RESEARCH ARTICLE

Comparative Analysis of Disease-Linked Single Nucleotide Polymorphic Markers from *Brassica rapa* for Their Applicability to *Brassica oleracea*

Young-Il Cho, Yul-Kyun Ahn*, Swati Tripathi, Jeong-Ho Kim, Hye-Eun Lee, Do-Sun Kim

Vegetable Research Division, National Institute of Horticultural & Herbal Science, Rural Development Administration, Suwon, Republic of Korea

* aykyun@korea.kr

Abstract

Numerous studies using single nucleotide polymorphisms (SNPs) have been conducted in humans, and other animals, and in major crops, including rice, soybean, and Chinese cabbage. However, the number of SNP studies in cabbage is limited. In this present study, we evaluated whether 7,645 SNPs previously identified as molecular markers linked to disease resistance in the *Brassica rapa* genome could be applied to *B. oleracea*. In a BLAST analysis using the SNP sequences of *B. rapa* and *B. oleracea* genomic sequence data registered in the NCBI database, 256 genes for which SNPs had been identified in *B. rapa* were found in *B. oleracea*. These genes were classified into three functional groups: molecular function (64 genes), biological process (96 genes), and cellular component (96 genes). A total of 693 SNP markers, including 145 SNP markers [BRH—developed from the *B. rapa* genome for high-resolution melt (HRM) analysis], 425 SNP markers (BRP—based on the *B. rapa* genome that could be applied to *B. oleracea*), and 123 new SNP markers (BRS—derived from BRP and designed for HRM analysis), were investigated for their ability to amplify sequences from cabbage genomic DNA. In total, 425 of the SNP markers (BRP-based on *B. rapa* genome), selected from 7,645 SNPs, were successfully applied to *B. oleracea*. Using PCR, 108 of 145 BRH (74.5%), 415 of 425 BRP (97.6%), and 118 of 123 BRS (95.9%) showed amplification, suggesting that it is possible to apply SNP markers developed based on the *B. rapa* genome to *B. oleracea*. These results provide valuable information that can be utilized in cabbage genetics and breeding programs using molecular markers derived from other *Brassica* species.

Introduction

The genus *Brassica* is one of the most important vegetable crop genera in the world. *Brassica* crops provide vegetables, oil, fodder, and condiments and are also valuable sources of dietary

fiber, vitamin C, and other beneficial factors, including several anticancer compounds [1,2]. In addition, *Brassica* species are popular for producing high-quality biodiesel owing to their relatively low levels of polyunsaturated and saturated fatty acids [3].

Among *Brassica* species, *Brassica rapa* (AA, 2n = 20), *Brassica nigra* (BB, 2n = 16), and *Brassica oleracea* (CC, 2n = 18) are diploid, whereas *Brassica juncea* (AABB, 2n = 36), *Brassica napus* (AACC, 2n = 38), and *Brassica carinata* (BBCC, 2n = 34) are amphidiploid (i.e., having combinations of the genomes of these diploid species) [4]. Thus, the *Brassica* genome provides substantial opportunities for studying the divergence of gene function and genome evolution associated with polyploidy, extensive duplication, and hybridization [5]. *Brassica rapa* has a small genome (529 Mb) compared with its close diploid relatives *B. oleracea* (696 Mb) and *B. nigra* (632 Mb) [6,7]. These characteristics are useful for the study of genomic traits. In response to the need for a simple genetic system with favorable genetic attributes for research on *Brassica* species, *B. rapa* has become a model species representing the *Brassica* A genome and is the focus of multifaceted genome projects with the goal of whole-genome sequencing based on the clone-by-clone strategy (http://www.brassica.info) [8].

Single nucleotide polymorphisms (SNPs) are the most common type of variation in DNA [9]. A SNP is a unique nucleotide difference between two DNA sequences. In theory, SNP variations could involve four different nucleotides at a particular site, but actually only two of these four possibilities are usually observed. Thus, in practice, SNPs are biallelic markers, and therefore the information content of a single SNP is limited compared to polyallelic simple sequence repeat (SSR) markers [10,11]. This disadvantage is overcome by the relatively greater abundance and stability of SNP loci compared to SSR loci. The abundance, ubiquity, and interspersed nature of SNPs together with the potential for automatic high-throughput analysis make them ideal candidate molecular markers for the construction of high-density genetic maps, quantitative trait loci (QTL) fine mapping, marker-assisted plant breeding, and genetic association studies [12,13]. In addition, SNPs located in known genes provide a fast alternative to analyzing the fate of agronomically important alleles in breeding populations, thus providing functional markers [14]. SNPs may be used as simple genetic markers, which may be identified in the vicinity of virtually every gene [13]. There is also great potential for the use of SNPs in the detection of associations between allelic forms of a gene and phenotypes, especially for common diseases with multifactorial genetics [15]. SNP discovery has been reported for several plant species, and the frequency of SNPs has shown variation depending on the different genomic regions in plants [8].

Genomes sequencing projects for *Brassica* species, including *B. rapa*, have produced vast amounts of sequence data that will provide useful information for genetic studies [3,16–18]. In total, 21,311 SNPs and 6,753 InDels in the gene space of the *B. rapa* genome were identified by re-sequencing 1,398 sequence-tagged sites (STSs) in eight genotypes [8]. In addition, more than 37,000 SNPs were identified through a comparison of two accessions of the model plant *Arabidosis thaliana* [19]. Cavell et al. [20] reported that the close sequence identity of coding regions (~87%) between the genomes of *Brassica* species and *A. thaliana* would allow for detailed comparative analyses. Such comparative mapping studies [21–23] have allowed for the assignment of orthologous segments in *Brassica* species and *A. thaliana*, enabling the identification of candidate genes that may directly account for *Brassica* QTL. These informational and genomic resources will promote the genome-wide study of DNA polymorphisms in *B. rapa* and will contribute significantly to *Brassica* crop improvement. Furthermore, the availability of *B. rapa* genomic sequence data offers an unprecedented opportunity to conduct detailed comparative analysis of the relationships between *Brassica* species genomes, and between the *Brassica* genome and *A. thaliana* genome. Various DNA markers, including AFLP [24], PCR-based markers [25,26], RFLP [27], and SSRs [28,29], have been studied in *B. oleracea*. However, there

is no information available regarding the comparative marker profile of *B. rapa* and *B. oleracea* genome.

In the present study, we evaluated whether 7,645 SNPs linked to disease resistance from the *B. rapa* genome may be applied to *B. oleracea*. A total of 693 SNP markers, including 145 SNP markers (BRH) developed from the *B. rapa* genome for high-resolution melt (HRM), 425 SNP markers (BRP) based on the *B. rapa* genome that could be applied to *B. oleracea*, and 123 new SNP markers (BRS) derived from BRP and designed for HRM analysis, were found to be useful tools for QTL fine mapping, the development of SNP markers linked to disease resistance, genomics-based breeding, and genetic association studies in *B. oleracea*.

Materials and Methods

Plant materials and DNA extraction

To evaluate the utility of SNP markers based on the *B. rapa* genome for *B. oleracea*, two cabbage varieties, Chungam45 and Bogam3, were selected from 53 cabbage accessions. Plants from each variety were container-grown in a greenhouse at the National Institute of Horticultural and Herbal Science of the Rural Development Administration (RDA). DNA was extracted from fresh, young leaves of two plants using a DNA extraction kit (Qiagen, Hilden, Germany). The relative purity and concentration of the extracted DNA were estimated with ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA), and the final concentration of each DNA sample was adjusted to 20 ng/µL.

Functional analysis

Previously, we developed 21,311 SNPs and 6,753 InDels using the gene space of the *B. rapa* genome by re-sequencing 1,398 STSs in eight genotypes [8]. The sequences of 7,645 of 21,311 SNP markers, which were linked to disease resistance based on the *B. rapa* genome and which aligned to BAC sequences of *B. rapa*, were obtained using FGENESH (http://www.softberry.com) based on a *B. rapa* matrix, in order to confirm the positions of the SNP primers and to analyze the information according to the positions of the corresponding genes. To analyze the biological functions of the predicted genes, the protein sequences of the corresponding genes were extracted from the BAC sequences collected in the *B. rapa* Genome Project (http://www.brassica-rapa.org/BRGP/chromosomeSequence.jsp). hese sequences were analyzed for function using the UniProt database (Table 1). A functional analysis of each protein was conducted according to its characteristics using MIPS, FunCat, Gene Ontology (GO), and Clusters of Orthologous Groups. The unigenes of 7,645 SNP markers, which were designed from the *B. rapa* genome, were mapped to the *B. oleracea* and *A. thaliana* genomes using the BLAST (version 2.2.24) program with an e-value of 1e-4 (top match: 1).

Primer design and PCR

The selected 7,645 SNP markers were previously designed from flanking exon sequences of the selected genes to amplify genic regions, including introns, by means of the Primer3 program [8,30]. In total, 693 SNP primers were developed and used in this study. Of these, 145 BRH primers were newly developed for HRM analysis using Primer3, and 425 BRP primers that could be applied to *B. oleracea* were selected from the 7,645 SNP primers. A total of 123 BRS primers were newly designed using CLC Genomic Workbench (CLC bio, Aarhus, Denmark) and Primer3. The amplification reactions were carried out in a total volume of 20 µl containing 40 ng of genomic DNA as template, 0.5 µM forward and reverse primers, and 2 × GoTaq Green Master Mix (Promega, Madison, WI, USA) following the manufacturer's recommended

Table 1. Databases used to compare 7,645 SNP markers developed from the *Brassica rapa* genome with the *Brassica oleracea* genome.

Analysis	Content	URL
Blocks	Multiply aligned ungaped segments corresponding to the most highly conserved regions of proteins	http://blocks.fhcrc.org
COG	Clusters of orthologous groups	http://www.ncbi.nlm.nih.gov/COG
EMBL	First DNA database in Europe	http://www.ebi.ac.uk/embl
FunCat	Functional catalog from MIPS	http://mips.helmholtz-muenchen.de/proj/funcatDB
PDB	Database of protein and nucleic acid three dimensional structure	http://www.rcsb.rog/pdb/home/home.do
Pfam	Protein domain database from the Sanger Center	http://www.sanger.ac.uk/Software/Pfam
PROSITE	Databases from the Swiss Institute of Bioinformatics	http://www.expasy.ch/prosite
SCOP	Structural classification of proteins	http://scop.mrc-lmb.cam.ac.uk/scop
UniProt	Universal Protein Resource(SWISS-PROT and TrEMBL)	http://www.ebi.uniprot.org/index.html
GO	Gene Ontology	http://www.geneontology.org

doi:10.1371/journal.pone.0120163.t001

protocols. PCR was conducted as follows: 95°C for 2 min, followed by 35 cycles of 95°C for 30 s; 55, 57, 60, 62, or 65°C for 30 s; and 72°C for 1 min, with a final extension at 72°C for 10 min using an Eppendorf Thermocycler (Eppendorf, Germany). Electrophoresis on a 1.0% agarose gel with ethidium bromide confirmed the presence of the amplified products.

Results and Discussion

Comparison of the *B. rapa* SNP primer sequences and with the *B. oleracea* genome

To evaluate the applicability of SNP markers designed from the *B. rapa* genome to *B. oleracea*, all reference sequences of *B. oleracea* (742,612), including mRNAs (1,772), ESTs (59,946), and genome survey sequences (GSSs) (680,894), were assembled from the NCBI database (http://www.ncbi.nlm.nih.gov/) (Table 2). The protein (1,735) and final protein sequences (nrProtein) (20,632) of *B. oleracea* as a non-redundant protein dataset were also assembled for ortholog analysis. In addition, 194,305 ESTs, 198,585 GSSs, 3,965 mRNAs, 2,062 proteins, and 55,420 nrProtein sequences of *B. rapa* were collected from the NCBI database. The close phylogenetic relationship between the *Brassica* species and the model plant *A. thaliana* suggests that the transfer of knowledge from *Arabidopsis* for *Brassica* crop improvement would be straightforward [5]. Extensive gene loss or gain events and large-scale chromosomal rearrangements, including segmental duplications or deletions, in the *Brassica* lineage complicated the orthologous relationships between loci from the two genomes [31]. Hybridization between species is another source of *Brassica* genome complexity. The complex genomic organization of *Brassica* species as a result of multiple rounds of polyploidy and genome hybridization makes the identification of orthologous relationships between genes difficult. The genomes of three diploid species, *B. rapa* (AA, 2n = 20), *B. nigra* (BB, 2n = 16), and *B. oleracea* (CC, 2n = 18), have triplicated homologus counterparts of corresponding segments in the *Arabidopsis* genome as a result of whole-genome triplication, which occurred approximately 12 to 17 million

Table 2. Datasets used to analyze the applicability and orthology of SNP markers based on the *Brassica rapa* genome to the *Brassica oleracea* genome.

Sequence	*Brassica rapa*	*Brassica oleracea*	*Arabidopsis thaliana*
EST	194,305	59,946	
Genome	198,585 (GSS[a])	680,894 (GSS[a])	Chromosome (5)
mRNA	3,965	1,772	-
Protein	2,062	1,735	32,615
nrProtein[b]	55,420	20,632	32,615

a. GSS: Genome Survey Sequence.

b. nrProtein: Final protein sequence as a non-redundant protein dataset for ortholog analysis

doi:10.1371/journal.pone.0120163.t002

years ago [8,32]. In addition, *B. oleracea* and *A. thaliana* diverged 15 to 20 million years ago [33].

The unigenes of 7,645 SNP markers of *B. rapa*, which are related to disease resistance, were mapped to the genomes of *B. oleracea* and *A. thaliana* using BLAST (version 2.2.24) algorithm with an e-value of 1e-4 (Table 3). Among 20,632 unigenes of *B. oleracea* and 32,615 unigenes of *A. thaliana*, 3,914 (18.9%) and 13,506 (41.4%) were orthologous with genes (or protein) specified by the *B. rapa* SNP markers. In total, 8,518 (41.2%) genes for *B. oleracea* and 1,946 (5.9%) genes for *A. thaliana* were unique, respectively. Of 7,645 SNP markers, 425 were applicable to *B. oleracea*, and 142 were applicable to *A. thaliana* (Table 4). The *Brassica* and *Arabidopsis* genomes share about 87% sequence identity in their coding regions [20]. This feature has been extensively exploited and has resulted in a large number of comparative mapping studies between *Brassica* crops and *Arabidopsis* [22]. Li et al. [4] tested sequence-tagged markers from *B. rapa* for homology with the genomic sequence of *A. thaliana*. They found that 223 markers had homologs in the genome of *A. thaliana*, and that these were distributed throughout the genome, except for one homolog, which was located on the short arm of chromosome 2. *Brassica rapa* is diploidy and has a small genome size (529 Mb) compared with its close diploid relative *B. oleracea* (696 Mb) [6]. These characteristics are useful for the study of genomic traits. Previous comparative mapping studies of *Brassica* and *Arabidopsis* using molecular markers revealed extensive synteny between *B. oleracea* and *A. thaliana*, suggesting that knowledge gained in one species can be productively applied to the other [34,35].

Functional annotation and GO

A total of 6,412 of 6,856 genes, which were mapped to 7,645 SNP markers, were annotated with the reference genome data of *B. rapa* using functional annotation analysis (data not shown). Of 7,645 SNP markers, 425 BRP that were applicable to 256 genes in *B. oleracea*, were mapped with the genes of *B. oleracea* (Fig. 1). These genes were classified into three functional

Table 3. Orthologous and unique genes between *Brassica oleracea* and SNP markers based on the *Brassica rapa* genome.

	Brassica rapa	*Brassica oleracea*	*Arabidopsis thaliana*
nrProtein[a] (%)	55,420 (100)	20,632 (100)	32,615 (100)
Orthologous genes (%)	17,810 (32.1)	3,914 (18.9)	13,506 (41.4)
Unique genes (%)	7,165 (12.9)	8,518 (41.2)	1,946 (5.9)

a. nrProtein: Final protein sequence as a non-redundant protein dataset for ortholog analysis

doi:10.1371/journal.pone.0120163.t003

Table 4. The applicability of SNP markers based on the *Brassica rapa* genome to *Brassica oleracea*.

Organism	*Brassica oleracea*	*Arabidopsis thaliana*
Template number	742,612	5
Number of successful PCR markers	425	142
PCR success rate (%)	5.56	1.86

doi:10.1371/journal.pone.0120163.t004

groups: molecular function (64 genes), biological process (96 genes), and cellular component (96 genes). Molecular function was subdivided into 12 categories: protein kinase activity (15), protein binding (8), nucleic acid binding (7), cation binding (5), receptor activity (5), active transmembrane transporter activity (4), hydrolase activity (4), magnesium ion binding (4), carboxypeptidase activity (3), polygalacturonase (3), purine nucleotide binding (3), and transmembrane receptor activity (3). Biological process was subdivided as follows: transcription
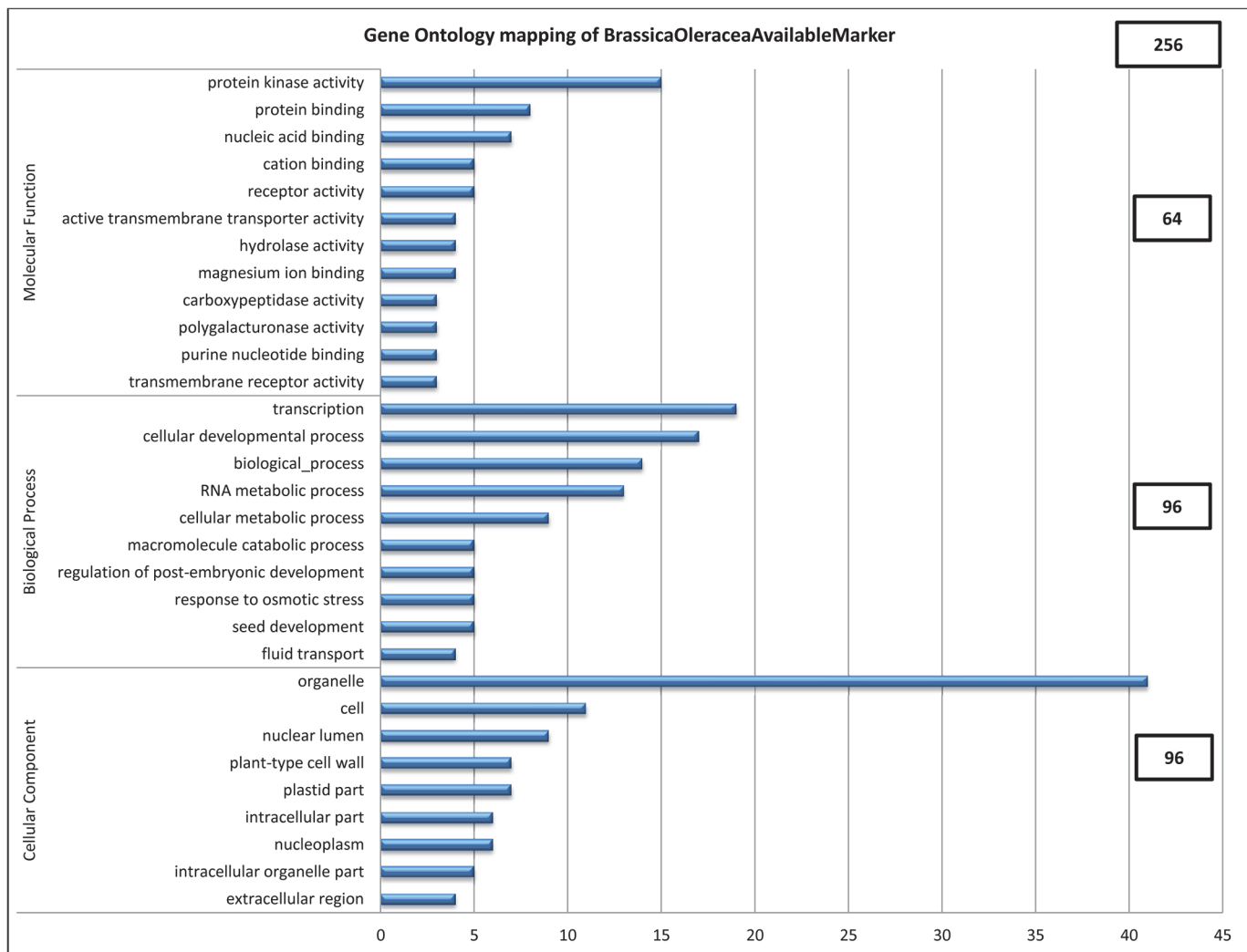

Fig 1. Gene Ontology (GO) mapping of cabbage genes using the sequence of the *Brassica oleracea* genome and 425 BRP based on the *Brassica rapa* genome.

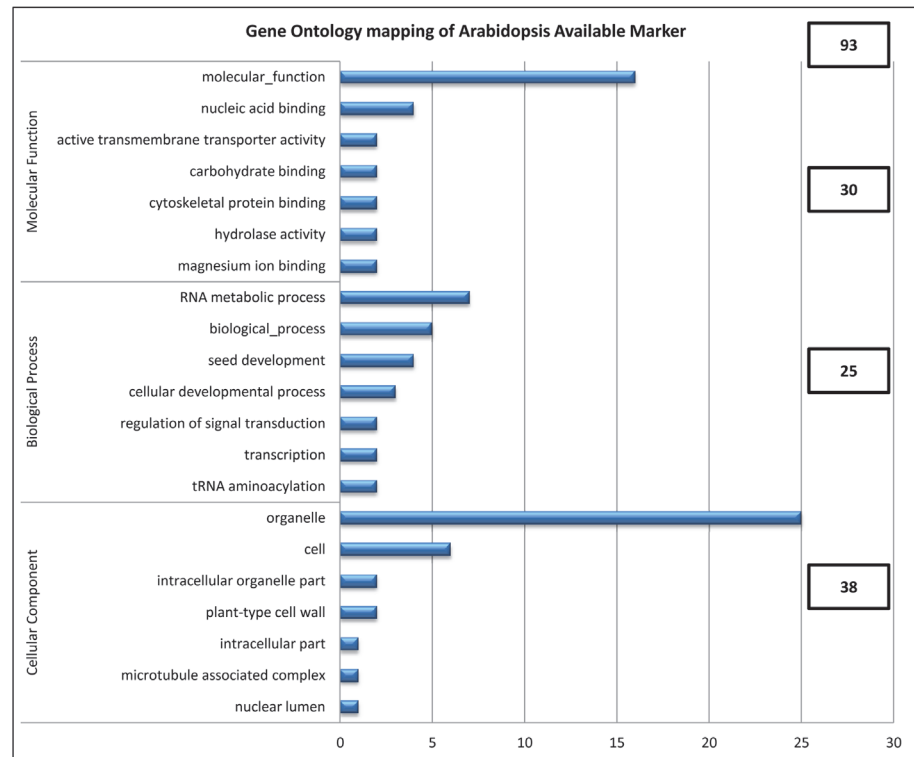doi:10.1371/journal.pone.0120163.g001

Fig 2. Gene Ontology (GO) mapping of *Arabidopsis thaliana* genes and the sequence of 142 SNP markers from *A. thaliana* based on the *B. rapa* genome.

(19), cellular developmental process (17), biological process (14), RNA metabolic process (13), cellular metabolic process (9), macromolecule catabolic process (5), regulation of post-embryonic development (5), response to osmotic stress (5), seed development (5), and fluid transport (4). Cellular component was subdivided into nine categories: organelle (41), cell (11), nuclear lumen (9), plant-type cell wall (7), plastid part (7), intracellular part (6), nucleoplasm (6), intracellular organelle part (5), and extracellular region (4).

The 142 SNPs that can be applied to 93 *A. thaliana* genes, as a result of GO annotation between the SNP sequences of *B. rapa* and sequence of the *A. thaliana* genome, were mapped with the genes of *A. thaliana* (Fig. 2). These genes were clustered into three functional categories: molecular function (30 genes), biological process (25 genes), and cellular component (38 genes). Molecular function was subdivided into seven categories: molecular function (16), nucleic acid binding (4), active transmembrane transporter activity (2), carbohydrate binding (2), cytoskeletal protein binding (2), hydrolase activity (2), and magnesium ion binding (2). Biological process was subdivided into seven categories: RNA metabolic process (7), biological process (5), seed development (4), cellular developmental process (3), regulation of signal transduction (2), transcription (2), and tRNA aminoacylation (2). Cellular component was subdivided seven categories: organelle (25), cell (6), intracellular organelle part (2), plant-type cell wall (2), intracellular part (1), microtubule associated complex (1), and nuclear lumen (1). Comparative genomics is rapidly emerging as a powerful tool for genome analysis and annotation [36]. The course of evolution for functional regions such as exons and regulatory sequences tends to be more conserved than that for nonfunctional regions; thus, local sequence
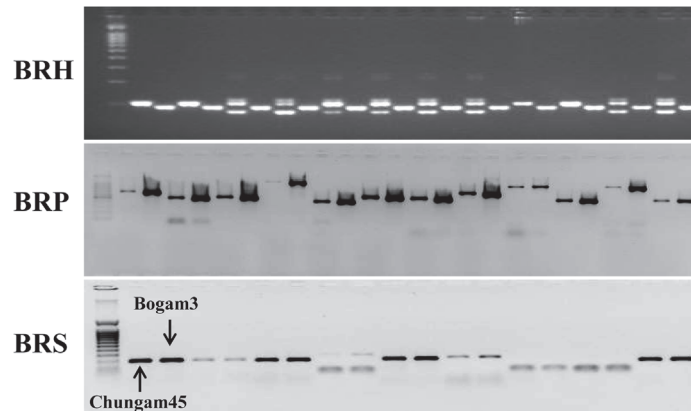
Fig 3. PCR amplification of genomic DNA from two cabbage varieties using SNP primers designed based on the *Brassica rapa* genome.

doi:10.1371/journal.pone.0120163.g003

similarities have implications for biological functionality. Nucleotide sequence conservation between *B. oleracea* and *A. thaliana* has been reported to be in the range of 75–90% for exons, compared to ≤ 70% for introns and intergenic regions. Thus, the genome-scale comparison of *Arabidopsis* with *Brassica* at the sequence level provides an excellent opportunity for the applicability of this phylogenetic footprinting approach to the annotation of plant genomes [37].

Application *B. rapa* SNP primers to *B. oleracea*

In this study, 693 SNP markers designed based on the *B. rapa* genome were tested for their applicability to *B. oleracea* using two cabbage varieties, Chungam45 and Bogam3 (Fig. 3 and Table 5). Of 145 BRH, 108 (74.5%) were amplified. In addition, 415 of 425 BRP (97.6%) were amplified, and 118 of 123 BRS (95.9%) were amplified using genomic DNA from two cabbage varieties. The amplification values for the BRP and BRS were higher than for the BRH. A total of 641 of 693 SNP markers from *B. rapa* were amplified using PCR, suggesting that these markers are beneficial molecular markers for *B. rapa* genetic analyses and breeding and that they can be applied to other *Brassica* species, including *B. oleracea*. These results provide valuable information that can be used for the utilization of *Brassica* in genomic studies and cabbage breeding.

Table 5. Amplification of SNP markers based on the *Brassica rapa* genome to *Brassica oleracea*.

Primer	No.	PCR amplification		
		Amplification	No amplification	Rate (%)
BRH[a]	145	108	37	74.5
BRP[b]	425	415	10	97.6
BRS[c]	123	118	5	95.9

a. BRH: SNP marker developed from the *Brassica rapa* genome for high-resolution melt (HRM) analysis.

b. BRP: SNP marker based on the *B. rapa* genome that could be applied to *B. oleracea*.

c. BRS: New SNP marker derived from a BRP designed for HRM analysis.

doi:10.1371/journal.pone.0120163.t005

Author Contributions

Conceived and designed the experiments: YIC JHK HEL DSK YKA. Performed the experiments: YIC ST JHK. Analyzed the data: YIC ST JHK. Wrote the paper: YIC ST. Guided overall experiments and manuscript: YKA.

References

1. Fashey JW, Florens HE (1995) The role of crucifers in cancer chemoprotection; Talalay P, Gustine DL, editors. USA: American Society of Plant Physiologists, Rockville, MD.

2. Higdon JV, Delage B, Williams DE, Dashwood RH (2007) Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis. Pharmacol Res 55: 224–236. PMID: 17317210

3. Li H, Chen X, Yang Y, Xu J, Gu J, et al. (2011) Development and genetic mapping of microsatellite markers from whole genome shotgun sequences in *Brassica oleracea*. Molecular Breeding 28: 585–596.

4. Li F, Kitashiba H, Inaba K, Nishio T (2009) A *Brassica rapa* linkage map of EST-based SNP markers for identification of candidate genes controlling flowering time and leaf morphological traits. DNA Res 16: 311–323. doi: 10.1093/dnares/dsp020 PMID: 19884167

5. Mun JH, Kwon SJ, Yang TJ, Kim HS, Choi BS, et al. (2008) The first generation of a BAC-based physical map of *Brassica rapa*. BMC Genomics 9: 280. doi: 10.1186/1471-2164-9-280 PMID: 18549474

6. Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, et al. (2005) Evolution of genome size in Brassicaceae. Ann Bot 95: 229–235. PMID: 15596470

7. Ramchiary N, Nguyen VD, Li X, Hong CP, Dhandapani V, et al. (2011) Genic microsatellite markers in *Brassica rapa*: development, characterization, mapping, and their utility in other cultivated and wild Brassica relatives. DNA Res 18: 305–320. doi: 10.1093/dnares/dsr017 PMID: 21768136

8. Park S, Yu HJ, Mun JH, Lee SC (2010) Genome-wide discovery of DNA polymorphism in *Brassica rapa*. Mol Genet Genomics 283: 135–145. doi: 10.1007/s00438-009-0504-0 PMID: 20024583

9. Brookes AJ (1999) The essence of SNPs. Gene 234: 177–186. PMID: 10395891

10. Griffin TJ, Smith LM (2000) Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. Trends Biotechnol 18: 77–84. PMID: 10652512

11. Schlotterer C (2004) The evolution of molecular markers—just a matter of fashion? Nat Rev Genet 5: 63–69. PMID: 14666112

12. Kwok PY (2001) Methods for genotyping single nucleotide polymorphisms. Annu Rev Genomics Hum Genet 2: 235–258. PMID: 11701650

13. Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5: 94–100. PMID: 11856602

14. Fusari CM, Lia VV, Hopp HE, Heinz RA, Paniego NB (2008) Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. BMC Plant Biol 8: 7. doi: 10.1186/1471-2229-8-7 PMID: 18215288

15. Jorde LB (2000) Linkage disequilibrium and the search for complex disease genes. Genome Res 10: 1435–1444. PMID: 11042143

16. Choi SR, Teakle GR, Plaha P, Kim JH, Allender CJ, et al. (2007) The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. Theor Appl Genet 115: 777–792. PMID: 17646962

17. Kim H, Choi SR, Bae J, Hong CP, Lee SY, et al. (2009) Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. BMC Genomics 10: 432. doi: 10.1186/1471-2164-10-432 PMID: 19751531

18. Yang TJ, Kim JS, Lim KB, Kwon SJ, Kim JA, et al. (2005) The Korea brassica genome project: a glimpse of the brassica genome based on comparative genome analysis with Arabidopsis. Comp Funct Genomics 6: 138–146. doi: 10.1002/cfg.465 PMID: 18629219

19. Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, et al. (2002) Arabidopsis map-based cloning in the post-genome era. Plant Physiol 129: 440–450. PMID: 12068090

20. Cavell AC, Lydiate DJ, Parkin IA, Dean C, Trick M (1998) Collinearity between a 30-centimorgan segment of Arabidopsis thaliana chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome 41: 62–69. PMID: 9549059

21. O'Neill CM, Bancroft I (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea var. alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. Plant J 23: 233–243. PMID: 10929117

22. Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, et al. (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171: 765–781. PMID: 16020789

23. Paterson AH, Lan TH, Amasino R, Osborn TC, Quiros C (2001) Brassica genomics: a complement to, and early beneficiary of, the Arabidopsis sequence. Genome Biol 2: REVIEWS1011.

24. van Hintum TJ, van de Wiel CC, Visser DL, van Treuren R, Vosman B (2007) The distribution of genetic diversity in a Brassica oleracea gene bank collection related to the effects on diversity of regeneration, as measured with AFLPs. Theor Appl Genet 114: 777–786. PMID: 17273846

25. Cheng J, Weijun K, Yun L, Jiabo W, Haitao W, et al. (2010) Development and validation of UPLC method for quality control of *Curcuma longa* Linn.: Fast simultaneous quantitation of three curcuminoids. J Pharm Biomed Anal 53: 43–49. doi: 10.1016/j.jpba.2010.03.021 PMID: 20395103

26. Sebastian RL, Howell EC, King GJ, Marshall DF, Kearsey MJ (2000) An integrated AFLP and RFLP Brassica oleracea linkage map from two morphologically distinct doubled-haploid mapping populations. Theoretical and Applied Genetics 100: 75–81.

27. Iniguez-Luy FL, Lukens L, Farnham MW, Amasino RM, Osborn TC (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa* and *B. oleracea*. Theor Appl Genet 120: 31–43. doi: 10.1007/s00122-009-1157-4 PMID: 19784615

28. Iniguez-Luy FL, Voort AV, Osborn TC (2008) Development of a set of public SSR markers derived from genomic sequence of a rapid cycling *Brassica oleracea* L. genotype. Theor Appl Genet 117: 977–985. doi: 10.1007/s00122-008-0837-9 PMID: 18651126

29. Soengas P, Cartea E, Lema M, Velasco P (2009) Effect of regeneration procedures on the genetic integrity of *Brassica oleracea* accessions. Molecular Breeding 23: 389–395.

30. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365–386. PMID: 10547847

31. Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, et al. (2006) Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of *Brassica rapa*. Plant Cell 18: 1339–1347. PMID: 16632644

32. Mun JH, Kwon SJ, Yang TJ, Seol YJ, Jin M, et al. (2009) Genome-wide comparative analysis of the *Brassica rapa* gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. Genome Biol 10: R111. doi: 10.1186/gb-2009-10-10-r111 PMID: 19821981

33. Yang YW, Lai KN, Tai PY, Li WH (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between Brassica and other angiosperm lineages. J Mol Evol 48: 597–604. PMID: 10198125

34. Babula D, Kaczmarek M, Barakat A, Delseny M, Quiros CF, et al. (2003) Chromosomal mapping of *Brassica oleracea* based on ESTs from Arabidopsis thaliana: complexity of the comparative map. Mol Genet Genomics 268: 656–665. PMID: 12589440

35. Lan TH, DelMonte TA, Reischmann KP, Hyman J, Kowalski SP, et al. (2000) An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. Genome Res 10: 776–788. PMID: 10854410

36. Ayele M, Haas BJ, Kumar N, Wu H, Xiao Y, et al. (2005) Whole genome shotgun sequencing of *Brassica oleracea* and its application to gene discovery and annotation in Arabidopsis. Genome Res 15: 487–495. PMID: 15805490

37. Quiros CF, Grellet F, Sadowski J, Suzuki T, Li G, et al. (2001) Arabidopsis and Brassica comparative genomics: sequence, structure and gene content in the ABI-Rps2-Ck1 chromosomal segment and related regions. Genetics 157: 1321–1330. PMID: 11238417