

Development of KASP Markers for Germplasm Characterization and Fingerprinting Identification of Broccoli in China

Yusen Shen

Institute of Vegetables, Zhejiang Academy of Agricultural Sciences

Jiansheng Wang

ZAAS

Huifang Yu

ZAAS

Xiaoguang Sheng

ZAAS

Zhenqing Zhao

ZAAS

Hong-Hui Gu (✉ guh2199@163.com)

Institute of Vegetables, Zhejiang Academy of Agricultural Sciences <https://orcid.org/0000-0001-5816-7530>

Research article

Keywords: broccoli, KASP, germplasm characterization, fingerprinting, next generation-sequence

Posted Date: September 4th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-60988/v1>

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

[Read Full License](#)

Abstract

Background: Broccoli (*Brassica oleracea* var. *italica*) is a vegetable widely cultivated in China. Many new-type broccoli cultivars were bred and developed by Chinese breeders during the recent three decades. However, the broccoli cultivar nomenclature and detailed information of genetic relationships among broccoli germplasms are unclear.

Results: The present study identified millions of SNPs by next-generation sequencing of 23 representative broccoli lines. Through several steps of selection, 100 SNPs were successfully converted into KASP markers, and used to evaluate the genetic diversity, genetic relationship, and population structure of 392 broccoli accessions, which represent the mainly broccoli breeding materials in China. The initial, introduced and improved accessions were well clustered, though some accessions were overlapped between groups, probably reflecting the fact that breeding activities led to genetic similarities. To make the KASP genotyping more efficient and cost-effective, 25 of the 100 KASPs were selected for fingerprinting of all accessions, and the 2D barcode contained fingerprinting information were generated for elite varieties.

Conclusion: The KASP markers developed in this study provided an efficient way for germplasm characterization, DNA fingerprinting, seed purity identification, and marker-assisted selection of broccoli in China.

Background

Broccoli (*Brassica oleracea* var. *italica*) is an economically important vegetable in many countries, and is becoming more popular as a human diet due to its rich source of nutritional value as well as anti-cancer glucosinolates in florets [1]. It was originated in the Mediterranean basin and introduced to China in the 20th century [2]. In recent years, the broccoli industry in China has made tremendous progress with about 80 000 ha cultivated area in 2019. In addition, many new-type broccoli cultivars were produced and developed by Chinese breeders during the recent three decades. While significant increases in productivity resulting from the use of improved varieties with narrow genetic base. Consequently, it is difficult to distinguish the similar broccoli varieties in seed trading market. The International Union for the Protection of New Varieties of Plants (UPOV) established distinctness, uniformity and stability (DUS) testing for new varieties before registration [3]. As supplementary, DNA fingerprinting can help improve knowledge of varieties with minimum distance in traditional characteristics [4].

Over the past three decades, several different DNA marker technologies, including amplified fragment length polymorphisms (AFLPs), simple sequence repeats (SSRs), and single nucleotide polymorphisms (SNPs), have been widely applied in research areas such as DNA fingerprinting of varieties, genetic diversity analysis, population structure analysis, and molecular marker-assisted breeding [5]. SSRs are routinely used for fingerprinting, because of their high level of polymorphism [6–9]. However, SSRs have some disadvantages, for instance, the throughput of detection is low, and the data integration or

comparison between different detection platforms is difficult [5]. While SNPs are the most abundant polymorphisms in any species [10, 11], and have been used in many genetic applications, including germplasm characterization (genetic diversity, genetic relationship, and population structure), quantitative trait loci (QTL) mapping, and marker-assisted selection [12]. Some SNP array-based marker sets could also be used for fingerprinting [5, 13–15], though this may not be their primary purpose.

Kompetitive allele specific PCR (KASP) assays are based on competitive allele-specific PCR and enable bi-allelic scoring of SNPs, insertions and deletions (InDels) at specific loci [16]. Compared with fixed chip array, KASP is a more flexible and cost-effective technology for SNP genotyping, especially for a small number of markers to genotype large number of samples. KASP assays have been used in various crops, such as rice [17, 18], wheat [19], and maize [12]. In some horticultural plants, KASP markers were used primarily for fine mapping of the candidate QTL [20–22]. To our knowledge, the development of KASP markers for broccoli has not been reported yet.

In this study, millions of SNPs were identified by re-sequencing of 23 broccoli lines. Through several steps of selection, 100 SNPs were successfully converted into KASP markers, and used for genetic diversity, genetic relationship, and population structure analysis of 392 broccoli accessions in China. Then 25 of the 100 KASPs were used for fingerprinting of all accessions. This KASP marker set provided an efficient way for broccoli germplasm characterization and marker-assisted selection.

Results

KASP marker development

Twenty-three broccoli lines were used for the next generation sequencing, with average sequence depth of 28.0×. A total of 346.188G raw data were initially obtained, and after quality filtering and adapter trimming, 345.577G clean data were available for further processing (Additional file 1: Table S1). A total of 2303.9 million paired reads (each line ranges from 87.7 to 124.9 million) were generated, and 2272.9 million (98.7%) paired reads were successfully aligned to the broccoli reference genome [23] (Additional file 1: Table S2). As a result, millions of SNP markers were detected in each line, with the number from 899,926 to 1,908,908 (Additional file 1: Table S3).

A total of 28,220 SNPs were selected after filtering out the SNPs with high missing data (> 20%), low levels of MAF (< 0.05) and the SNPs with multiple variation before or after 50bp of the loci. The PIC values of all these SNPs were calculated, and 13,621 SNPs with the PIC between 0.2 and 0.5 were retained. Of these, 8,768 (64.4%) SNPs could be successfully designed as KASP markers. According to the physical position of the SNPs and the genomic structural annotation, 2,515 SNPs located the exonic (including non-synonymous, stop gain or stop loss), intergenic, upstream or downstream of the functional genes. These SNPs may effected some important agronomic traits that could be used for molecular breeding. Ultimately, 500 SNPs that uniformly distributed across nine chromosomes were selected to develop KASP markers.

Genotyping of 392 broccoli accessions

To evaluate the quality of the KASP markers, 23 broccoli lines used for next generation-sequencing (NGS) were genotyped by the selected 500 KASP markers. As a result, 347 (69.4%) KASP markers could be genotyped successfully, that is, the genotype for most of the accessions was clear (Fig. 1A). While 54 KASPs presented inconformity genotype between NGS and KASP assay, including some KASPs with no polymorphism (Fig. 1B). The remaining 293 KASPs were further used for genotyping the 392 broccoli accessions. To make the results reliable, markers showed more than 10% missing values or had ambiguous SNP calling were removed (Fig. 1C), and 100 markers with high quality and evenly distribution were used for further analysis (Fig. 2; Additional file 2)

In all 392 accessions, the heterozygous marker ratio was 5.5%- 87.2%, and 387 accessions (98.7%) were labeled with more than 30% heterozygous marker ratio, indicating that most of these accessions were heterozygous (Fig. 3A). The MAF values of all KASPs were 0.11- 0.48, with an average of 0.32. Most of the KASPs (88%) have the MAF values more than 0.2. The PIC values of all KASPs were 0.20- 0.50, and the percentage of PIC values between 0.45 and 0.50 was relatively high (47 %) (Fig. 3B; Additional file 2)

Accession assessment, genetic diversity and population evaluation

Based on the genotyping results of 100 KASP markers, the phylogenetic tree was constructed by Neighbor-joining method using MEGA software version 10.0.4, with 1 000 bootstrap replications. All 392 broccoli accessions were clustered into three groups (Fig. 4). The cluster I contained 103 accessions, and most of them were improved accessions. The cluster II contained eight accessions, which were introduced from Japan, including bck2, bck3, bck5, bck6 and so on. The cluster III is the major group that contains 281 accessions, and most of them were initial accessions. Most of the improved or introduced accessions presented strong growth potential, high-round flower head with thin and uniform size buds. While the initial accessions were faded in some characteristics.

Principal component analysis (PCA) indicated that cluster II (introduced accessions) was contained within cluster III (initial accessions), and cluster I (improved accessions) was overlapped with the other groups (Fig. 5). The first axis explained 8.1% and the second axis explained 6.5% of the overall variance, respectively.

The population structure of 392 broccoli accessions was classified by Structure software version 2.3.4. The population number K was set to 1-10, and each hypothetical K value was calculated five times. The ΔK of Evanno was maximal at $K = 2$ (Fig. 6A). Therefore, the 392 broccoli accessions were divided into 2 groups: POP1 and POP2 (Fig. 6B). POP1 was an improved-type broccoli subgroup consisting of 155 broccoli accessions, and POP2 was an initial-type broccoli subgroup consisting of 237 broccoli accessions. The F_{ST} value of POP1 and POP2 was 0.0008, which could strongly explain the genetic distance between these two populations. Comparing with the results of phylogenetic tree, 98 of 103 accessions in cluster I belongs to POP1. All accessions in cluster II, and 223 of 281 accessions in cluster III belong to POP2 (Additional file 4).

Above all, these 100 KASP markers were effective at discriminating the population structure of the accessions. However, different types of accession were overlapped, probably reflecting the fact that breeding activities led to genetic similarities.

Selection core KASP markers for fingerprinting of broccoli accessions

To build a rapid and cost-effective way of variety identification, we selected 25 KASP markers from the genotyping database as the fingerprinting for every accession. These KASP markers were highly effective for distinguishing among 392 examined accessions, and evenly distributed across the nine chromosomes (Fig. 2). For each accession, the genotype based KASP barcode was used to generate corresponding 2D barcode using online tool (available at www.cli.im). Then this barcode can be scanned to obtain the information used for creating the 2D barcode. Figure 7 depicts barcode of a representative variety of broccoli used in the present study.

Discussion

KASP marker development by next-generation sequencing

Markers have been extensively used for genetic diversity assessment, population structure determination, QTL mapping or molecular breeding. As the third generation marker type, SNPs are the most commonly used markers because of its high density in genomes, and more available high-throughput genotyping method. In the present study, we identified millions of SNPs by next-generation sequencing of 23 representative broccoli lines. Mining SNPs by sequencing so many lines can make SNPs a wide range of variation.

Hundreds of SNPs can meet the demand of germplasm characterization and fingerprinting. To make it cost-effective and more flexible, KASP assay technology, which produced by LGC company, was used for SNP genotyping. While not all SNPs can be converted into KASP markers. Design success rate and work success rate are two parameters for conversion efficiency evaluation. The SNP sites that can be used to design primers to the total number of SNP sites are called the design success rate. And the number of SNP sites that can generate genotype calls via primers to the number of SNP sites with successful designed primers called the work success rate [17]. In this study, we drove up the filtering standards, which the GC content of the primers should be more than 0.3, so that the design success rate was only 64.4%. But this rate could not affect the final result or experimental cost, because we designed the KASP primers before synthesis. In addition, 347 (69.4%) of the 500 KASP markers could successfully genotyped for the broccoli accessions. Compared with some other reports, this lower work success rate might be because of the inconsistent DNA quality and quantity provided by different breeding groups. To solve this problem, we selected 100 high quality KASP markers for germplasm characterization. To our knowledge, this was the first time that hundreds of KASP markers were developed in broccoli. Thanks to the supporting of these breeding groups, so many accessions could represent nearly all types and important agronomic traits of broccoli in China.

Characterization of 392 representative broccoli accessions in China

Broccoli is a popular vegetable in China, not only because of its good taste, but also the rich nutrients in the flower head and stem. While broccoli breeding in China has a history of only 30 years, and a few elite varieties were used as core germplasms, resulting the narrow genetic base of accessions. In this study, based on the 100 selected KASP markers, 392 accessions were clustered into two (Fig. 6) or three sub-groups (Fig. 4). In particular, PCA results showed that group II was encompassed by group III, and group I was overlapped with other groups (Fig. 5), indicating that the initial and introduced type did not have obvious boundaries.

Increasing favourable alleles in a breeding population, while reducing inbreeding and avoiding genetic drift, is of great importance for attaining long-term breeding gains [24]. The genotyping of Chinese broccoli accessions can help improve knowledge of the genetic relationships among accessions, which contribute to reduce chances of introgression of undesirable alleles or guide selection of parents to reduce mating among related lines [25].

Application value of fingerprinting for broccoli accessions

As the narrow genetic base of broccoli, it is sometimes difficult to distinguish varieties according to their agronomic performance. Marker-based fingerprinting have advantages in identifying varieties, and are free from environmental impact, compared with customary field inspection [26]. Similar to other horticultural species, spurious varieties are easily mixed with the genuine ones in broccoli seed market, which will eventually harm the interest of consumers and breeders. In the present study, the barcodes of all broccoli accessions were established, especially for some widely used varieties. This barcode contained fingerprinting information of the accessions, and can ensure the authenticity of varieties.

In addition, commercial F_1 cultivars of broccoli were cytoplasmic male sterile lines, and the parents or other heterogeneous seeds will be easily mixed during seed production, including the cause of pollen contamination, mechanical admixture, or artificial mislabeling. Loss of parent materials can cause irreparable damage to the breeders. To solve these problems, it is essential to identify the seed purity before commercial distribution. Though we have developed 100 KASPs for broccoli fingerprinting, in our experience, 3-5 KASPs were adequate for seed purity identification. The marker assembly can be selected from the 100 KASPs according to the parental genotypes of a specific variety.

The exchange of germplasm resources is the prerequisites for human beings to effectively develop and utilize their resource value, avoid duplication and waste, and create more wealth for human beings. While protecting the intellectual property of breeders needs to be taken into consideration before germplasm exchange. From this perspective, fingerprinting can not only provide the identification for every germplasm, but also the evidence for the proportion of shared benefits. That is, breeders or merchants can acquire corresponding benefits from a hybrid variety, depending on the ratio of the same fingerprinting markers as its parents.

As the high-efficiency and cost-effective advantages, KASP assay is a promising technology for molecular breeding. In this study, most of the SNP loci selected for KASPs were based on the physical positions of functional genes. For example, the KASP markers GZKASP_1 and HFKASP_6 were developed from the reported loci controlling clubroot and black rot resistance in cabbage, respectively [27, 28]. Some KASPs were located the homologous genes of *Arabidopsis thaliana*, controlling glucosinolate, flowering, anthocyanin content and so on. Further study should be performed to evaluate the function of these KASPs in broccoli, so that they could be applied in marker-assisted selection.

Materials And Methods

Materials for next-generation sequencing and fingerprinting

Based on their phylogenetic relationships, 23 broccoli lines were selected for NGS. Ten of these materials were DH lines, and the others were high-generation inbred lines. These materials are representative in many important agronomic traits of broccoli, such as head color, head shape, leaf angle, stem height, flowering time and so on.

A total of 392 broccoli accessions provided by 11 breeding groups in China, were used for fingerprinting identification. These materials could nearly cover all ecological types of broccoli in China.

Selection of SNP loci

After obtaining the SNP loci from next generation sequencing of the 23 broccoli lines, statistical values, including minor allele frequency (MAF), polymorphic information content (PIC) and heterozygous marker ratio for each SNP were estimated using PowerMarker v3.25 [29].

Several steps were performed to select the available SNP loci for KASP marker development. (1) Filtered out the SNPs which in more than 20% of the materials were missing data. (2) SNPs with the MAF values under 0.05 were deleted. (3) Filtered out the SNPs that other variation existed before or after 50bp of the marker position. (4) SNPs with the PIC values under 0.2 were deleted. (5) Design KASP primers based on SNP loci, and the GC content lower than 0.3 were deleted. (6) To get the functional KASP markers, loci within the exonic, or at the intergenic, upstream or downstream of the functional genes were selected. (7) The remaining loci were selected according to the physical position, and cover the chromosomes as evenly as possible.

KASP marker genotyping

Genomic DNA for genotyping was extracted from young leaf tissues by a modified cetyltrimethylammonium bromide method [30]. The DNA quality and quantity were determined by a NanoDrop 2000 Spectrophotometer, and 10-30 ng/ml was suitable for KASP assay.

The PCR amplification was performed in 96 well microplates using a 10.14 μ l reaction volume. Each well contained 5 μ l template DNA, 5 μ l 2 \times KASP Master mix, and 0.14 μ l KASP assay mix. After dispensing

genotyping mix onto the reaction plate, the PCR was run following a program including three steps. Firstly, activation at 94 °C for 15 min. Secondly, denaturation at 95°C for 20 s, and elongation at 61-55 °C (drop 0.6 °C per cycle) for 60 s, with 10 cycles. Thirdly, denaturation at 95°C for 20 s, and elongation at 55 °C for 60 s, with further 26 cycles.

The PCR product was detected by a fluorescent microporous plate detector to determine the genotype of each SNP, and the SNP result was checked manually using amplification curves and final allele discrimination.

Abbreviations

UPOV: The International Union for the Protection of New Varieties of Plants; DUS: distinctness, uniformity and stability; AFLPs: amplified fragment length polymorphisms; SSRs: simple sequence repeats; SNPs: single nucleotide polymorphisms; QTL: quantitative trait loci; KASP: Kompetitive allele specific PCR; InDels: insertions and deletions; NGS: next generation-sequence; PCA: Principal component analysis; MAF: minor allele frequency; PIC: Polymorphic Information Content

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

The Major Scientific Joint Research of Improved Varieties (Broccoli) of China (2018-2022), National Key Research and Development Program of China (2017YFD0101805, 2017YFE0114500), Natural Science Foundation of Zhejiang Province (LQ20C150001), Science and Technology Department of Zhejiang Province (2016C02051-5), and Zhejiang Academy of Agricultural Sciences for subject construction (A1).

Authors' contributions

HG designed the experiments, supervised the study and revised the manuscript. YS analyzed the experimental data and wrote the manuscript. JW and ZZ collected the accessions and selected the core germplasms for re-sequencing. HY and XS designed the KASP markers and performed the research. All authors discussed the results and contributed to the final manuscript

Acknowledgments

We thank the breeding groups in China for kindly providing the DNA of broccoli accessions, including Chinese Academy of Agricultural Sciences, Tianjin Kerun Vegetable Research Institute, Beijing Academy of Agriculture and Forestry Sciences, Shanghai Academy of Agricultural Sciences, Jiangsu Academy of Agricultural Sciences, Taizhou Academy of Agricultural Sciences, Wenzhou Academy of Agricultural Sciences, Zhejiang Mitsuo Seed Co.,Ltd, Ningbo Weimeng Seed Co.,Ltd, and Zhenjiang Xinyuanda Horticulture Technology Co.,Ltd.

References

1. Wang J, Gu H, Yu H, Zhao Z, Sheng X, Zhang X. Genotypic variation of glucosinolates in broccoli (*Brassica oleracea* var. *italica*) florets from China. *Food Chem.* 2012;133:735–41.
2. Branca F, Bahcevandziev K, Perticone V, Monteiro A. Sources of resistance to downy mildew (*Peronospora parasitica* (Pers. ex Fr.) Fr.) in Sicilian germplasm of cauliflower and broccoli. *Biodivers Conserv.* 2005;14:841–8.
3. Jones H, Norris C, Smith D, Cockram J, Lee D, O’Sullivan DM, et al. Evaluation of the use of high-density SNP genotyping to implement UPOV Model 2 for DUS testing in barley. *Theor Appl Genet.* 2013;126:901–11.
4. Jones H, Mackay I. Implications of using genomic prediction within a high-density SNP dataset to predict DUS traits in barley. *Theor Appl Genet.* 2015;128:2461–70.
5. Tian HL, Wang FG, Zhao JR, Yi HM, Wang L, Wang R, et al. Development of maizeSNP3072, a high-throughput compatible SNP array, for DNA fingerprinting identification of Chinese maize varieties. *Mol Breed.* 2015;35:136.
6. Pandino G, Lombardo S, Moglia A, Portis E, Lanteri S, Mauromicale G. Leaf polyphenol profile and SSR-based fingerprinting of new segregant *Cynara cardunculus* genotypes. *Front Plant Sci.* 2015;5:800.
7. Ouyang P, Kang D, Mo X, Tian E, Hu Y, Huang R. Development and characterization of high-throughput EST-based SSR markers for *Pogostemon cablin* using transcriptome sequencing. *Molecules.* 2018;23:2014.
8. Zheng X, Cheng T, Yang L, Xu J, Tang J, Xie K, et al. Genetic diversity and DNA fingerprints of three important aquatic vegetables by EST-SSR markers. *Sci Rep.* 2019;9:14074.
9. Rauscher G, Simko I. Development of genomic SSR markers for fingerprinting lettuce (*Lactuca sativa* L.) cultivars and mapping genes. *BMC Plant Biol.* 2013;13:11.
10. Deschamps S, Campbell MA. Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. *Mol Breed.* 2010;25:553–70.
11. Trebbi D, Maccaferri M, de Heer P, Sørensen A, Giuliani S, Salvi S, et al. High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theor Appl Genet.* 2011;123:555–69.

12. Ertiro BT, Ogugo V, Worku M, Das B, Olsen M, Labuschagne M, et al. Comparison of Kompetitive Allele Specific PCR (KASP) and genotyping by sequencing (GBS) for quality control analysis in maize. *BMC Genomics*. 2015;16:908.
13. Thomson MJ, Singh N, Dwiyantri MS, Wang DR, Wright MH, Perez FA, et al. Large-scale deployment of a rice 6 K SNP array for genetics and breeding applications. *Rice*. 2017;10:40.
14. Xu C, Ren Y, Jian Y, Guo Z, Zhang Y, Xie C, et al. Development of a maize 55 K SNP array with improved genome coverage for molecular breeding. *Mol Breed*. 2017;37:20.
15. Ellis D, Chavez O, Coombs J, Soto J, Gomez R, Douches D, et al. Genetic identity in genebanks: Application of the SolCAP 12K SNP array in fingerprinting and diversity analysis in the global in trust potato collection. *Genome*. 2018;61:523–37.
16. Hao X, Yang T, Liu R, Hu J, Yao Y, Burlyaeva M, et al. An RNA sequencing transcriptome analysis of grasspea (*Lathyrus sativus* L.) and development of SSR and KASP markers. *Front Plant Sci*. 2017;8:1873.
17. Yang G, Chen S, Chen L, Sun K, Huang C, Zhou D, et al. Development of a core SNP arrays based on the KASP method for molecular breeding of rice. *Rice*. 2019;12:21.
18. Steele KA, Quinton-Tulloch MJ, Amgai RB, Dhakal R, Khatiwada SP, Vyas D, et al. Accelerating public sector rice breeding with high-density KASP markers derived from whole genome sequencing of indica rice. *Mol Breed*. 2018;38:38.
19. Rasheed A, Wen W, Gao F, Zhai S, Jin H, Liu J, et al. Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor Appl Genet*. 2016;129:1843–60.
20. Rett-Cadman S, Colle M, Mansfeld B, Barry CS, Wang Y, Weng Y, et al. QTL and transcriptomic analyses implicate cuticle transcription factor SHINE as a source of natural variation for epidermal traits in cucumber fruit. *Front Plant Sci*. 2019;10:1536.
21. Liao N, Hu Z, Li Y, Hao J, Chen S, Xue Q, et al. *Ethylene-responsive factor 4* is associated with the desirable rind hardness trait conferring cracking resistance in fresh fruits of watermelon. *Plant Biotechnol J*. 2020;18:1066–77.
22. Zhu H, Zhai W, Li X, Zhu Y. Two QTLs controlling clubroot resistance identified from bulked segregant sequencing in pakchoi (*Brassica campestris* ssp. *chinensis* Makino). *Sci Rep*. 2019;9:9228.
23. Belser C, Istace B, Denis E, Dubarry M, Baurens F-C, Falentin C, et al. Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat Plants*. 2018;4:879–87.
24. Padi FK, Ofori A, Takrama J, Djan E, Opoku SY, Dadzie AM, et al. The impact of SNP fingerprinting and parentage analysis on the effectiveness of variety recommendations in cacao. *Tree Genet Genomes*. 2015;11:44.
25. Collard BCY, Mackill DJ. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos Trans R Soc B Biol Sci*. 2008;363:557–72.
26. Tommasini L, Batley J, Arnold G, Cooke R, Donini P, Lee D, et al. The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. *Theor Appl Genet*. 2003;106:1091–101.

27. Lee J, Izzah NK, Choi B-S, Joh HJ, Lee S-C, Perumal S, et al. Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Res.* 2016;23:29–41.
28. Kifuji Y, Hanzawa H, Terasawa Y, Ashutosh, Nishio T. QTL analysis of black rot resistance in cabbage using newly developed EST-SNP markers. *Euphytica.* 2013;190:289–95.
29. Liu K, Muse S V. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics.* 2005;21:2128–9.
30. Hanania U, Velcheva M, Sahar N, Perl A. An improved method for isolating high-quality DNA from *Vitis vinifera* nuclei. *Plant Mol Biol Report.* 2004;22:173–7.

Figures

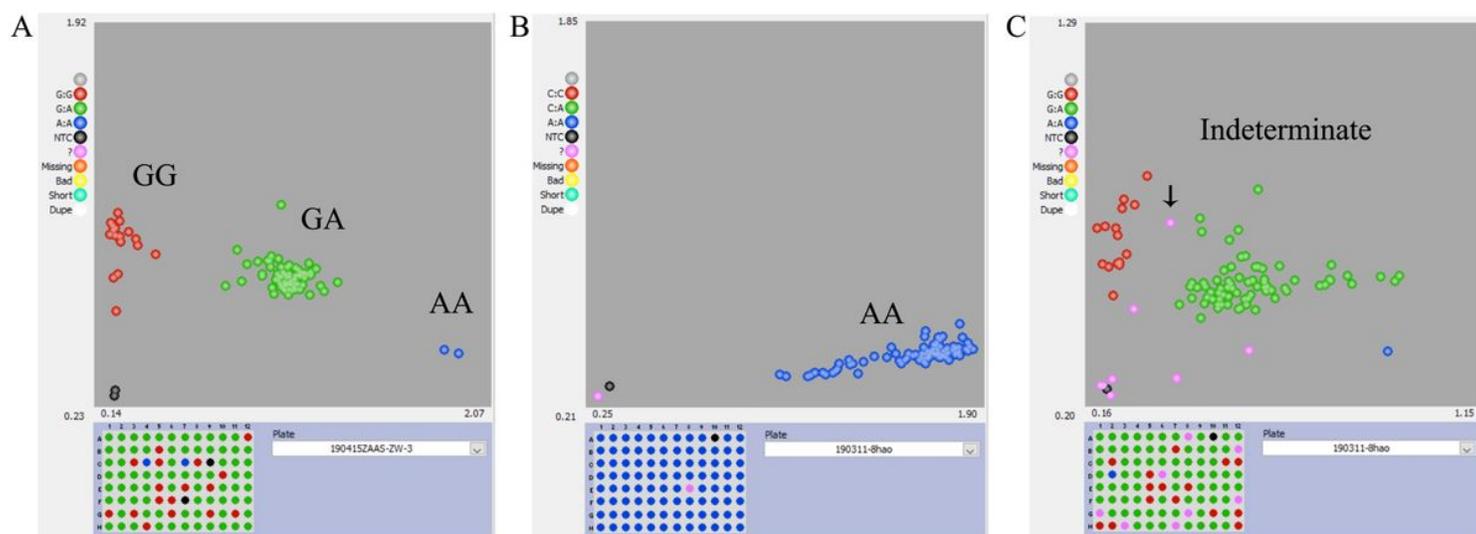


Figure 1

Representative results of KASP genotyping assay. (A) KASP marker with good polymorphism that was successfully developed. (B) KASP marker with no polymorphism that should be discarded. (C) KASP marker with more than 10% missing values or had ambiguous SNP calling that should be discarded.

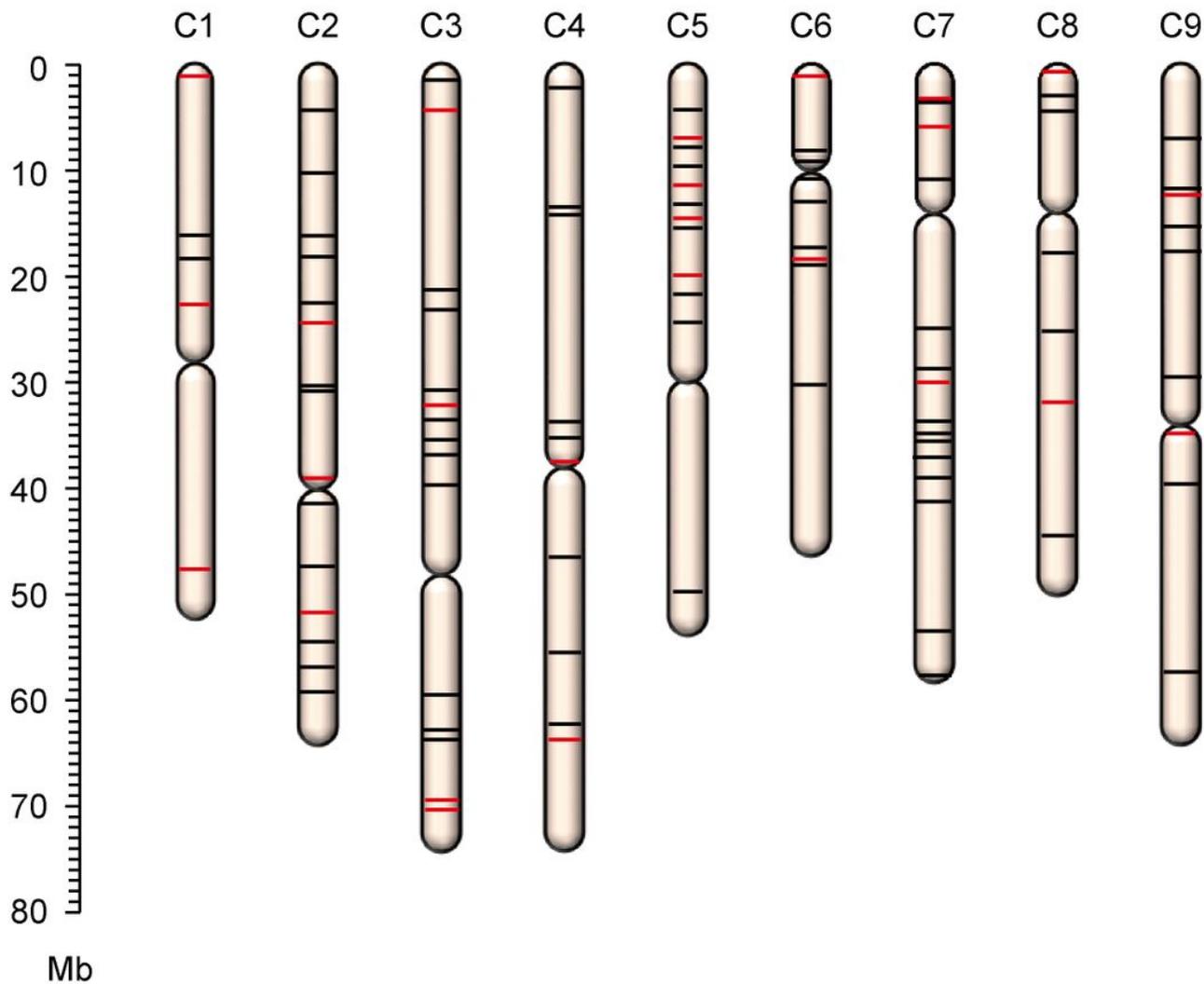


Figure 2

Distribution of KASP markers among broccoli chromosomes. The horizontal lines perpendicular to a chromosome represent the KASP markers developed in this study, among which the red ones were core KASP markers selected for fingerprinting of broccoli accessions.

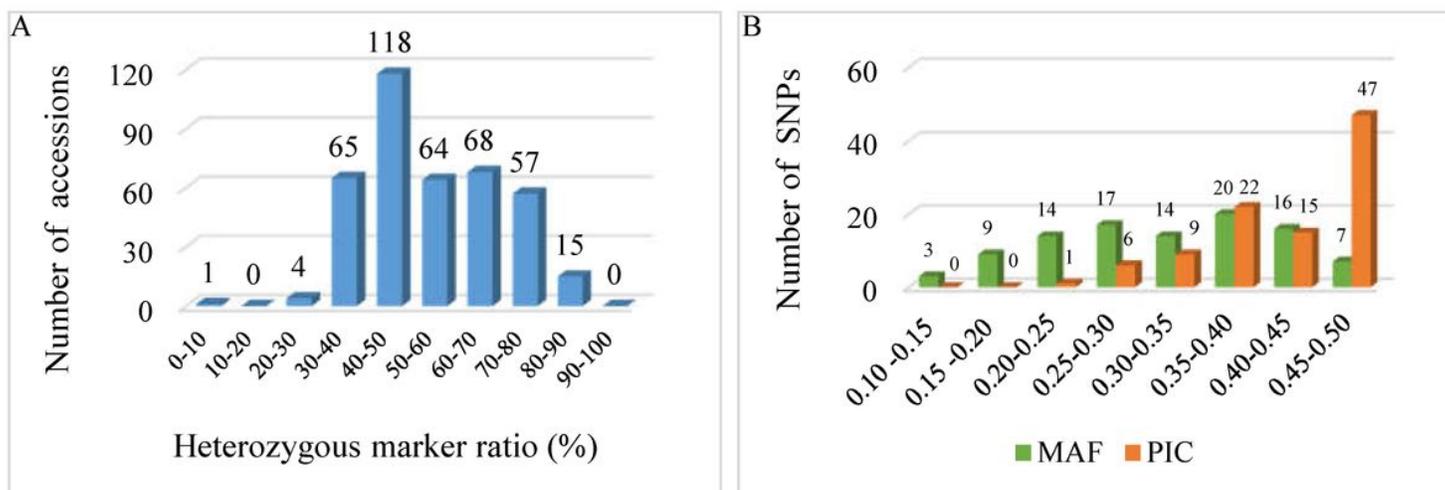


Figure 3

Genetic diversity of 100 KASP markers. (A) Heterozygosity (B) Minor allele frequency and polymorphism information content.

Tree scale: 0.1

Colored ranges

- CAAS
- KR
- BAAFS
- SAAS
- ZAAS
- JAAS
- TAAS
- WAAS
- MZA
- WM
- XYD
- Japan

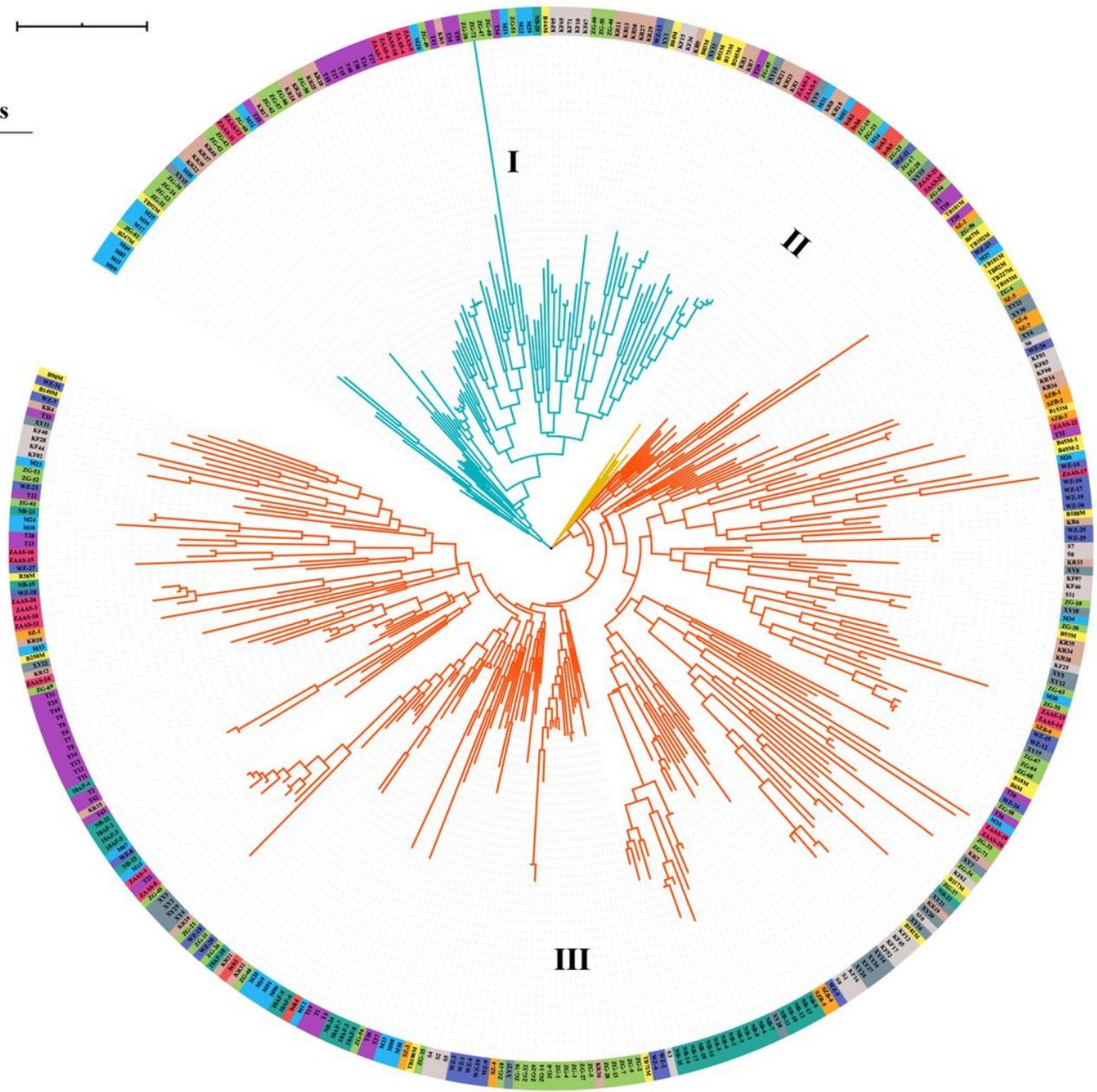


Figure 4

phylogenetic tree constructed using the Neighbor-Joining method based on 392 broccoli accessions. The colors in outer circle represent different sources of the accessions. CAAS: Chinese Academy of Agricultural Sciences; KR: Tianjin Kerun Vegetable Research Institute; BAAFS: Beijing Academy of Agriculture and Forestry Sciences; SAAS: Shanghai Academy of Agricultural Sciences; ZAAS: Zhejiang Academy of Agricultural Sciences; JAAS: Jiangsu Academy of Agricultural Sciences; TAAS: Taizhou Academy of Agricultural Sciences; WAAS: Wenzhou Academy of Agricultural Sciences; MZA: Zhejiang Mitsuo Seed Co.,Ltd; WM: Ningbo Weimeng Seed Co.,Ltd; XYD: Zhenjiang Xinyuanda Horticulture

Technology Co.,Ltd; Japan: Sakata Seed Corporation, Japan. The colors in the tree correspond to subpopulations. Cluster I with blue color was improved accessions, cluster II with yellow color was introduced accessions, cluster III with red color was initial accessions.

Individuals – PCA

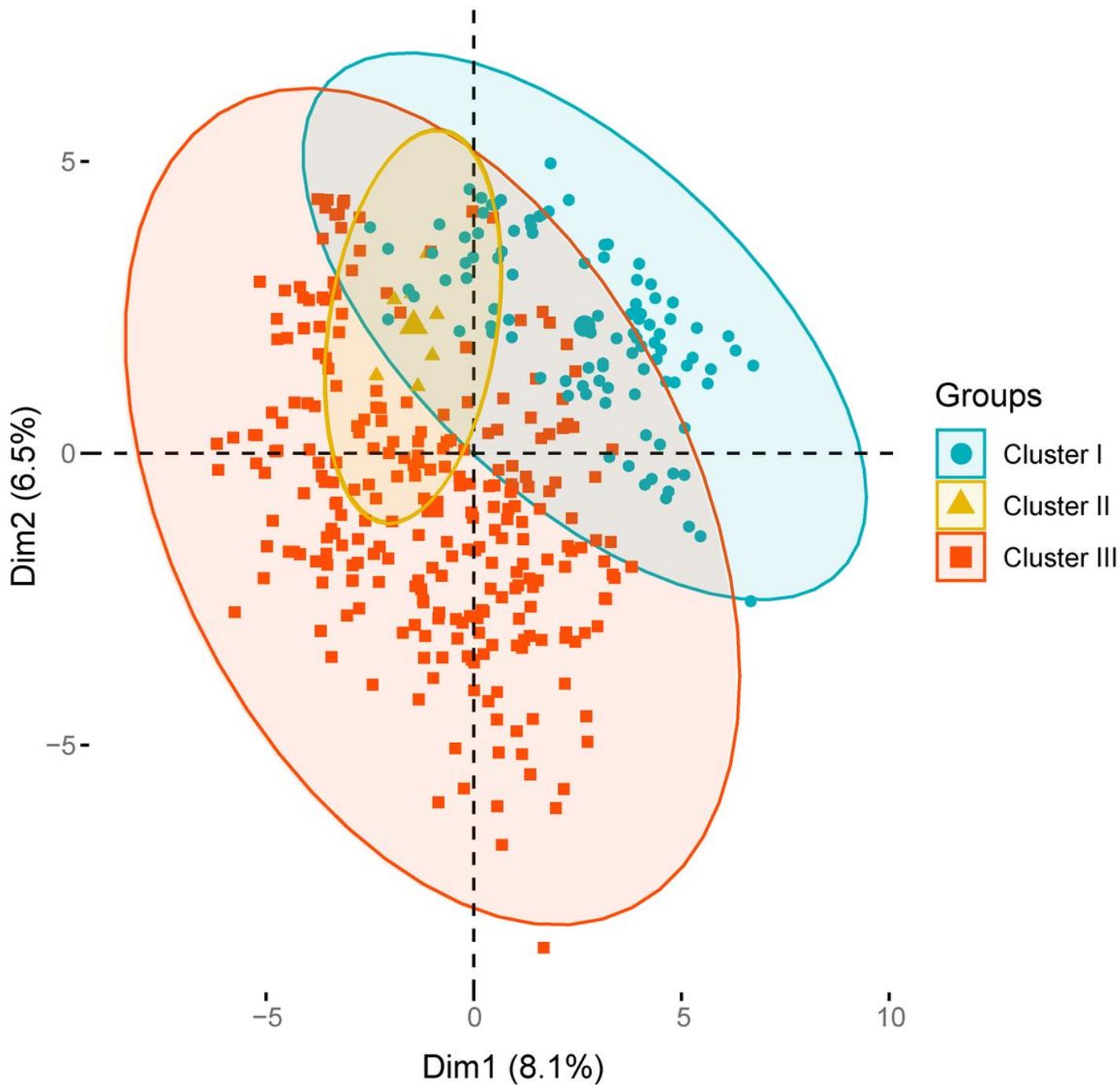


Figure 5

Principal component analysis (PCA) of the 392 broccoli accessions genotyped with 100 genomic KASPs. Cluster I with blue color was improved accessions, cluster II with yellow color was introduced accessions, cluster III with red color was initial accessions.

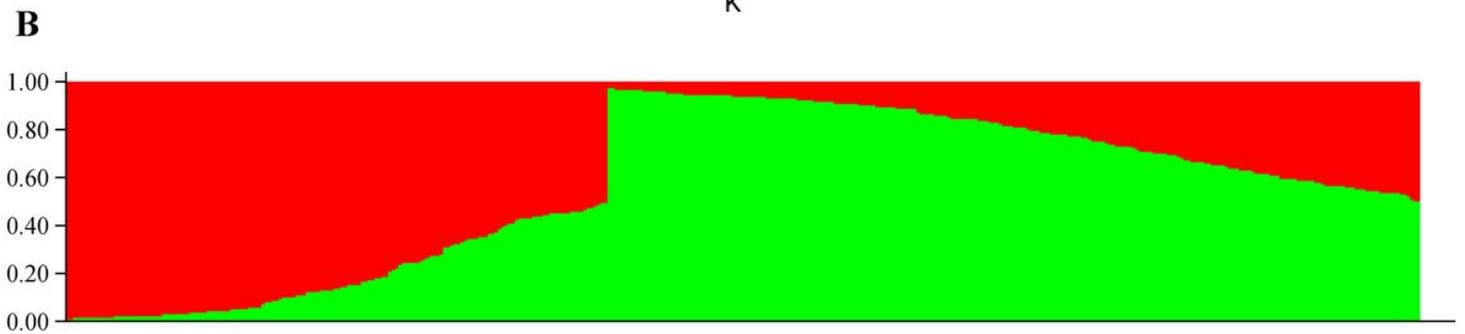
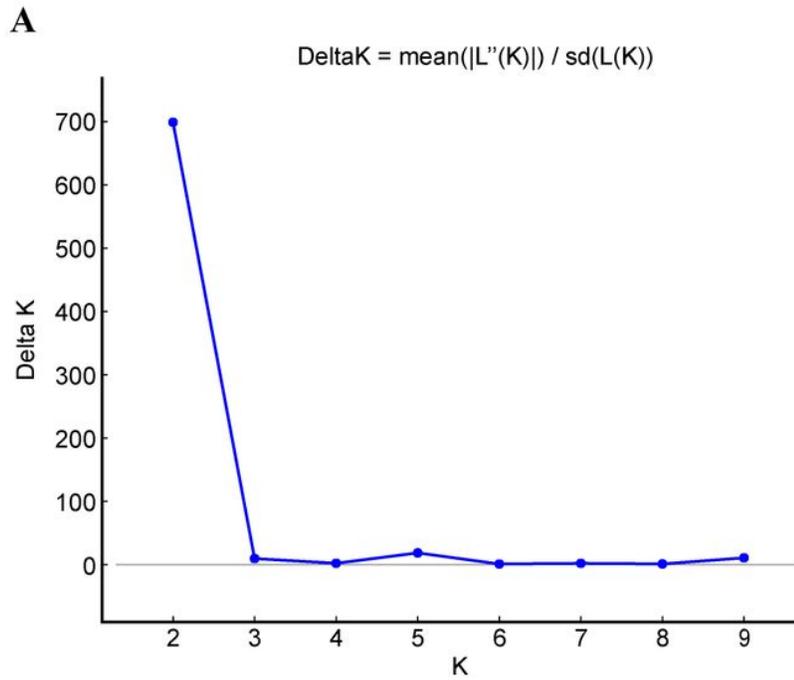


Figure 6

Population structure of 392 broccoli accessions based on 100 KASPs. (A) ΔK values plotted as the number of subpopulations. (B) Subpopulations ($K = 2$) inferred using STRUCTURE.



ZAAS-2

Figure 7

Barcode (2D) of a representative variety of broccoli used in the present study.

Supplementary Files

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

- [Supplementaryinformation.docx](#)
- [Additionalfile4.accessionsofStructure.pdf](#)
- [Additionalfile3.Genotypingofallbroccoli.xlsx](#)
- [Additionalfile2.InformationofKASPmarkers.xlsx](#)
- [Additionalfile1.Statisticsofressequencingresults.xlsx](#)