativecommons.org/licenses/by/4.0/>).