Construction of DNA fingerprint for chili pepper varieties using SNAP markers

R T Terryana*, H Rijzaani, T P Priyatno, I Manzila and P Lestari

Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), Indonesian Agency for Agricultural Research and Development, Ministry of Agriculture, Jalan Tentara Pelajar No. 3A, Bogor 16111, West Java, Indonesia

*E-mail: re2n_terryana@ymail.com

Abstract. Establishing the genetic identity of crop varieties has been considered essential for protecting plant breeder and farmer rights, particularly in developing countries like Indonesia. DNA fingerprint using molecular markers is important to give an unambiguous characteristic pattern as a valuable tool for genetic identification. In this study, eight Single Nucleotide Amplified Polymorphism (SNAP) markers were developed and applied to fingerprint 23 varieties of chili pepper. Polymorphism Information Content (PIC) detected in each primer ranged from 0.14 to 0.36 with an average of 0.17. The average of gene diversity was 0.20 among all varieties for total SNAP markers. A phylogenetic tree was subsequently constructed based on their genotypic scores for selected six markers, which separated the 23 varieties into three major groups. The cluster consisted of 2, 5 or 16 varieties. The DNA fingerprints were translated into capital letters representing presence and absence of allele, and they revealed the specific identity of five varieties. A number of varieties possessed the same DNA fingerprint profiles indicating their close genetic distance. Eventhough these SNAP markers were not able to distinguish each variety according to its unique allelic composition, this study could serve as preliminary information to establish genetic fingerprints of chili pepper varieties in Indonesia. Similar studies in the future will benefit from the SNAP found in this study.

Keywords: chili pepper, DNA fingerprint, SNAP.

1. Introduction

Chili pepper belongs to the genus *Capsicum* of the family Solanaceae. It is one of the most important vegetable-spice crops cultivated in tropical regions such as Indonesia. Fruits of chili pepper plants are among the most heavily consumed spices in the world due to their unique colour, taste, pungency, flavour and aroma [1]. Chili pepper is a facultative cross-pollinated crop, and hence, exhibits wide variability for different qualitative and quantitative traits [2,3]. The genus *Capsicum* has a broad genetic diversity, most of these are found growing in the wild and are believed not to have been domesticated. Chili pepper that has been domesticated and cultivated widely in the world comprises five species: *C. annuum*, *C. frutescens*, *C. chinense*, *C. baccatum* and *C. pubescens*. Among these species, *C. annuum* and *C. frutescens* are the main commercial chili pepper traded and cultivated in Indonesia [4].

Genetic identification is critically important in crop plant variety protection. Protection can be granted if the genetic identity of a variety has been proved to be distinct from existing varieties. The uniqueness of a variety is established by tests for distinctiveness. Due to technical limitations, the authentic genetic identification is mainly based on morphological and physiological characters, which are affected by environmental conditions and are often subjective decisions. As a result, different varieties may be difficult to effectively distinguish and arbitrate due to lack of effective species identification methods. Thus, it is an urgent need to establish a set of steady, reliable and easily accessible identification methods for chili pepper varieties to effectively protect their intellectual property rights [5]. Development of molecular marker technology would make it possible to quickly and accurately identify varieties at DNA level, since this technology is not affected by environmental conditions and should be more reproducible and objective [6].

Molecular markers can display the differences in nucleotide sequences which are suitable for DNA fingerprinting of crop varieties [7]. In the last decade, several molecular markers, such as Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Insertions and Deletions (Indel), Inter Simple Sequence Repeats (ISSR) and Simple Sequence Repeats (SSR), were used in chili pepper for developing high density genetic maps [8,9,10], genetic diversity evaluation [11] and are prospective for Marker-assisted Selection [12]. Among all molecular markers, Single Nucleotide Polymorphism (SNP) is the most abundant, robust and feasible for automated high-throughput genotyping [13].

SNP is a single nucleotide DNA variation at specific locations throughout the plant genome. The easiest, most rapid, simplest and allele-specific marker that can be developed utilizing SNP is the Single Nucleotide Amplified Polymorphism (SNAP) marker [14,15]. SNAP marker uses modified allele-specific primers with a mismatched base pair within four bases of the 3'-end in addition to the 3'-end base complementary to the SNP site. The SNAP markers can be developed and applied to construct genetic fingerprint, analyze genetic diversity, kinship and pollen dispersal of target plants [16]. Constructing the DNA fingerprint for chili pepper varieties not only would identify chili pepper species, but also could provide their genetic distance. However, DNA fingerprinting by SNAP markers in chili pepper varieties has not been carried out. In this study, the DNA fingerprints for 23 varieties of chili pepper were constructed by using SNAP markers to provide a reliable scientific basis for the molecular identification and the intellectual property protection of the varieties.

2. Materials and methods

2.1. Plant materials and DNA extraction

A total of 23 chili pepper varieties were used as the plant genetic materials. Detailed information of the chili pepper varieties including name, subspecies, year of release and pedigree is available (Table 1). All chili pepper varieties were grown in a greenhouse until three or four leaves stage of seedling. Genomic DNA extractions from fresh young and healthy leaves were done in Laboratory of Molecular Biology, ICABIOGRAD using cetyltrimethylammonium bromide (CTAB) method [17] with some minor modifications. The quality of extracted DNA was estimated using NanoDrop^{TM} 2000 Spectrophotometer (Thermo Fisher Scientific, USA) and run in 1% (w/v) agarose gel. The samples were visualized under Geldoc-UV Imager (Thermo Fisher Scientific, USA).

2.2. SNAP primer designing

The gene-specific SNAP primers were developed based on previously identified SNP sites (www.genom.litbang.pertanian.go.id) in the genome of chili pepper. The identified SNPs having bi-allelic alternative alleles were selected, and their fragment sequences were adjusted as required for submission for SNAP primer design using the WebSNAPER program (http://ausubellab.mgh.harvard.edu). In WebSNAPER, PCR product with optimum size of 325–375 and absolute size of 300–500 were chosen, while other criteria followed WebSNAPER instruction. After the submission process, optional SNAP primers output with reference and alternate alleles could be seen on display and combination of SNAP primer pairs corresponding to the SNP appeared.

Candidates of SNAP primer pairs with high stability were selected and tested using optimum PCR reaction and program as recommended by WebSNAPER. A pair of primers specific to the corresponding allele with a single band and consistent to the SNP existed in chili pepper varieties could be used as SNAP marker (Table 2).

2.3. DNA amplification

DNA amplification was performed in a T1 Thermocycler (Biometra, Germany). The PCR was performed in 10 µl reaction solution containing 40 ng DNA template, 5 µl Kapa2G Fast Ready Mix (Kapa Biosystems, USA), 0.5 µl each of the forward and reverse primers, and 2 µl sterile $_{dd}H_2O$. PCR conditions for amplification were as follow: pre-denaturation at 94°C for 5 min, 28 or 38 cycles consisting of denaturation at 94°C for 30 s, annealing and extension at 62°C for 1 min, then final extension at 72°C for 10 min and finally stored at 4°C. PCR products for each sample were separated by using 1.5% agarose gel in 1× TAE buffer at 90 V for 90 min to estimate each allele in the SNP site.

2.4. DNA fingerprint based on SNAP and genetic diversity analysis

The molecular data collected from eight SNAP primers were converted into binary format (presence of allele as "1" and absence of allele as "0" representing the reference alleles and alternate alleles, respectively) for analysis with PowerMarker V3.25. Characteristic of the SNAP primer pairs for constructing chili pepper DNA fingerprint were evaluated in the 23 chili pepper varieties in terms of major allele frequency, Nei's gene diversity and Polymorphic Information Content (PIC) using PowerMarker V3.25 software [18]. Molecular Evolutionary Genetics Analysis (MEGA) version 5.0 software [19] was used to develop an Unweighted Pair Group Method of Arithmetic Mean (UPGMA) for evaluating genetic relationships among chili pepper varieties.

Table 1. Detailed information on chili accessions used in this study.

Variety	Species	Year of release	Pedigree
Tanjung-1	*Capsicum annuum*	2001	Natural segregant from Brebes local variety
Tanjung-2	*C. annuum*	2008	Natural segregant from Brebes local variety
Lembang-1	*C. annuum*	2001	Lines selection from Pangalengan local variety
Lingga	*C. annuum*	2011	Lines selection of LV3491
Ciko	*C. annuum*	2011	Lines selection of LV2699
Kencana	*C. annuum*	2011	Lines selection of LV6401
Gelora	*C. annuum*		
Canon	*C. frustecens*	2016	Mass selection from CR017382620115110
AVPP 0207	*C. annuum*		Introduction from AVRDC
Taringe	*C. frustecens*	2008	Mass selection from CR020.0.3.1.2.0
Kresna	*C. frustecens*	2011	
Lembang	*C. frustecens*		Lembang local variety
Landung	*C. annuum*	2011	
Sempurna	*C. annuum*		Natural segregant from Sumatra local variety
Tunduk	*C. frustecens*		
Madun	*C. frustecens*	2013	Lines selection of CR021
Midun	*C. frustecens*		
Andalas	*C. annuum*	2011	Lines selection of CK835
Rama	*C. frustecens*	2011	Lines selection of CR729
Vitra	*C. annuum*		Natural segregant from Sumatra local variety
Tripang	*C. frustecens*		
Prima Agrihorti	*C. frustecens*	2015	Lines selection of R29
Lembang	*C. annuum*		Lembang local variety

Table 2. List of designed and selected Single Nucleotide Amplified Polymorphism (SNAP) primer sets.

Primers	Chr	Primer sequences (5'–3')	Ref	Alt	Ann (°C)	Cycle	Product size (bp)
CaSNAP6_3151	6	F: TTTAATTTTCAAATATCATTGTTCACT TCGAAAACG R: TCCTTCTTAATCACGAAATCAACCCA CTTTCT	A	C	62	38	353
CaSNAP1_0181	1	F: GAAATGCTGAAATAAGTAGCAATAA GAAGCAAAATG R: TTTAAAGCCTTGAGATAAAAGCATAT GTTCTGGAAG	G	A	62	28	375
CaSNAP1_3421	1	F: TATTCAATATTAGGTGAAATGCTCTA GTTGCTCACG R: GGCATTATTCTTAATGCCATTCCACA TAACTAAAAA	C	A	62	28	363
CaSNAP1_5962	1	F: GATCAAATAATGTCATCGGACATGC TCG R: CTGATTTGCGTTTAACTTTGAGAATC CATTTGT	G	A	62	28	373
CaSNAP11_2961	11	F: GAGGCATTGGTGCCTAATCAGGGAT CCTGCTTGTCTGCCCCTCAAAATAGAA R: CCTGCTTGTCTGCCCCTCAAAATAGA A	T	C	62	28	346
CaSNAP11_1679	11	F: TCTGCTGATACCTATTTACCATACTTA TTGAAGACA R: AAAAACATACGGTTACTGATGGCGG ATAGG	A	G	62	28	332
CaSNAP9_4829	2	F: TTTATATTGCCTTACCTATCATTCCTT CACTCTAGC R: TACGCCGAATGGTTGGACTCGCTATA	C	T	62	28	340
CaSNAP9_5132	9	F: AAGTTTGAAATATAGCTTATGCATGC GGGTG R: GAAACTCACCTAAGATATACTATTGA CTCCCCCGAT	G	T	62	38	355

Chr = chromosome, Ref = reference, Alt = alternate, Ann = annealing.

3. Results and discussion

3.1. SNAP markers reliability for DNA fingerprinting construction

All eight SNAP primer pairs could amplify the target sequences in 23 chili pepper varieties (Figure 1). In this study, only reference allele was converted to SNAP, confirming the allele presence and absence depending on the SNP detected in each variety, then translated into nucleotide base scoring profile. Based on the 8 SNAP markers tested on 23 chili pepper varieties, six markers (CaSNAP6_3151, CaSNAP1_0181, CaSNAP11_2961, CaSNAP11_1679, CaSNAP9_4829 and CaSNAP9_5132) revealed polymorphism while 2 SNAP markers (CaSNAP1_3421 and CaSNAP1_5962) were monomorphic, amplifying products for both reference and alternate allele in all the 23 chili pepper varieties. These result indicated that these six markers were suitable for constructing SNAP fingerprint profiles of the 23 chili pepper varieties.

Subsequently, genetic properties for SNAP markers were calculated, including major allele frequency, gene diversity and PIC (Table 3). The average of genes diversity was 0.20 of total genotypes for total SNAP. The usefulness of molecular markers could be measured based on their PIC [20]. PIC is described as the value of a marker for detecting polymorphism in a population and it depends on the number of detectable alleles and distribution of their frequencies. PIC of 6 SNAP

markers used ranged from 0.14 (CaSNAP6_3151 and CaSNAP9_5132) to 0.36 (CaSNAP9_4829). SNAP is co-dominant marker and bi-allelic. However, their PIC is not high as multi-allele microsatellites. As suggested by Guidelines for Molecular Marker Selection and Database Construction, co-dominant markers are favoured as molecular markers for DNA fingerprinting [21].

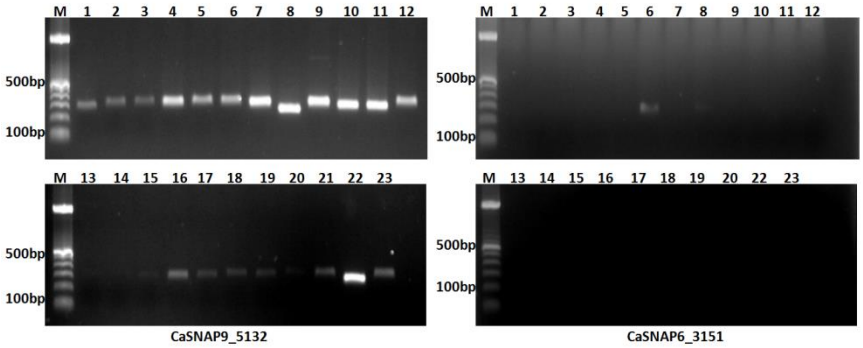


Figure 1. Bands of amplified pattern results after electrophoresis for CaSNAP9_5132 and CaSNAP6_3151, separated by 1.5% agarose gel.

Table 3. Summary of descriptive statistics of 8 SNAP markers in 23 chili pepper varieties.

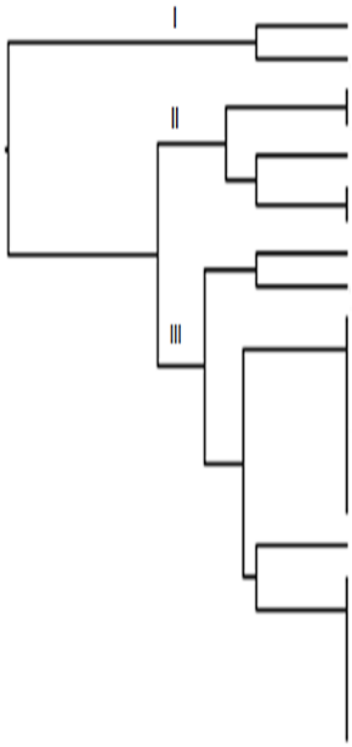
Marker	Major allele frequency	Gene diversity	PIC
CaSNAP6_3151	0.91	0.15	0.14
CaSNAP1_0181	0.86	0.22	0.20
CaSNAP1_3421	1.00	0.00	0.00
CaSNAP1_5962	1.00	0.00	0.00
CaSNAP11_2961	0.86	0.22	0.20
CaSNAP11_1679	0.69	0.42	0.33
CaSNAP9_4829	0.60	0.47	0.36
CaSNAP9_5132	0.91	0.15	0.14
Mean	0.85	0.20	0.17

3.2. DNA fingerprint of chili pepper varieties based on SNAP markers

To identify the genetic diversity between the 23 chili pepper varieties, a phylogenetic tree was subsequently constructed from the six selected SNAP markers based on their genotypic scores using the UPGMA method. UPGMA separated the 23 varieties into three major groups (Figure 2). The first major group consisted of two varieties (Landung and Sempurna), the second major group comprised five varieties (Tripang, Midun, Vitra, Ciko and Tanjung-2) and the remaining varieties (Gelora, Tunduk, Madun, Andalas, Rama, Prima Agrihorti, Lembang [*C. frustecens*], Kresna, Taringe, AVPP-0207, Lingga, Lembang-1, Tanjung-1, Lembang [*C. annuum*], Kencana and Canon) belonged to the third major group. The results of the grouping indicated that the 23 chili pepper varieties of different species could not be distinguished clearly. Some of *C. annuum* species did not separate from *C. frustecens* species.

DNA fingerprinting with molecular markers allows precise, objective and rapid varietal identification. A DNA fingerprinting of 23 chili pepper varieties was constructed with six selected SNAP markers. Based upon the amplicon profile generated by analyzing 23 varieties of chili pepper using six primer pairs, a 6-digit DNA fingerprint for six primer pairs was constructed (Figure 2). For this purpose, the assigned allele for 6 primer pairs was placed from left to right in capital letters. Digits from left to right corresponded to the allele at loci CaSNAP6_3151, CaSNAP1_0181,

CaSNAP11_2961, CaSNAP11_1679, CaSNAP9_4829 and CaSNAP9_5132. For example, fingerprint code for the Tanjung-1 variety was CGTACG, which was from left to right signified scored allele of CaSNAP6_3151, CaSNAP1_0181, CaSNAP11_2961, CaSNAP11_1679, CaSNAP9_4829 and CaSNAP9_5132, respectively.



Variety	Fingerprint Code
Landung	CACGTT
Sempurna	CATGTT
Ciko	CGCGTG
Vitra	CGCGTG
Tanjung-2	CGTGCG
Midun	CGTGTG
Tripang	CGTGTG
Kencana	AGTATG
Canon	AGTACG
Lembang (*C. frustecens*)	CGTACG
Kresna	CGTACG
Taringe	CGTACG
AVPP-0207	CGTACG
Lingga	CGTACG
Tanjung-1	CGTACG
Lembang-1	CGTACG
Lembang (*C. annuum*)	CATATG
Prima Agrihorti	CGTATG
Rama	CGTATG
Andalas	CGTATG
Madun	CGTATG
Gelora	CGTATG
Tunduk	CGTATG

Figure 2. Phylogenetic tree and DNA fingerprint code of 23 chili pepper varieties resulted from UPGMA cluster analysis based on SNAP marker.

The DNA fingerprint code of genotypes would reflect how closely the varieties are related to each other. The six selected SNAP primers were not able to distinguish some of chili pepper varieties (Figure 2.). For instance, seven varieties (Lembang [*C. frustecens*], Kresna, Taringe, AVPP-0207, Lingga, Tanjung-1 and Lembang-1) have the same DNA fingerprint code (CGTACG), and the other six varieties (Prima Agrihorti, Rama, Andalas, Madun, Gelora and Tunduk) also have the same code (CGTATG). Further work is needed to develop the DNA fingerprinting of these chili pepper varieties using SNAP markers.

4. Conclusions

A preliminary DNA fingerprinting database of the 23 chili pepper varieties was built in this study using six SNAP markers, which could be expanded as the number of additional varieties and molecular markers increase. Phylogenetic tree of SNAP markers divided the 23 varieties into three major groups. However, a number of varities possessed the same DNA fingerprint profiles, indicating their close genetic distance. Eventhough these SNAP markers were not able to identify each variety according to its unique code, this study could be useful as preliminary information to establish the genetic identity of chili peppers variety in Indonesia in the future.

5. Acknowledgement
The authors gratefully acknowledge financial support from Indonesian Agency for Agricultural Research and Development through Indonesian National Budget fiscal year 2017 under the project hosted by the Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development.

6. References

[1] Eggink P M 2013 *A Taste of Pepper: Genetics, Biochemistry and Prediction of Sweet Pepper Flavor* (Wageningen University)
[2] Datta S and Jana J C 2010 Genetic variability, heritability and correlation in chilli genotypes under Terai zone of West Bengal *SAARC J. Agric.* **8** 33–45
[3] Sharma V K, Semwal C S and Uniyal S P 2010 Genetic variability and character association analysis in bell pepper (*Capsicum annuum* L.) *J. Hortic. For.* **2** 58–65
[4] Agustina S, Widodo P and Hidayah H S 2014 Analisis fenetik kultivar cabai besar *Scr. Biol.* **1** 117–25
[5] Zhang C C, Wang L Y, Wei K and Cheng H 2014 Development and characterization of single nucleotide polymorphism markers in *Camellia sinensis* (Theaceae) *Genet. Mol. Res.* **13** 5822–31
[6] Gao L, Jia J and Kong X 2016 A SNP-based molecular barcode for characterization of common wheat *PLoS Genet.* 1–12
[7] Xu Y and Crouch J H 2008 Marker-assisted selection in plant breeding: from publications to practice *Crop Sci.* 391–407
[8] Moulin M M, Rodrigues R, Bento C S, Viana A P and Londrina U E De 2015 Construction of an integrated genetic map for *Capsicum baccatum* L. *Genet. Mol. Res.* **14** 6683–94
[9] Tan S, Cheng J, Zhang L, Qin C, Nong D and Li W 2015 Construction of an interspecific genetic map based on indel and SSR for mapping the QTLs affecting the initiation of flower primordia in pepper (*Capsicum* spp.) *PLoS One* 1–15
[10] Mimura Y, Inoue T, Minamiyama Y and Kubo N 2012 An SSR-based genetic map of pepper (*Capsicum annuum* L.) serves as an anchor for the alignment of major pepper maps *Breed. Sci.* **98** 93–8
[11] Cheng J, Zhao Z, Li B, Qin C, Wu Z, Trejo-Saavedra D L, Luo X, Cui J, Rivera-Bustamante R F, Li S and Hu K 2016 A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum Sci. Rep.* **6** 1–12
[12] Dias G B, Gomes V M, Moraes T M S, Zottich U P and Rabelo G R 2013 Characterization of *Capsicum* species using anatomical and molecular data *Genet. Mol. Res.* **12** 6488–501
[13] Alkan C, Coe B P and Eichler E E 2014 Genome structural variation discovery and genotyping *Nat. Rev. Gen.* **12** 363–76
[14] Lestari P and Koh H E E J 2013 Development of new CAPS/dCAPS and SNAP markers for rice eating quality *Hayati J. Biosci.* **20** 15–23
[15] Ruangchai J, Prakit S, Sompong C, Takehiko S, Sugunya W, Akito K and Peerasak S 2011 A SNP in *GmBADH2* gene associates with fragrance in vegetable soybean variety “Kaori” and SNAP marker development for the fragrance *Theor. Appl. Genet.* **122** 533–41
[16] Drenkard E, Richter B G, Rozen S, Stutius L M, Angell N A, Mindrinos M, Cho R J, Oefner P J, Davis R W and Ausubel F M 2000 A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in *Arabidopsis Plant Physiol.* **124** 1483–92
[17] Doyle J and Doyle J 1990 Isolation of plant DNA from fresh tissue *Focus (Madison)* **12** 13–5
[18] Liu K and Muse S V 2005 PowerMarker: an integrated analysis environment for genetic marker analysis *Bioinformatics* **21** 2128–9
[19] Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S 2011 MEGA5: molecular

evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods research resource *Mol. Biol. Evol.* **28** 2731–9

[20] Botstein D, White R L, Skolnick M and Davis R W 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32** 314–31

[21] Jordens R 2005 Progress of plant variety protection based on the International Convention for the Protection of New Varieties of Plants (UPOV Convention) *World Pat. Inf.* **27** 232–43