Research Article

Development of InDel markers to identify *Capsicum* disease resistance using whole genome resequencing

Sandeep Karna・Yul-Kyun Ahn

Abstract In this study, two pepper varieties, PRH1 (powdery mildew resistance line) and Saengryeg (powdery mildew resistance line), were resequenced using next generation sequencing technology in order to develop InDel markers. The genome-wide discovery of InDel variation was performed by comparing the whole-genome resequencing data of two pepper varieties to the Capsicum annuum cv. CM334 reference genome. A total of 334,236 and 318,256 InDels were identified in PRH1 and Saengryeg, respectively. The greatest number of homozygous InDels were discovered on chromosome 1 in PRH1 (24,954) and on chromosome 10 (29,552) in Saengryeg. Among these homozygous InDels, 19,094 and 4,885 InDels were distributed in the genic regions of PRH1 and Saengryeg, respectively, and 198,570 and 183,468 InDels were distributed in the intergenic regions. We have identified 197,821 polymorphic InDels between PRH1 and Saengryeg. A total of 11,697 primers sets were generated, resulting in the discovery of four polymorphic InDel markers. These new markers will be utilized in order to identify disease resistance genotypes in breeding populations. Therefore, our results will make a one-step advancement in whole genome resequencing and add genetic resource datasets in pepper breeding research.

Keywords chromosome, InDel markers, next generation sequencing, *Capsicum* disease

S. Karna
Vegetable Research Division, National Institute of Horticultural & Herbal Science, Rural Development Administration, 100, Nongsaengmyeong-ro, Iseo-myeon, Wanju-gun, Jeollabuk-do 55365, Korea
Present Address: BK21 PLUS Life Science, School of Dentistry, Seoul National University, 101 Daehak-Ro, Jongno-Gu, Seoul 03080, Korea

Y. K. Ahn (✉)
Deptartment of vegetable crops, Korea National College of Agriculture and Fisheries, Jeonju-si, Jeollabuk-Do 54874, Korea
e-mail: aykyun@korea.kr

Introduction

Whole-genome sequencing is a laboratory procedure that can be used to determine the entire make-up of an organism's genome, thus facilitating an improved understanding of the genetic variation both within and between species. Next-generation sequencing (NGS) technology has had an incomparable impact on genomic research, since it was first introduced in 2005, and over last 10 years, the substantial advancement in sequencing technology and bioinformatics has instigated a transition from classical conservation genetics to conservation genomics (Allendorf et al. 2010; Primmer 2009). Recently, the Sanger method has been partially replaced with several NGS technologies that provide improved cost-effectiveness and better prospects for the development of DNA-based molecular markers (Fakrudin et al. 2006).

Such molecular markers are important tools in both basic and applied research for identifying variation, investigating genetic diversity, genotypic fingerprinting, phylogenetic analysis, determining functional genetic variation, and marker-assisted breeding to improve the horticultural, nutritional, and medicinal value of crops (McCouch et al. 1997; Nagaraju et al. 2002; Vos et al. 1995). Insertion/Deletion (InDel) markers are the most abundant and uniformly distributed form of structural variation in genomes (Pacurar et al. 2012), and after single nucleotide polymorphisms (SNPs), InDels are the second most common type of genetic variation in both humans and plants (Pena and Pena 2012; Vali et al. 2008). InDels are also ubiquitously distributed, with massive variation in size, and occur nearly as frequently as SNPs (Montgomery et al. 2013). Accordingly, with advancements in NGS, InDels have become the most-favored and −frequently used markers for plant breeding (Jiang et al. 2013; Ollitrault et al. 2012; Vasemagi et al. 2010).

Pepper fruits (*Capsicum* spp.) are an important component of fresh vegetable markets, owing to their use as vegetables, spices, condiments, and coloring agents, with additional value,

Fig. 1 Parental lines of powdery mildew are inoculated. (A) Saengryeg, susceptible line and (B) PRH1, resistant line

owing to their use as medicines and as a source of vitamins (Marin et al. 2004; Mejia et al. 1998; Sun et al. 2007). *Capsicum* belongs to the Solanaceae family, which includes over 3,000 species that have 12 chromosomes but significantly different genome sizes. The most common species for chili pepper are *C. annuum, C. frutescens, C. chinense, C. pubescens,* and *C. baccatum* (Pickersgill 1997). *Capsicum* disease such as bacterial wilt and powdery mildew causes significantly reduce yield and quality of pepper fruit. Therefore, it is important to develop InDel markers in *Capsicum* and to apply these markers to associate important traits, such as disease resistance. In this study, we have used two pepper varieties PRH1 (powdery mildew resistance line) and Saengryeg (powdery mildew resistance line) (Fig. 1). The objective of this research was to develop novel InDel markers that related to the powdery mildew disease resistant.

Material and Methods

Isolation of genomic DNA from pepper plants

Young leaves were collected from PRH1 and Saengryeg plants that were cultivated in a greenhouse at the National Institute of Horticultural and Herbal Science, Rural Development Administration, Jeonju, Republic of Korea. The leaves (200 mg) were pulverized in liquid nitrogen, and high-quality DNA was extracted using the CTAB method (Doyle and Doyle 1987), with slight modification. Briefly, the pulverized samples were incubated in CTAB buffer (2% CTAB, 20 mM EDTA, 100 mM Tris-HCl pH 8, 1.4 M NaCl, and 1% w/v PVP at 65°C for 10 min. After cooling to room temperature, chloroform was added, and the sample mixtures were thoroughly mixed and centrifuged at 13,000 xg for 5 min at 4°C. The supernatants were transferred to new
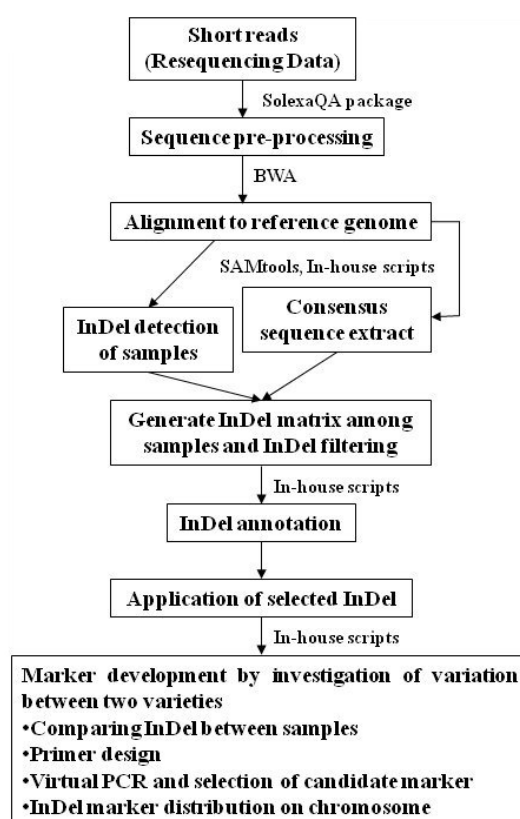

Fig. 2 Flow chart for the discovery of InDel

tubes, equal volumes of absolute ethanol were added, and the mixtures were centrifuged at 13,000 xg for 5 min at 4°C. This time, the supernatants were discarded, followed by the addition of 70% ethanol (v/v) and centrifugation at 13,000 xg for 5 min at 4°C. Again, the supernatants were discarded, and the precipitated DNA pellets were dried at room temperature.

Subsequently, the resulting DNA pellets were used as starting material for DNA purification, using the GenElute

plant DNA isolation kit (Sigma-Aldrich Co., St. Louis, MO, USA). The quality of the purified DNA was assessed *via* electrophoresis on a 1% agarose gel. The concentration of extracted DNA was estimated using a NanoVue spectrophotometer (GE Healthcare Bio-Science, Little Chalfont, UK), the DNA quality was further examined using the NanoVue spectrophotometer, by ensuring a single absorbance peak at 260 nm, a 260/280 absorbance ratio of 1.8-2.0, and no evidence of substantial band shearing or contamination (either RNA or polysaccharide).

DNA library construction and massively parallel sequencing

The purified genomic DNA was randomly sheared using an S2 focused-ultrasonicator (Covaris, Woburn, MA, USA) to yield DNA fragments in the target range of 400-500 bp, and the average size of the fragmented molecules was assessed using an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). The resulting overhangs were converted to blunt ends, using the TruSeq DNA Sample Preparation Kit v2 (Illumina, San Diego, CA, USA), followed by a cleanup step, using AMPure XP beads (Beckman Coulter Genomics, Danvers, MA, USA). In order to increase the success of ligation between the fragmented DNA and index adapters, as well as to reduce self-ligation of the blunt-end fragments, the 3′ ends were adenylated and immediately ligated to the index adapters, after which the fragments were purified using AMPure XP beads. The sample ligation products were then size selected (400 ~ 500 bp) on a 2% agarose gel, followed by gel extraction, column purification, and PCR amplification, using adapter-specific primers. The PCR products were purified using AMPure XP beads (Beckman Coulter Genomics) and the average molecule size of the libraries was assessed using an Agilent Bioanalyzer 2100 (Agilent Technologies). Finally, each library (n = 2) was loaded on a HiSeq 4000 platform, and we performed high-throughput sequencing to ensure a 10-fold average sequencing depth for each sample.

Sequence pre-processing

After sequencing, reads were assessed for quality and impractical parts of the reads were discarded, since high-quality reads are required for *de novo* genome assembly. The raw reads were trimmed using the Solexa QA v.1.13 package (Cox et al. 2010), and since it is common for the bases at either end of Illumina reads to drop in quality, we removed the read ends with Phred quality scores below Q = 20 (i.e., above 0.05 probability of error), using the DynamicTrim

algorithm in Solexa QA. In addition, we also removed 5′ and 3′ stretches of ambiguous 'N' nucleotides and a minimum sequence length of 25 bp was applied, using the LengthSort algorithm. These data were then used for downstream analysis, and the 3.48 Gb reference genome of *C. annuum* cv. CN334 (v1.55) was downloaded from the Sol Genomics Network (http://www.sgn.cornell.edu/) (Kim et al. 2014).

Alignment and detection of InDels

Genome-wide InDels were identified using an improved BWA-SAMtools workflow (Li and Durbin 2010; Li et al. 2009). To align our reads to the *C. annuum* cv. CN334 (v1.55) reference genome, we applied the Burrows-Wheeler Aligner (BWA 0.6.1-r104; (Li and Durbin 2010), using the default values for mapping, except for seed length (-l) = 30, maximum differences in the seed (-k) = 1, number of threads (-t) = 16, maximum number of gap extensions (-e) = 50, mismatch penalty (-M) = 6, gap open penalty (-O) = 15, and gap extension penalty (-E) = 8. Mapped reads were extracted from the resulting BAM file using SAMtools 0.1.16 (Li et al. 2009) for further analyses.

SAMtools, which has been widely applied for data mining, owing to its diverse modules for mapping statistics, variant calling, and file conversion, was used to ensure reliable mapping of the reads, which is important for variant calling. Using the varFilter command, InDels were only called for variable positions with a minimal mapping quality (-Q) of 30. The minimum and maximum of read depths were set 3 and 100, respectively, and depending on the ratio of InDel reads to mapped reads, variant types were classified into three categories: homozygous InDels (for variants with over 90% identical reads), heterozygous InDels (for variants with over than 40% and less than 60% identical reads), and others. A variant was classified as homozygous when non-reference alleles were detected and heterozygous when both reference and non-reference alleles were detected. Further discovered InDels were classified into genic and intergenic regions because InDel classification is important to reduce number of false positive candidate.

In addition, we investigated the InDel distribution pattern in various types of genomic regions, i.e., intergenic regions, introns, coding sequences, intergenic coding sequences, and investigated the distribution of homozygous- and heterozygous-type InDels in both on the chromosomes of both pepper varieties. Furthermore, InDels that were polymorphic between the two samples were selected as candidate markers if their reads contained sufficient sequences on both sides

of the InDel and if there were no structural variations adjacent to the InDel that would interfere with primer design. Then, to design flanking primers, we used an in-house script and Primer3 (v2.3.5) (Untergasser et al. 2012), with the following parameters: primer length = 18-24 bp, with 20 bp as the optimum; primer GC% = 20-80%, with 50% as the optimum; primer Tm = 55-65°C, with 60°C as the optimum; and product size range = 80 ∼ 600 bp. After the designed primers were mapped to the genome sequence and only aligned primers were selected as candidate for InDel marker.

HRM analysis

HRM analyses were performed in 20 ml of total reaction mixture containing 2 ml of DNA extract (200 ng), 1 x of SsoFast Evagreen Supermix (Bio-Rad Laboratories, Hercules, CA, USA), and 200 nM of forward and reverse primers. The reactions were performed in a fluorometric thermal cycler CFX96 real-time systems (Bio-Rad Laboratories, Hercules, CA, USA), following program: 98°C for 2 min, 45 cycles at 98°C for 5 s and 60°C for 10 s. PCR products were denatured at 95°C for 30 s and than annealed at 65°C for 5 min, in order to allow correct annealing of the DNA double strands. The melt curve ranged from 65°C to 95°C with temperature increments of 0.5°C every 10 s and with acquisition of fluorescence data at the end of each melting temperature. The data were analyzed using the Precision Melt Analysis Software 1.2 (Bio-Rad Laboratories, Hercules, CA, USA) to create melting curves as a function of temperature and difference curves for easier visualization of the clusters. Samples were analyzed in triplicates in two independent assays.

Results and Discussion

The advancement of NGS technology has provided more efficient genome sequencing at a significantly lower cost than ever before. NGS technology produces enormous quantities of DNA sequence data, thus providing powerful tools for discovering high-density markers. NGS also facilitates the resequencing of a large number of genomes, which provides exceptional prospects for discovering DNA polymorphisms and enables us to discover genome-wide variation on a large scale, thus providing an outstanding resource to the plant science community (Albert and Chang 2014). Along with the rapid advancement in sequencing technology, significant progress has also been made in the performance, sensitivity, and specificity of InDel identification software (Neuman et

al. 2013). InDel identification accuracy can also be improved by increasing read depth per locus, and alignment parameters can be adjusted to minimize non-specific amplification, by increasing mismatches during the discovery of unique loci in the genome. Hence, accuracy might be improved by reducing the number of mismatches in the analysis of variation between populations.

A total of 334,236 and 318,256 InDels were identified in PRH1 and Saengryeg, respectively. In PRH1, 217,690 homozygous InDels, 19,534 heterozygous InDels, and 97,012 other types of InDels were identified, and in Saengryeg, 188,356 homozygous InDels, 20,525 heterozygous InDels, and 109,375 other types of InDels were recognized. Furthermore, the InDels were evenly distributed across the chromosomes (n = 12). The largest numbers of homozygous and heterozygous InDels were found on chromosome 1 (24,954 and 2,242, respectively) in PRH1 and on chromosome 10 (29,552 and 2,553, respectively) in Saengryeg, whereas the smallest numbers of homozygous and heterozygous InDels were found on Chromosome 8 in both varieties (12,000 and 999 in PRH1, respectively, and 5,264 and 823 in Saengryeg; Table 1, Fig. 3).

InDels have been recognized as the second most abundant source of genetic markers and are widely spread across genomes. In addition, InDels are capable of causing genetic diseases, since their presence frequently affects translation by generating frame-shifts that create new stop codons, and as a result, InDels are the most common structural variant to cause pathogenesis (Stenson et al. 2012), by modifying gene expression and functionality. A total of 333,998 and 318,126 InDels were assigned to genomic structures in PRH1 and Saengryeg, respectively. We identified 302,289 intergenic InDels and 31,709 genic InDels in PRH1 and 309,452 intergenic InDels and 8,674 genic InDels in Saengryeg (Table 2). In addition, we found that most of the InDels were located in intergenic regions, followed by introns, coding sequences, and intergenic coding sequences, with total number of a total of 19,086 and 4,884 genic InDels in PRH1 and Saengryeg, respectively. Recently, InDels have become important DNA markers, owing to their profusion, stability, efficiency, codominance, and readily automated lab techniques. As a consequence, InDel markers have become valuable for genetic diversity analysis, high-density genetic map construction, QTL mapping, and the design of genomic selection strategies (Hou et al. 2010, Mills et al. 2006, Pacurar et al. 2012, Temnykh et al. 2001, Wu et al. 2013).

In total, 782,576 InDel polymorphisms were identified between PRH1 and Saengryeg. Among them, 197,821 InDels were located at polymorphic loci, and 20,181 InDels were

Table 1 Chromosome wise distribution of InDel for PRH1 and Saengryeg

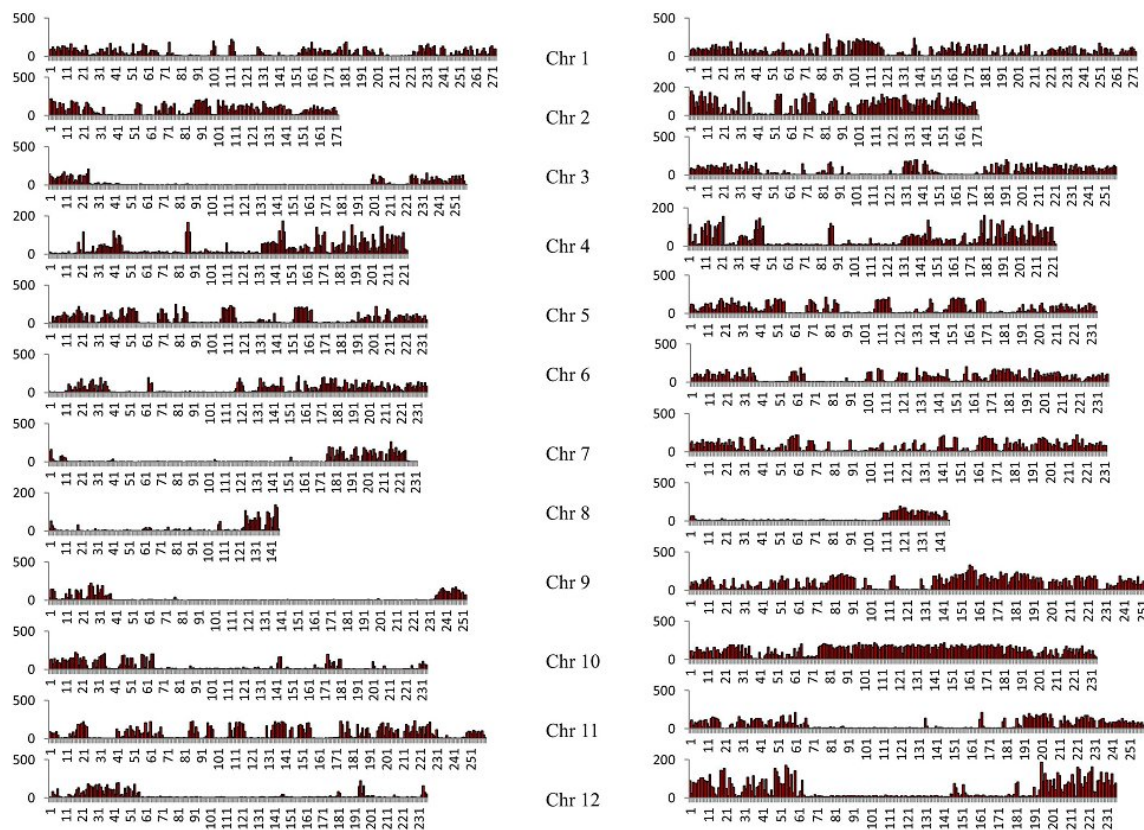| Pepper Chromosome	PRH1		Saengryeg	
	Homozygous	Heterozygous	Homozygous	Heterozygous
Chromosome 1	24, 954	2,242	20,642	2,062
Chromosome 2	17,504	1,597	11,993	1,441
Chromosome 3	23,279	1,974	14,880	1,853
Chromosome 4	16,394	1,610	8,631	1,155
Chromosome 5	15,502	1,492	15,859	1,741
Chromosome 6	18,250	1,690	15,116	1,756
Chromosome 7	19,954	1,660	17,269	1,981
Chromosome 8	12,000	999	5,264	823
Chromosome 9	15,061	1,346	25,195	2,000
Chromosome 10	17,975	1,815	29,552	2,553
Chromosome 11	18,971	1,446	14,268	1,627
Chromosome 12	17,820	1,647	9,684	1,518
Total	217,664	19,518	188,353	20,510

Fig. 3 InDel distribution per 1Mb chromosome (left for PRH1 and right for Saengryeg). The distribution of InDel detected with resequencing of pepper varieties along 12 chromosomes. The horizontal x-axis represents the length (Mb) of the chromosome and y-axis represents the number of InDel

located at non-polymorphic loci. A total of 11,697 primer sets for high resolution melting (HRM) analysis were generated for the polymorphic loci. In addition, we discovered four polymorphic InDel markers on chromosome 1, 4, 8, and

10 (Table 3). As per our parameters for selecting HRM marker primers, a total number of 11,597 marker pairs were selected. We have discovered four InDel markers that were located on chromosomes 1, 4, 8 and 10. Finally, we check

Table 2 Summary of InDel classification by genome structure for PRH1 and Saengryeg

Sample	Total no. of InDels	Region	Total	Homozygous	Heterozygous	Others
PRH1	333,998	Intron CDS	7	3	0	4
		Intergenic CDS	8	4	1	3
		CDS intergenic	10	5	1	4
		CDS intron	3	1	0	2
		CDS	3,364	1,756	319	1,289
		Intron	28,308	17,317	1,687	9,304
		Genic region	31,709	19,094	2,009	10,606
		Intergenic region	302,289	198,570	17,509	86,210
Saengryeg	318,126	Intron CDS	0	0	0	0
		Intergenic CDS	2	2	0	0
		CDS intergenic	0	0	0	0
		CDS intron	0	0	0	0
		CDS	1,355	703	124	528
		Intron	7,315	4,179	538	2,598
		Genic region	8,674	4,885	662	3,127
		Intergenic region	309,452	183,468	19,848	106,136

Table 3 The sequence and annotations of four InDel markers that were develop in this study

InDel marker	Position	Genic / Intergenic	Gene feature	5' Primer sequence	3' Primer sequence
InDel 1	Chromosome 1 (21,146,352)	Intergenic		AACTTGGTAGCAATTTTATTGGGT	TGGAGACAATGTGCATAAGTCTCT
InDel 2	Chromosome 4 (220,561,193)		Intron	TCTTTCCAACCAGTCTGCTCT	ACTTCTGATGTTTTGCCTGT
InDel 3	Chromosome 8 (144,732,433)		Intron	AACGGACAGGAAACACTTGA	TGATGATGCAGCTAGCTGGG
InDel 4	Chromosome 10 (182,186,018)		Intron	AGAGATGCAGTTTGCCACCA	GAGGGCAAACTAGGTGAGCT
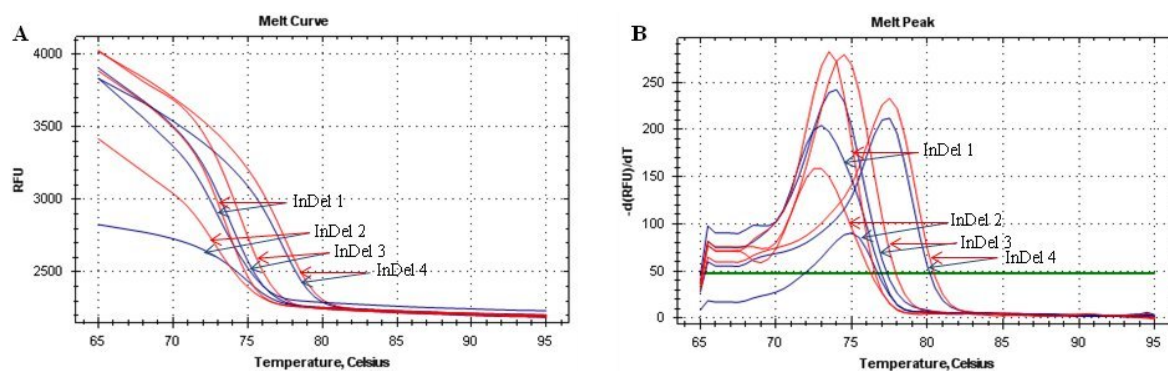
Fig. 4 High-resolution melting (HRM) analysis with InDel markers (red: Saengryeg and blue: PRH1). (A) melt curve and (B) melt peak using InDel markers 1 to 4

DNA amplification of discovered markers using High-resolution melting (HRM) analysis. Tested primers generate different melting profiles allowing the differentiation of *capsicum* varieties (Fig. 4). We have also performed bio-assay for powdery mildew experiment that is provided on supplementary Fig. 1.

In the present study, we identified a large set of high quality InDel markers from the DNA libraries of the *C. annuum* varieties PRH1 and Saengryeg and the majority was located within intergenic regions (> 90%), whereas < 10% were located in genic regions. Polymorphism among homologous-type InDels is the foundation for discovering InDel markers (Liu et al. 2015). InDel polymorphisms exhibit high variability, co-dominant inheritance, and other characteristics of user-friendly-type markers (Pacurar et al. 2012). This study provides extensive InDel data (supplementary Table 1) that can be used for constructing high-resolution genetic maps, performing genome-wide association studies, and designing genomic selection strategies in *C. annuum.*

Acknowledgement

References

Albert VA, Chang TH (2014) Evolution of a hot genome. Proc Natl Acad Sci U S A 111:5069-5070

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697-709

Cox MP, Peterson DA, Biggs PJ (2010) SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11:485

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11-15

Fakrudin B, Vijaykumar, Kambalpally KB. Patil BV, Kuruvinashetti MS (2006) DNA-based marker systems and their utility in entomology. Entomologica Fennica 17

Hou X, Li L, Peng Z, Wei B, Tang S, Ding M, Liu J, Zhang F, Zhao Y, Gu H, Qu LJ (2010) A platform of high-density INDEL/ CAPS markers for map-based cloning in Arabidopsis. Plant J 63:880-888

Jiang GQ, Yao XF, Liu CM (2013) A Simple CELI Endonuclease-Based Protocol for Genotyping both SNPs and InDels. Plant Molecular Biology Reporter 31:1325-1335

Kim S, Park M, Yeom SI, Kim YM et al. (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270-278

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595

Li H, Handsaker B, Wysoker A, Fennell T et al (2009) Genome Project Data Processing: The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079

Liu J, Qu J, Yang C, Tang D, Li J, Lan H, Rong T (2015) Development of genome-wide insertion and deletion markers for maize, based on next-generation sequencing data. BMC Genomics 16:601

Marin A, Ferreres F, Tomas-Barberan FA, Gil MI (2004) Characterization and quantitation of antioxidant constituents of sweet pepper (*Capsicum* annuum L.). J Agric Food Chem 52: 3861-3869

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol 35:89-99

Mejia LA, Hudson E, deMejia EG, Vazquez F (1998) Carotenoid content and vitamin-A activity of some common cultivars of Mexican peppers (*Capsicum* annuum) as determined by HPLC. J. Food Sci. 53:1448-1451

Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, Devine SE (2006) An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res 16: 1182-1190

Montgomery SB, Goode DL, Kvikstad E, Albers CA et al. (2013) The origin, evolution, and functional impact of short insertion-deletion variants identified in 179 human genomes. Genome Res 23: 749-761

Nagaraju J, Kathirvel M, Kumar RR, Siddiq EA, Hasnain SE (2002) Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc Natl Acad Sci U S A 99:5836-5841

Neuman JA, Isakov O, Shomron N (2013) Analysis of insertion-deletion from deep-sequencing data: software evaluation for optimal detection. Brief Bioinform 14:46-55

Ollitrault F, Terol J, Martin AA, Pina JA, Navarro L, Talon M, Ollitrault P (2012) Development of indel markers from Citrus clementina (Rutaceae) BAC-end sequences and interspecific transferability in Citrus. Am J Bot 99:e268-273

Pacurar DI, Pacurar ML, Street N, Bussell JD, Pop TI, Gutierrez L, Bellini C (2012) A collection of INDEL markers for map-based cloning in seven Arabidopsis accessions. J Exp Bot 63: 2491-2501

Pena HB, Pena SD (2012) Automated Genotyping of a Highly Informative Panel of 40 Short Insertion-Deletion Polymorphisms Resolved in Polyacrylamide Gels for Forensic Identification and Kinship Analysis. Transfus Med Hemother 39:211-216

Pickersgill B (1997) Genetic resources and breeding of *Capsicum* spp. Euphytica. 96:129-133

Primmer CR (2009) From conservation genetics to conservation genomics. Ann N Y Acad Sci 1162:357-368

Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN (2012) The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. Curr Protoc Bioinformatics Chapter 1: Unit 13

Sun T, Xu Z, Wu CT, Janes M, Prinyawiwatkul W, No HK (2007) Antioxidant activities of different colored sweet bell peppers

(*Capsicum* annuum L.). J Food Sci 72:S98-102

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441-52

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3-new capabilities and interfaces. Nucleic Acids Res 40:e115

Vali U, Brandstrom M, Johansson M, Ellegren H (2008) Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. BMC Genet 9:8

Vasemagi A, Gross R, Palm D, Paaver T, Primmer CR (2010) Discovery and application of insertion-deletion (INDEL) polymorphisms for QTL mapping of early life-history traits in Atlantic salmon. BMC Genomics 11:156

Vos P, Hogers R, Bleeker M, Reijans M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-4414

Wu DH, Wu HP, Wang CS, Tseng HY, Hwu KK (2013) Genome-wide InDel marker system for application in rice breeding and mapping studies. Euphytica. Euphytica: 131-143