Development of InDel Markers and Fingerprinting of Bitter Gourd (*Momordica Charantia*) Based on Whole Genome Re-sequencing

Liwen Su

Guangxi University Guangxi Agriculture College

Lianlian Ma

Guangxi University Guangxi Agriculture College

Xiaochun Huang

Guangxi University Guangxi Agriculture College

Wenting Wu

Guangxi University Guangxi Agriculture College

Haixuan Lv

Guangxi University Guangxi Agriculture College

Zhikui Cheng

Guangxi University Guangxi Agriculture College

Jiquan Gou

Guangxi University Guangxi Agriculture College

Yong Chen

Guangxi University Guangxi Agriculture College

Peng Wang

Guangxi University Guangxi Agriculture College

Zhengguo Liu (✉ liu-zhengguo@126.com)

Guangxi University Guangxi Agriculture College <https://orcid.org/0000-0001-8910-5164>

Research Article

Keywords: Bitter gourd, Resequencing, InDel marker, Genetic diversity, Fingerprint, Germplasm

Posted Date: August 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-810603/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Seed Science and Technology on January 1st, 2022. See the published version at <https://doi.org/10.15258/sst.2022.50.1.11>.

Abstract

Bitter gourd (*Momordica charantia* L.) is one of the most important vegetable crops in many Asian and African countries. Here, InDel markers developed by comparing whole genome re-sequencing data were used to analyze the genetic diversity of a bitter gourd germplasm sample from various geographical origins. To verify the reliability of the set of InDel markers identified, 220 pairs of InDel primers were designed. The primers were preliminarily detected by 8% polyacrylamide gel electrophoresis, and 25 pairs of primers with better polymorphism were screened. Using the 25 primer combinations, the 53 bitter gourd accessions were effectively distinguished and the InDel fingerprint of DNA was constructed. Concomitantly, the degree of purity of different crosses was determined based on the differences in specific bands among genotypes. The unweighted pair-group method with arithmetic means showed that the 53 bitter gourd materials may be divided into three groups, with a similarity coefficient of 0.645 as the threshold. Therefore, this study can provide many InDel markers for genotypic identification, genetic relationship analysis, and genetic map construction of bitter gourd.

1. Introduction

Bitter gourd is a common, annual, climbing herbaceous medicinal plant, and a melon vegetable suitable for cultivation in temperate and tropical regions. Bitter gourd originated in tropical Africa [1] and is mainly distributed in tropical and subtropical regions of Asia and Africa, as well as in the Amazon region of Brazil, the Caribbean, and South Africa [2, 3].

Several wild bitter gourd genetic resources have been found in the southern foothills of the Himalayas. Particularly, India is home to some of the wildest bitter gourd germplasm, followed by Southeast Asian countries, such as Thailand and Myanmar, and the Yunnan and Hainan areas in China[4]. Wild bitter gourd fruit is small, thin, and prickly; furthermore, because of its strong bitter taste, it is difficult to eat but after years of domestication, modern cultivated bitter gourd varieties have developed thick and long flesh, an increased weight, and a milder bitter taste and are more suitable for eating. Expanding research efforts have confirmed a high nutritional value and medicinal potential of bitter gourd for reducing blood glucose levels in patients with diabetes [5] .

Genetic resources provide abundant genetic variation for cultivar improvement as well as genetic and biological research. Germplasm is an important part of genetic diversity, the basic premise of agricultural origin and development, and the raw material for various breeding pathways, which largely determine the breeding effect [6]. As in any other horticultural crop, any major breakthrough in bitter gourd breeding depends on the development and utilization of key germplasm. Breeding high-quality bitter gourd varieties requires access to extensive genetic resources. In addition, genetic diversity and genetic relationship analysis are the main concerns while evaluating germplasm resources. Genetic diversity is a reflection of plant adaptability to environmental changes, which in turn is reflected at the morphological, cytological, physiological, and biochemical levels [7].

Analysis of genetic diversity and genetic relationship of agronomic traits within a germplasm sample is of great significance for the development and utilization of germplasm resources. Thus, for example, Huang et al. [8] used 28 morphological traits to conduct cluster analysis of 33 bitter melon accessions, and divided them into three groups, namely, a wildtype group, a dense tumor, small-fruit type group and a long-fruit type group. They found that there were obvious regional differences among groups. Further, the genetic distance between Chinese bitter melon germplasm and germplasm resources from India and Southeast Asia was relatively large. Some bitter melon varieties reflect differences in morphological characteristics that are quite different from the classification of morphological experience.

More recently, Kang et al. [9] used ISSR molecular markers to study the genetic diversity among 48 bitter melon samples. They found that Jiangxi variety 608 and Fujian variety 615 were similar in skin maturity, color, and shape but the characteristics of the fruit surface differed. Similarly, Dhillon et al. [10] used 50 pairs of SSR primers to conduct cluster analysis on 114 bitter melon accessions from Asia. The genetic similarity coefficients ranged between 0.61 and 0.88, and they were divided into four groups at a threshold value of 0.66. Further, Wang [11] used morphological markers and SRAP and SSR molecular markers to conduct cluster analysis on 14 accessions of *Momordica charantia*. He found that the similarity of clustering based on morphological markers and molecular markers reached 78.57%; furthermore, the genetic differences among the 14 accessions were not significant. However, the clustering results obtained based on the two molecular markers were very close, indicating that molecular marker-based clustering might be more reliable. Therefore, it is better to study the diversity among germplasm accessions using molecular markers than using morphological characteristics, as the former can accurately indicate the genetic relationship of bitter melon at the molecular level and determine the extent of homology.

Some varieties with similar morphological characteristics can be clustered together, thus confirming the correctness of the traditional morphological classification framework, which can be supplemented and refined by clustering agronomic traits based on differences at the DNA level. Therefore, the evaluation, innovation, and utilization of molecular markers within bitter melon germplasm certainly warrants further support.

Insertion-deletion length polymorphism (InDel) refers to the difference between two materials relative to the whole species genome. Compared with each other, there will much likely be a certain number of nucleotide insertions or deletions in the genome of one material which do not exist in the other. As a new generation of markers, InDel markers are widely distributed, numerous, highly polymorphic, stable in variation, low in price, and co-dominant in the genome of a species. To date, InDel markers have been successfully applied to the study of rice, cucumber, and pepper germplasm, among other crops. For example, Hayashi et al. [12] developed a set of InDel markers for the blast-resistance gene in rice. In turn, Li et al. [13] explored InDel loci based on re-sequencing results and used 134 pairs of InDel primers to detect the effectiveness of 16 typical cucumber genotypes. Their results showed that 116 pairs of InDel primers fully revealed the diversity and specificity of the 16 genotypes. Similarly, Li et al. [14] developed InDel markers for pepper genetic mapping using two inbred lines, BA3 and B702, with genome-wide re-

sequencing; they obtained a linkage genetic map based on the InDel markers developed, which consisted of 12 linkage groups with a genetic distance of 1178.01 cM and an average distance of 5.01 cM between bin markers.

With the release of the bitter melon genome, SSR markers [15] and InDel markers [16] have been developed for the whole species genome. However, InDel markers have not been reported for genetic diversity analysis of bitter melon germplasm. According to the whole genome series of bitter melon Dali-11, two materials, namely, No. 12 from Guangxi and H-13 from Guangdong were re-sequenced to develop InDel markers. The fruits of these two materials are green and straight tumor, but their sources and fruit lengths are different. The fruit of No. 12 is longer (35–40 cm) and shows late maturity, while the fruit of H-13 is shorter (approximately 20 cm) and shows early maturity. There is strong heterosis between the two materials, which is important to re-sequence and develop InDel markers. Concomitantly, this set of markers was used to screen out the core markers to construct the molecular fingerprints of 53 bitter melon germplasm accessions from different regions in China, as well as other countries, which can be used for variety identification and for determination of hybrid seed purity. Finally, the genetic relationship of the tested materials was analyzed to provide a theoretical reference for the selection of parents in the process of breeding superior varieties.

2. Materials And Methods

2.1. Plant materials

All 53 bitter melon (Table S1) accessions and seeds of three hybrids, namely, Zhongyu, Tiantianhao 11, and Tiantianhao 12, were provided by the Liu Zhengguo Research Group, of the College of Agriculture at Guangxi University. Two of the materials are from Thailand, one from Malaysia, and the rest from various regions of China.

Genomic DNA was extracted from young leaves using the CTAB method after modifications [17]. The concentration of DNA was determined using an K5800 ultramicro spectrophotometer (BJKA, Yima Optoelec Co., Ltd., China). Using an OD_{260}/OD_{280} of 1.8–2.0 as the reference value, any DNA sample with OD_{260}/OD_{280} values <1.7 , or >2.0 , was purified to 1.8–2.0 and diluted to $25 \text{ ng } \mu\text{L}^{-1}$.

2.2. DNA library construction sequencing and resequencing data quality assessment

After the genomic DNA of bitter melon accessions No. 12 and H-13 was qualified, the DNA was fragmented using ultrasonication; then, the fragmented DNA was purified, repaired at the end, added with A at the 3' end, and connected to the sequencing joint. The fragment size was selected using agarose gel electrophoresis, and the sequencing library was formed by PCR amplification. The constructed library was subjected to library quality inspection, and the qualified library was sequenced by Illumina. Sequenced or raw reads containing a joint, i.e., low-quality reads, were filtered to obtain clean reads to ensure the quality of subsequent information analysis.

2.3. Sequencing comparative analysis

The sequencing reads obtained by re-sequencing need to be relocated to the reference genome to analyze the existent variation. The BWA [18] software was used for the comparison between the short sequence and the reference genome obtained by high-throughput sequencing (as per the sequencing platform of Illumina) with the reference genome, and information such as sample comparison rate was counted. Comparison rate: Clean Reads that could be located on the reference genome accounted for all clean reads. If the reference genome is selected properly and there is no pollution in the relevant experimental process, the comparison rate of sequencing reads will be higher than 70%. In addition, the alignment rate is related to the genetic relationship between the sequenced species and the reference genome, the assembly quality of the reference genome, and the sequencing quality of reads. The closer the species, the more complete the assembly of the reference genome; furthermore, the higher the quality of the sequencing reads, the more reads can be located in the reference genome, and the higher the alignment rate.

According to the position information of small InDel loci on the reference genome detected by the sample, compared with the gene and CDS location information of the reference genome, we can annotate whether the InDel locus occurs in the intergenic region, in the gene region, or in the CDS region, and whether it is a frame-shift mutation. The annotation of small InDel is performed by the SnpEff software. An InDel with a transcoding mutation may lead to changes in gene function.

2.4. Screening and analysis of InDel primers

InDel primers were designed according to the predicted InDel loci while re-sequencing data, and 20 pairs of InDel primers were randomly selected from each chromosome, which were evenly distributed on the chromosome, with a total of 220 pairs covering the whole genome of bitter melon. The primers were synthesized by Beijing Qingke Biotechnology Co., Ltd.

The optimized amplification system comprised a total volume of 12 μ L containing 5 μ L 2 $\times$ T5 Super PCR Mix (PAGE), 2 μ L genomic DNA, 1 μ L upstream and downstream primers, and 4 μ L ddH₂O. The reaction process comprised pre-denaturation at 94°C for 45 min; denaturation at 94°C for 30 s, renaturation at 52°C for 30 s, extension at 72°C for 30 s, 35 cycles; then, extension at 72°C for 5 min, and preservation at 16°C. PCR amplification products were detected by 8% polyacrylamide gel electrophoresis (DYY-6D, Beijing Liuyi Biotechnology Co., Ltd, China) performed at 300 V and 150 mA for 150 min. After electrophoresis, silver staining, statistical band analysis, and imaging were performed.

2.5. Construction of the molecular fingerprints of 53 bitter melon genetic accessions

Twenty-five pairs of InDel primers were used to construct the molecular fingerprint of the bitter melon germplasm sample under study. Then, according to the separation of PCR amplification products of the 53 bitter melon accessions analyzed by PAGE, the assignment rules were as follows. Regarding the same

migration position, strong bands were labeled “1,” while bands without, or with weak bands, were denoted as “0.” Thus, the assignment of bands amplified by 25 marker primers in each variety was combined into a series of numbers, which was used as the molecular identity card of the corresponding variety.

According to the rules of complementary bands between parents and their different hybrids, suitable primers were screened from the constructed fingerprint to identify the purity of the tested samples. Based on the results of InDel molecular marker PAGE, the sample type was determined according to whether the electrophoresis band of the sample was complementary to the electrophoresis band of the parents. The seed purity of Zhongyu, Tiantianhao 11, and Tiantianhao 12 was determined by the extent of homology of specific bands shared with each parent.

2.6. Data analysis

The PopGen32 software [19] was used to calculate Nei’s gene diversity index (H) and the Shannon-Wiener diversity index (I). Polymorphism information content (PIC) was calculated using the formula $PIC = (X/X - 1) (1 - P_{ij}^2)$, where X represents the number of samples and P_{ij} represents the frequency of the i^{th} site that appears in the j^{th} gene. The amplified bands were statistically analyzed, and the 0,1 matrix (0 represents no band and 1 represents band) was stored in an Excel table according to the band type. The genetic similarity matrix was calculated by the Ntsys 2.10 software, and cluster analysis was conducted according to the unweighted pair-group method of arithmetic average (UPGMA) [20].

3. Results

3.1. Quality assessment of re-sequencing data

Table 1 shows the analysis of the sequencing data obtained and summarizes the results of sequencing quality analysis. This analysis for 20.05 Gbp Clean Data revealed high sequencing quality ($Q_{20} \geq 96.85\%$, $Q_{30} \geq 91.49\%$) and a GC content of approximately 36.80–36.85%; furthermore, the quality of the sample base distribution in the first four bases and the last 10 bases will be lower than the quality of the intermediate-base sequencing, but the quality values Q are higher than 30%; AT, CG base pairs basically did not separate, the curve is gentle, and the sequencing results are normal (Table 1, Fig. 1). According to these results, the genome sequencing data is sufficient, C and G are distributed in the normal range, and sequencing quality is high, which altogether, support the subsequent analysis. The Dali-11 analysis of the reference genome of bitter melon showed that the average comparison rate between the sample and the reference genome was 99.22%, indicating that the library was constructed normally and without pollution. The average coverage depth was 26X, and the genome coverage was 92.71%, at least one-base coverage. The comparison results are normal (Table 2), which might be used for subsequent detection of variations and correlation analysis. According to the coverage depth of each chromosome site, it can be seen that the genome is evenly covered, indicating that sequencing randomness was adequate (Fig. 1).

Table 1 Evaluation statistics of sample sequencing data.

Variety	Clean_Reads	Clean_Base	Error rate (%)	Q20 (%)	Q30 (%)	GC (%)
H-13	32508450	9738384766	0.01	96.87	91.63	36.80
No. 12	34425836	10313295526	0.01	96.85	91.49	36.85

Table 2

Variety	Clean_Reads	Clean_Base	Error rate (%)	Q20 (%)	Q30 (%)	GC (%)
H-13	32508450	9738384766	0.01	96.87	91.63	36.80
No. 12	34425836	10313295526	0.01	96.85	91.49	36.85
Analysis of sequencing depth and coverage.						

3.2. Detection of small InDel between samples and reference genomes

A total of 59098 and 73683 InDel were detected in the whole genome of bitter melon H-13 and No. 12, respectively. The number of InDel in the coding region was 540 and 671, respectively. Further comparative analysis revealed that the number of InDel across the genome between H-13 and No. 12 was 96348. The InDel locus length distribution according to the length of InDel in the CDS region and the whole genome of the samples is shown in Fig. 3. The length distribution trend of InDel loci across the whole genome of the two varieties was basically the same (Fig. 2). The number of InDel loci decreased with increasing InDel fragment length, except for the gene region. There are more types of +1, +1, +2, +2, +3, and +3 mutations in the CDS region, and relatively few types of +3 and +3 mutations in the genome. Using the SnpEff software to annotate the whole genome small InDel of H-13 and No. 12 (Fig. 2), we found 301 coding-shift mutations in the CDS region of small InDel of H-13, which accounted for 56.90% of all mutations. Followed by the deletion and insertion of codons, a total of 152 (28.73%), we found fewer mutations in start codons and stop codons, a total of 14 (2.65%). Small InDel in the CDS region of No. 12 were also coding-shift mutations accounting for 345 (52.11%), a total of 207 codon deletions and insertions (31.27%), and 28 mutations (2.72%) of the fewer start codons and stop codons. These small InDel caused a total of 1211 gene mutations (Table 3).

Table 3

Variety	Total_reads	Mapped (%)	Properly_mapped (%)	Mean depth	Coverage 1X (%)	Coverage 5X (%)	Coverage 10X (%)
H-13	65016900	99.12	93.63	26	92.64	92.32	91.38
No. 12	68851672	99.33	93.18	27	92.78	92.5	91.74
Genome-wide and coding region InDel statistics.							

3.3. Polymorphism analysis of InDel markers

Twenty representative bitter gourd accessions were used as test materials to comprehensively reveal the genetic background of the bitter gourd germplasm collected by our research group; 220 pairs of InDel primers were synthesized for polymorphism detection and 25 pairs of primers with better polymorphism were screened (Table S2). Subsequently, 25 pairs of polymorphic InDel primers screened were used to identify 53 bitter gourd accessions and 49 polymorphic loci were detected. Each pair of primers produced an average of 1.96 polymorphic loci, and each pair of primers produced 1 ~ 3 polymorphic loci. The amplified bands of primers MC02-1, MC05-2, and MC011-4 were clear and polymorphic (Fig. 3). All other polymorphic primers were effective. The PIC of 25 InDel primers ranged from 0.473 to 0.700, with an average of 0.590. The PIC value of primer MC05-13 was the highest and that of primer MC05-2 was the lowest. Shannon's diversity index ranged from 0.196 to 0.693 with an average of 0.600. Nei's gene diversity index (H) ranged from 0.141 to 0.500, with an average of 0.430. Based on these findings, we concluded that these 25 pairs of InDel markers showed adequate polymorphism with potential for use in the molecular identification of bitter gourd varieties and in determining the degree of purity of a given cross.

3.4. Germplasm identification and molecular fingerprint of bitter gourd

The InDel amplification of 53 bitter gourd germplasm accessions was statistically analyzed, and the number and specificity of bands amplified by each primer were quite different. Thus, for example, primer MC01-5 effectively distinguished between 16 – 1 and LM-3-1; primer MC02-1 effectively distinguished between 10-2-5-4, KS-1, and LH-3; primer MC05-2 effectively distinguished between 16 – 1, LM-3-1, PE-1, 10-2-1, and PD-1; primer MC11-4 effectively distinguished between I-7, 3-1-6, and No. 8; furthermore, the combination of these 25 primers effectively identified all the germplasm accessions under study. According to the principle of using the smallest primer combination to identify the largest possible number of differing genotypes, the DNA fingerprints of 53 bitter gourd accessions were constructed using

the combination of these 25 primers (Table S3) for the identification of the accessions in the germplasm sample under study.

Self-pollination of parents producing false hybrids sometimes occurs during hybrid seed production. According to the amplification of characteristic heterobelts by each primer in the fingerprint, our results showed that primers MC02-1, MC05-2, and MC11-4 effectively determined the degree of hybrid purity in different crosses. As shown in Fig. 4, primer MC02-1 helped to identify the F₁ hybrid produced with No. 12 as female parent and H-13 as male parent. In turn, the results showed that numbers A14 and A29 are actually false hybrids, and primer MC05-2 helped to identify the hybrid produced with 10-2-1 as female parent and 10-2-5-3 as male parent. The commercial name given to this hybrid is Tiantianhao 11. The identification results showed that this is not a false hybrid. Primer MC11-4 was used to identify the hybrid produced with 3-1-6-2-1 as female parent and 10-2-5-3 as male parent. The commercial name given to this hybrid was Tiantianhao 12, and again, the identification results showed that this is not a false hybrid plant.

3.5. Genetic relationship analysis of bitter melon germplasm

The genetic similarity coefficient of Jaccard was calculated by the NTSYS-pc 2.10 analysis software, and the UPGMA method was used for cluster analysis to construct the resulting phylogenetic tree (Fig. 5). The results showed that the 53 bitter melon germplasm accessions analyzed clustered into three groups with a similarity coefficient of 0.645 as threshold. The first group included three accessions, namely, 16 – 1, MX-1, and LH-4, which accounted for 5.7% of the total. The second group included 40 accessions: 10-2-5-3, 10-2-5-4, I-7, PE-1, BO5367, 9-1-4-3, I-4-1-2-7, KS-1, BO-1, LH-3, I-4-1-2-6, 10-2-1, DJ-7-1, I-9-6-1, ITY-1, I-6-3, H-13, YI-1-2-1, BWB-1-1, BH-2, LH-1, HYT, PD-1, YHF, No. 5, HX-1, Yuxiu, Huayu 3, No. 58, HY-1, XHY, Huabi 2, BP6201, I-13-1-1, YXH, LI-1-1, BP6221, HXU, BP6214, and HB-1, altogether accounting for 75.5% of the total. Finally, the third group included 10 accessions: LM-3-1, LM-4-1, DJ-9-3, LM-1-1, BBL-1-2, 2-3-1, BP6237, 3-1-6, No. 12, and No. 8, which altogether accounted for 18.8% of the total.

4. Discussion

4.1. InDel markers are of great value in genetic diversity analysis and genetic map construction

Our experiments demonstrated the feasibility and convenience of using genome re-sequencing data to verify InDel polymorphism. The identified InDel loci could be validated to produce informative genetic markers with reliable, high polymorphism rates. More importantly, as the coordinates of InDels were known relative to the reference genome, it was possible to develop genetic markers in specific genome regions to help in effectively constructing genetic maps and using them for fine mapping of bitter melon

accessions No. 12 and H-13. The number of InDel detected in the whole genome of bitter gourd genotypes H-13 and No. 12 were 59098 and 73683, respectively, while the number of InDel detected across these two entire genomes together was 96348. In addition, most InDels were located in the intergenic region, followed by the upstream region of the gene (within 5 K), the downstream region of the gene (within 5 K), the intron, the gene, and the gene coding sequence. The small InDel in the CDS region also accounted for a large proportion of code-shift mutations. The InDel with code-shift mutations will lead to the change of the protein triplet codon reading-frame, which in turn may lead to a change of gene function. Therefore, InDel markers are of great value in genetic diversity analysis [21], high-density genetic map construction [22], and gene mapping [23] .

The genetic diversity of 53 bitter gourd germplasm accessions was analyzed by 25 pairs of InDel markers selected in the experiments reported herein. The results showed that the average PIC was 0.590, and that the selected primers were rich in polymorphism and might provide sufficient genetic information. The specific molecular fingerprints of the 53 bitter gourd accessions were constructed using the bands amplified by 25 pairs of primers, which was the external manifestation of the characteristic molecular structure and is beneficial to the identification of varieties in the process of plant variety protection. Simultaneously, specific differences in band profiles between varieties can be used to identify the degree of seed purity of the hybrid generation obtained by crossing the two parental varieties. In this study, three pairs of InDel marker primers were used to determine the degree of seed purity of the three hybrids. Therefore, the 53 characteristic branch fingerprints constructed can help in identifying varieties and help determine the degree of seed purity in large-scale seed production. Thus, seed production units may determine seed purity at any time to ensure seed quality to promote the adoption of a variety in the market.

4.2. UPGMA-cluster analysis of germplasm is useful in bitter gourd breeding

Based on the correlation between molecular genetic distance and heterosis, the level of genetic diversity between parents is considered a means to predict hybrid performance and heterosis [24, 25]. UPGMA-cluster analysis results showed that, at a genetic similarity coefficient of 0.645, the 53 bitter gourd materials tested here can be divided into three groups. The observed clustering had certain correlation with fruit color, fruit nodule characteristics, and geographical origin, consistently with previous results [26, 27]. The three materials in Group I had green peel; two of them were from Guangxi, China. Group II contained germplasms from different regions and with different fruit color and fruit tumor traits, and there were certain traits and regional similarities in some germplasms with close genetic relationships. Group III was green except for the white peel of BBL-1-2. It was also found that 3-1-6,12 and eight bitter gourds from Guangdong were closely related.

The two re-sequenced materials No. 12 and H-13 were clustered in different groups; the genetic distance was 0.563, the genetic relationship was distant, and the genetic background was complex. The two materials showed strong heterosis for yield. Therefore, when the molecular markers of the parent

populations have been developed, heterosis prediction can be enhanced, which may not predict the best hybrid combinations, but it will certainly reveal the practical value of allocating existing and new hybrid bitter melon germplasm to heterosis populations, thereby increasing the opportunity to develop ideal hybrids from the best heterosis populations. Similar conclusions have been reached for rice and maize [28, 29]. Conversely, white bitter melon did not show a similar aggregation trend, likely because the origin of white bitter melon may involve diversification and its genetic basis may be relatively rich, consistent with results reported by Zhou et al. [30].

In this study, InDel markers were used to analyze the genetic diversity within a bitter melon germplasm sample and the genetic relationship among bitter melon genetic resources was studied at the molecular level. The evaluation of this genetic diversity showed that the InDel markers developed in our study will be a valuable marker source for the study of the genetic background and the molecular variation within bitter melon germplasm with other sources in the future.

5. Conclusions

In this study, A novel set of InDel primers with high polymorphism potential were developed based on resequencing. The application of these markers for the evaluation of genetic diversity among 53 bitter melon genotypes and the construction of molecular fingerprinting indicated that they were of great significance for mapping, variety protection and hybrid seed purity identification in the genetic analysis of bitter melon. Therefore, the newly developed InDel marker of bitter melon between the fine germplasm H-13, NO.12 was of great significance and provided a valuable bitter melon resource for the identification of genetic characteristics of bitter melon genotype.

Declarations

Author contribution

Zhengguo Liu and Lianlian Ma conceived and designed the study. Liwen Su conducted all experiments. Peng Wang, Xiaochun Huang, Wenting Wu, Haixuan Lv, Zhikui Cheng, and Jiquan Gou participated in some experiments. Liwen Su and Lianlian Ma wrote the manuscript. Z.G.L. and P.W revised the manuscript. Liwen Su and Lianlian Ma contributed equally. All the authors reviewed the manuscript.

Funding

This work was funded by Excavation of High Quality Germplasm Resources of Characteristic Vegetables and Breeding of New Varieties-Wax Gourd and Chieh-qua (AE3310012701).

Data Availability

The data sets supporting the results of this article are included within the article.

Declarations

Consent to Participate Not applicable.

Consent to Publish Not applicable.

Conflict of Interest The authors declare no competing interests.

References

1. Schaefer, H., & Renner, S. S. (2010). A three-genome phylogeny of Momordica (Cucurbitaceae) suggests seven returns from dioecy to monoecy and recent long-distance dispersal to Asia. *Molecular Phylogenetics & Evolution*, 54, 553–560
2. Basch, E., Gabardi, S., & Ulbricht, C. (2003). Bitter gourd(Momordica charantia): A review of efficacy and safety. *American Journal of Health-System Pharmacy*, 60, 356
3. Grover, J. K., & Yadav, S. P. (2004). Pharmacological actions potential uses of Momordica charantia: A review. *Journal of Ethnopharmacology*, 93, 123–132
4. Dey, S. S., Singh, A. K., Chandel, D., & Behera, T. K. (2006). Genetic diversity of bitter gourd (Momordica charantia L.) genotypes revealed by RAPD markers and agronomic traits. *Scientia Horticulturae*, 109, 28
5. Trakoon-Osot, W., Sotanaphun, U., Phanachet, P., Porasuphatana, S., & Komindr, S. (2013). Pilot study: Hypoglycemic and antiglycation activities of bitter gourd (Momordica charantia L.) in type 2 diabetic patients. *Journal of Pharmacy Research*, 6, 859–864
6. Liu, Z. J., Niu, Y., Zhu, J., Liu, Z. H., & Yang, Y. (2016). Study on genetic diversity of agronomic traits and genetic relationships among core collections of bitter gourd. *Agricultural Science & Technology*, 17, 1134–1138
7. Zhang, J. N., Chang, X. P., Hao, C. Y., Liu, G. R., & Jing, R. L. (2010). Genetic diversity of common wheat germplasm resources with drought resistance in northern China. *Journal of Plant Genetic Resources*, 11, 253–259
8. Huang, R., Sun, D., Zhang, M., Fang, F., & Luo, H. (2008). Genetic diversity of Momordica charantia and its cluster analysis based on morphological traits. *Guangxi Agricultural Sciences*, 39, 351–356
9. Kang, J. B., Zhu, H. S., Li, D. Z., Li, Y. P., Wang, W. C., & Wen, Q. F. (2010). Determination of Momordica's germplasm diversity by ISSR. *Fujian Journal of Agricultural Sciences*, 25, 597–601
10. Dhillon, N., Sanguansil, S., Schafleitner, R., Wang, Y. W., & McCreight, J. D. (2016). Diversity among a wide Asian collection of bitter gourd landraces and their genetic relationships with commercial hybrid cultivars. *Journal of the American Society for Horticultural Science*, 141, 475–484
11. Wang, G. L. (2020). Morphological classification and molecular identification of germplasm resources in Momordica charantia I. *Botanical Research*, 9, 341–353
12. Hayashi, K., Yoshida, H., & Ashikawa, I. (2006). Development of PCR-based allele-specific and InDel marker sets for nine rice blast resistance genes. *Theoretical and Applied Genetics*, 113, 251–260

13. Li, S. G., Shen, D., Liu, B., Qiu, Y., Zhang, X. H., Zhang, Z. H. ... Li, X. X. (2013). Development and application of cucumber InDel markers based on genome re-sequencing. *Journal of Plant Genetic Resources*, 14, 278–283
14. Li, W. P., Cheng, J. W., Wu, Z. M., Qin, C., Tan, S., Tang, X. ... Hu, K. L. (2015). An InDel-based linkage map of hot pepper (*Capsicum annuum*). *Molecular Breeding*, 35, 32
15. Cui, J. J., Cheng, J. W., Nong, D. G., Peng, J. Z., Hu, Y. F., He, W. M. ... Hu, K. L. (2017). Genome-Wide analysis of simple sequence repeats in bitter melon (*Momordica charantia*). *Frontiers in Plant Science*, 8, 1103
16. Cui, J. J., Peng, J. Z., Cheng, J. W., & Hu, K. L. (2020). Development and validation of genome-wide InDel markers with high levels of polymorphism in bitter melon (*Momordica charantia*). *BMC Genomics*, 22, 190
17. An, C. Y., Xie, D. S., Peng, Q. W., & He, X. M. (2011). Comparison of genomic DNA extraction methods in Chieh-qua (*Benincasa hispida*). *China Cucurbits and Vegetables*, 24, 1–4
18. Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25, 1754–1760
19. Fonteles, T. V., Costa, M., Jesus, A., Miranda, M., Fernandes, F., & Rodrigues, S. (2012). Power ultrasound processing of cantaloupe melon juice: Effects on quality parameters. *Food Research International*, 48, 41–48
20. Sun, M. (1997). Genetic diversity in three colonizing orchids with contrasting mating systems. *American Journal of Botany*, 84, 224
21. Moonsap, P., Laksanavilat, N., Sinumporn, S., Tasanasuwan, P., Kate-Ngam, S., & Jantasuriyarat, C. (2019). Genetic diversity of Indo-China rice varieties using ISSR, SRAP and InDel markers. *Journal of Genetics* 98.
22. Cambiaso, V., Pratta, G. R., Pereira Da Costa, J. H., Zorzoli, R., Francis, D. M., & Rodríguez, G. R. (2019). Whole genome re-sequencing analysis of two tomato genotypes for polymorphism insight in cloned genes and a genetic map construction. *Scientia Horticulturae*, 247, 58–66
23. Wang, Y. X., Wu, K., Fang, Y. X., Chen, H. Q., Ni, S., Fu, X. D., & Zhu, X. D. (2014). Genetic analysis, gene mapping and breeding application of small grain and dense panicle mutant sep1 in rice. *Chinese Journal of Rice Science*, 28, 229–234
24. Lee, M., Godshalk, E. B., Lamkey, K. R., & Woodman, W. W. (1989). Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. *Crop Science*, 29, 1067–1071
25. Sekhon, M. S., & Gupta, V. P. (1995). Genetic distance and heterosis in Indian mustard: Developmental isozymes as indicators of genetic relationships. *Theoretical and Applied Genetics*, 91, 1148–1152
26. Gao, S., Lin, B. Y., Xu, D. X., Fu, R. Q., Lin, F., Lin, Y. Z., & Pan, D. M. (2010). Genetic Diversity of Bitter Gourd (*Momordica charantia* L.) based on RAPD and ISSR. *Journal of Plant Genetic Resources*, 11, 78–83

27. Zhao, X. J., Song, J. W., & Hu, K. L. (2013). Genetic diversity analysis of bitter gourd germplasm resources based on SRAP markers. *Journal of Plant Genetic Resources*, 14, 79–85
28. Lanza, L., Souza, C., Ottoboni, L., Vieira, M., & Souza, A. (1997). Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. *Theoretical and Applied Genetics*, 94, 1023–1030
29. Xie, F. M., He, Z. Z., Esguerra, M. Q., Qiu, F. L., & Ramanathan, V. (2013). Determination of heterotic groups for tropical Indica hybrid rice germplasm. *Theoretical and applied genetics*, 127, 407–417
30. Zhou, K., Zhang, C., Luo, J., & He, X. (2013). SRAP analysis on genetic diversity of bitter gourd (*Momordica charantia* L.). *Guangdong Agricultural Sciences*, 40, 136–140

Figures

Figure 1

A, Sample base error rate and base type distribution; B, Sample depth and chromosome coverage depth distribution.

Figure 2

A, InDel length map of whole genome and coding region; B, InDel annotation statistics. InDel length (less than 10 bp) in the ordinate, greater than 0 for insertion, less than 0 for deletion, transverse coordinates for the corresponding number.

Figure 3

Polymorphism map of primers MC02-1, MC05-2, and MC11-4 among different bitter melon germplasm accessions. 1–53 16-1,10-2-5-3, 10-2-5-4, LM-3-1, I-7, PE-1, 2-3-1, BO5367, 9-1-4-3, 3-1-6, I-4-1-2-7, KS-1, BO-1, LH-3, I-4-1-2-6, 10-2-1, DJ-7-1, I-9-6-1, ITY-1, I-6-3, H-13, No.12, YI-1-2-1, BWB-1-1, No. 8, MX-1, BP6237, BH-2, LH-1, HYT, PD-1, YHF, DJ-9-3, No. 5, HX-1, LH-4, Yuxiu, Huayu 3, No. 58, HY-1, XHY, Huabi 2, BP6201, I-13-1-1, YXH, LM-4-1, LM-1-1, LI-1-BP6221, HXU, BBL-1-2, BP6214, HB-1.

Figure 4

Determination of the degree of purity of three bitter melon hybrids. A, Primer MC2-1 was used to identify the degree of purity of hybrid combinations with No. 12 as female parent and H-13 as male parent. M, DL1 500 DNA marker; 1–2: H-13 paternal parent; 3–4: 12 female parent; 5–54: Hybrids; B, Primer MC5-1 was used to determine the degree of purity of ‘Tiantianhao 11’ bitter melon; M, DL1 500 DNA marker; 1–2: 10-2-1-1-1-2 paternal parent; 3–4: 10-2-5-3 female parent; 5–58: Hybrid C: primer MC11-4 for ‘Tiantianhao 12’ bitter melon purity identification; M: DL1 500 DNA marker; 1–2: 10-2-5-3 male parent; 3–4:3-1-6-2-1 female parent; 5–54: hybrid.

Figure 5

UPGMA clustering of bitter melon germplasm. 1–53: 16-1, 10-2-5-3, 10-2-5-4, LM-3-1, I-7, PE-1, 2-3-1, BO5367, 9-1-4-3, 3-1-6, I-4-1-2-7, KS-1, BO-1, LH-3, I-4-1-2-6, 10-2-1, DJ-7-1, I-9-6-1, ITY-1, I-6-3, H-13, No. 12, YI-1-2-1, BWB-1-1, No. 8, MX-1, BP6237, BH-2, LH-1, HYT, PD-1, YHF, DJ-9-3, No. 5, HX-1, LH-4, Yuxiu, Huayu 3, No. 58, HY-1, XHY, Huabi 2, BP6201, I-13-1-1, YXH, LM-4-1, LM-1-1, LI-1-BP6221, HXU, BBL-1-2, BP6214, HB-1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterials.doc](#)