

Short Communication

Molecular markers for *Capsicum frutescens* varieties cultivated in Borneo

K. F. Rodrigues and H. K. Tam

Biotechnology Research Institute, University of Malaysia Sabah, Locked Bag 2073, Kota Kinabalu 88999, Sabah, Malaysia.

Accepted 21 October, 2017

This paper describes the isolation and characterization of microsatellite loci in *Capsicum frutescens* varieties cultivated in Borneo, Malaysia. A total of 58 putative microsatellites distributed across 29 sequences were isolated from a genomic library enriched for microsatellites. Ten locus specific polymorphic molecular markers were developed and applied to determine the levels of genetic diversity across the cultivated population. The expected heterozygosity ranged from 0.4316 - 0.6684 (mean 0.5958) and the observed heterozygosity ranged from 0 - 1 (mean 0.6000). There was a deficiency of heterozygotes within the total population which implies that most of the cultivated varieties in the region have been derived as a result of inbreeding. The estimation of the total genetic diversity of a cultivated crop variety within a geographic region can provide an indication of the extent to which traditional cultivators are relying on inbred varieties of seeds for crop production.

Key words: *Capsicum frutescens*, bird's eye chili, inosine, microsatellites, heterozygosity.

INTRODUCTION

The bird's eye chili (*Capsicum frutescens*) is cultivated across South East Asia for its culinary and medicinal applications. The aim of this investigation was the estimation of the levels of genetic diversity within cultivated varieties of *C. frutescens* among traditional cultivators in Sabah, North Borneo, Malaysia. The farming community in this region is comprised of small holders who have relied on the traditional method of seed production, which does not involve hybridization and selection of determinate hybrids.

Therefore, it was hypothesized that the level of genetic diversity within the region would provide an indicator of the level of exchange of germplasm between individual cultivators as well as the extent of out-crossing. Previous attempts to characterize the molecular genetic diversity in the genus *Capsicum* have relied on AFLPs (Toquica et al., 2003), microsatellites (Lee et al., 2004) and combinations of AFLPs and RFLPs (Kang et al., 2001), however an overview of the NCBI GenBank indicated that there were no specific sequences containing microsatellites available for *C. frutescens*, thus prompting the need to

isolate novel genomic loci containing microsatellite repeat motifs and develop specific molecular markers for application in population genetic studies. A total of forty leaf samples of *C. frutescens* were obtained from cultivators across the state of Sabah, Malaysia.

The anthers of *C. frutescens* exhibit a distinct purple color as compared to the anthers of *C. annuum* that are white. DNA was extracted according to the method described by Doyle and Doyle (1987). DNA concentration and purity were estimated using a NanoVue Spectrophotometer (GE Healthcare Lifesciences). A genomic library enriched for microsatellite loci was constructed by PCR amplification of DNA from one specimen (CFcp21) with degenerate primers incorporating inosine. Seven degenerate primers with a range of repeat motifs, BRIC10 [5'- GGIII(CT)₁₀I - 3'], BRIAGG6 [5'- GGIII(AGG)₆I - 3'], BRIGAA6 [5'- GGIII(GAA)₆I - 3'], BRIAG10 [5'- GGIII(AG)₁₀I - 3'], BRIGTT6 [5'- GGIII(GTT)₆I - 3'], BRIGAT6 [5'- GGIII(GAT)₆I - 3'] and BRICTT6 [5'-GGIII (CTT)₆I - 3'] (where I = inosine) were used in PCR amplification to isolate a range of repeat motifs. PCR amplification was carried out in a total volume of 20 µl containing 50 ng of template DNA, 2.0 mM MgCl₂, 1X PCR buffer containing 10 mM Tris-HCl (pH = 8.0), 50 mM KCl, 1 U Taq DNA polymerase

*Corresponding author. E-mail: kennethr@ums.edu.my.

(Qiagen), 20 pmol of degenerate primer (BRIC10, BRIAGG6, BRIGAA6, BRIAG10, BRIGTT6, BRIGAT6, BRICTT6), 0.25 mM of dATP, dGTP, dCTP and dTTP. Amplification was carried out in a thermocycler (MJ Research thermal cycler) with an initial denaturation at 96°C for 5 min followed by 30 cycles of 30 s at 94°C, 40 s at 58°C, 90 s at 72°C and a final extension step of 10 min at 72°C. The amplification products were separated by electrophoresis on a 2% TBE Agarose gel with 100 bp and 1 Kb DNA ladders (Promega) as size standards. The gel was stained with Ethidium Bromide (5 µg/ml) and visualized using an ALPHAIMAGER 2000 (Alpha Innotech Corp. USA). PCR products containing distinct amplicons with a size in excess of 300 bp were purified using a PCR purification kit (Qiagen) and 2 µL of the purified PCR products were blunted with 1.0 U T4 DNA Polymerase (Fermentas) ligated onto a pJET1.2 blunt cloning vector and transformed into chemically competent *Escherichia coli* TOP 10 cells according to the instructions from the manufacturer (Fermentas). A total of 50 recombinant clones derived from each of the two species were selected randomly from Luria-Bertani plates containing Ampicillin (100 mg/l). Plasmids were extracted and purified using the GeneJET plasmid miniprep kit according to the instructions from the manufacturer (Fermentas). Plasmids containing inserts were identified by PCR amplification of each of the purified plasmids. PCR amplification was carried out in a total volume of 20 µl containing 50 ng of template plasmid DNA, 1.5 mM MgCl₂, 1X PCR buffer containing 10 mM Tris-HCl (pH = 8.0), 50 mM KCl, 0.5 U Taq DNA polymerase (Qiagen), 10 pmol of specific primer pJET1.2F (5'-CGACTCACTATAGGGAGAGCGGC-3') and pJET1.2R (5'-AAGAACATCGATTTTCCATGGCAG-3'), 0.2 mM of dATP, dGTP, dCTP and dTTP. Amplification was carried out in a thermocycler (MJ Research thermal cycler) with an initial denaturation at 96°C for 5 minutes followed by 30 cycles of 30 s at 94°C, 30 s at 58°C, 60 s at 72°C and a final extension step of 10 min at 72°C. PCR products were electrophoretically resolved on a 2% TBE Agarose gel.

Plasmids with an insert size in excess of 300 bp were sequenced with the primers pJET1.2F and pJET1.2R using the BigDye Terminator v3.0 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) on an ABI Prism 377 DNA sequencer. All sequences containing repeat motifs were deposited at the GenBank (NCBI). A similar search was done using the NCBI blastn algorithm. Specific primer pairs were designed to flank repeat motifs using the online software PRIMER 3 (Rozen and Skaletsky, 2000). Twenty five pairs were synthesized and applied to characterize the population. PCR amplification of specific loci was carried out in a total volume of 20 µl containing 50 ng of template DNA, 1.5 mM MgCl₂, 1X PCR buffer containing 10 mM Tris-HCl (pH = 8.0), 50 mM KCl, 0.5 U Proofreading Taq DNA polymerase (Qiagen), 10 pmol of specific primer, 0.2 mM of dATP, dGTP, dCTP

and dTTP. Amplification was carried out in a thermocycler (MJ Research thermal cycler) with an initial denaturation at 96°C for 5 min followed by 30 cycles of 30 s at 94°C, 30 s at 58°C, 30 s at 72°C and a final extension step of 5 min at 72°C. The amplification products were separated by electrophoresis on a 10% Poly Acrylamide Gel with a 25 bp DNA ladder (Promega) as a size standard. The gel was stained with Ethidium Bromide (5 µg/ml), bands within a range of 50 bp above and below the expected size were scored using the gel documentation system ALPHAIMAGER 2000 and the data recorded was analyzed using the POPGENE version 1.31 software (Yeh and Boyle 1999).

The isolation of microsatellites using degenerate primers containing inosine proved to be reliable and effective in insulating microsatellite loci in *C. frutescens*. The incorporation of an inosine residue at the 3' end of the primer facilitated extension in the event of a single nucleotide mismatch at the terminal end of the microsatellite repeat motif, a similar approach has been applied previously (Fujiwara et al., 1995). A total of 29 clones containing microsatellite repeat motifs yielded 34 dinucleotide, 24 trinucleotide and 18 cryptic simple repeat sequences. There was no significant similarity between any of the deposited sequences with those currently available at the NCBI GenBank. Twenty five microsatellite loci were randomly selected to characterize the population, of which 10 were found to be poly-morphic; the remaining 15 loci were monomorphic. The number of alleles ranged from three to eight. The expected heterozygosity ranged from 0.4316 - 0.6684, with a mean of 0.5958 and the observed heterozygosity ranged from 0 - 1, with a mean of 0.6000. The details are shown in Table 1. The Ewens-Watterson test for neutrality indicated that no loci deviated significantly from neutrality. There was no significant linkage between any of the two loci tested implying that the loci have segregated independently of each other. None of the loci deviated significantly from Hardy-Weinberg Equilibrium (HWE) when computed using the software POPGENE. The average F_{IS} was 0.7533 indicating that there is a deficiency of heterozygotes in the overall population. In a similar studies, involving crop genetic diversity of *Hordeum vulgare* in Eritrea (Backes et al., 2009) and *Sesamum indicum* L. in South East Asia (Pham et al., 2009) a high level of genetic diversity was found in crops which have been cultivated on a small-scale over a large geographic range. However in the case of *C. frutescens* varieties cultivated in Borneo, Malaysia, a deficiency of heterozygotes was observed, implying that there is almost no evidence of out-crossing. One of the possible reasons for this phenomenon may be the isolation of the *C. frutescens* cultivars within the island ecosystem of Borneo. These cultivars have possibly been introduced into Borneo by immigrants and then established in small farming communities as is supported by the evidence of heterozygote deficiency. There is a high likelihood that

the loss in heterozygosity may lead to a decrease in the level of fitness as reported in several crops species (Udall and Wendel, 2006), and this can result in a subsequent decline in crop productivity and disease resistance.

The research strategy applied in this investigation can be applied to determine the total genetic diversity of a cultivated crop within a region, and facilitate interventions, such as the introduction of new varieties for the augmentation of genetic diversity levels. The molecular markers developed have the potential for application in characterization of *C. frutescens* populations.

REFERENCES

- Backes G, Orabi J, Wolday A, Yahyaoui A, Jahoor A (2009). High genetic diversity revealed in barley (*Hordeum vulgare*) collected from small-scale farmer's fields in Eritrea. *Genet. Res. Crop. Evol.*, 56: 85-97.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem. Bull.*, 19: 11-15.
- Fujiwara H, Fujiwara K, Hashimoto K (1995). PCR with Deoxyinosine containing primers using DNA polymerases with proofreading activity. *PCR Methods Applications*, 4: 239- 240.
- Kang BC, Nahm SH, Huh JH, Yoo HS, Yu JW, Lee MH, Kim BD (2001). An interspecific (*Capsicum annuum* × *C. chinense*) F2 linkage map in pepper using RFLP and AFLP markers. *Theor. Appl. Genet.*, 102: 531-539.
- Lee JM, Nahm SH, Kim YM, Kim BD (2004). Characterization and molecular genetic mapping of microsatellite loci in pepper. *Theor. Appl. Genet.*, 108: 619- 627.
- Pham TD, Bui TM, Werlemark G, Bui TC, Merker A, Carlsson AS (2009). A study of genetic diversity of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia estimated by RAPD markers. *Genet. Resour. Crop. Evol.*, 56: 679-690.
- Rozen S, Skaletsky HJ (2000). PRIMER3 on the www for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), Humana Press, Totowa, New Jersey. pp. 365-386.
- Toquica SP, Rodriguez F, Martinez E, Duque MC, Tohme J (2003). Molecular characterization by AFLPs of *Capsicum* germplasm from the Amazon Department in Columbia, characterization by AFLPs of *Capsicum*. *Genet. Res. Crop Evol.*, 50: 639-647.
- Udall JA, Wendel JF (2006). Polyploidy and crop improvement. *The Plant Genome*, 1: 3-14.
- Yeh FC, Boyle T, Rongcai Y, Ye Z, Xian JM (1999). POPGENE version 3.1.