Development and application of microsatellites in plant breeding

Maria Celeste Gonçalves-Vidigal1* and Luciana Benchimol Rubiano2

ABSTRACT - Molecular markers are powerful tools for analyzing genome diversity within a species, and to evaluate genetic relationships between individuals and populations. Among them, microsatellites (SSRs) are one of the most important polymorphic markers that can be used effectively to distinguish germplasm accessions. These markers present high informative content due to their codominant inheritance, multiallelism, mendelian pattern and good genome coverage. The enrichment methodology for microsatellite development has a superior efficiency in plants, especially when performed using biotin-labeled microsatellite oligoprobes and streptavidin-coated magnetic beads. The development of EST-SSR markers has become a fast and relatively inexpensive way but it is limited to species for which this type of database exists. Given the high polymorphism level of microsatellites when compared to other markers, SSRs have been used to study population structure, for genetic diversity analysis, genetic mapping and marker assisted selection.

Key words: common bean, genetic diversity, microsatellites, plant breeding.

INTRODUCTION

Hybridization method is very commonly used in crop breeding in order to obtain segregant populations portraying interesting agronomical traits and broad genetic variability. This method combines available genes from two or more different genitors in a unique individual (Allard 1999).

The choice of genitors is very important in this process because it assures the obtaining of promising segregant populations. This procedures is based on genetic dissimilarity of the subject under study. In plants, genetic diversity may be inferred through quantitative (phenotypic or genotypic traits) and predictive characteristics. The quantitative genetic diversity is usually calculated from a set of genotypic or phenotypic characters, throught the use multivariate statistics (Das and Gupta 1984). Predictive methods are those based on morphological, physiological or molecular differences, quantifying then in some similarity/dissimilarity measurement, which may express the degree of parental genetic diversity (Cruz et al. 2004).

However, due to high environmental influence on morphoagronomic characteristics, many researchers have chosen predictive techniques for estimating genetic diversity as they reduce the necessity of hybridization among progenitors (Ceolin et al. 2007), diminishing labor and expenses with a great number of crosses and field experiments.

Molecular markers are powerful tools for analyzing genome diversity within a species, and to evaluate genetic relationships between individuals and populations. They have been used for mapping of genomic regions containing

---

1 Universidade Estadual de Maringá, Departamento de Agronomia, Av. Colombo, 5790, 87.020-570, Maringá, PR, Brazil. * E-mail: mcgvidigal@uem.br
2 Instituto Agronômico (IAC), Centro de Análises e Pesquisa Tecnológica do Agronegócio dos Grãos e Fibras, C. P. 28, 13.001-970, Campinas, SP, Brazil
genes of agricultural interest (Charcosset and Moreau 2004). In recent years, genetic divergence studies in different crops have often used microsatellite as the main tool for genetic diversity analysis.

Microsatellites or SSRs (Simple Sequence Repeats, Tautz 1989) are one of the most indicated techniques to study polymorphism between DNA sequences. These molecular markers are based on PCR reaction that detects loci variations of repetitive sequences. They present high levels of polymorphism, codominant inheritance, multiallelism, mendelian pattern and good genome coverage. Microsatellites require low amount of DNA, can be easily automated for high throughput screening, may be exchanged between laboratories, and are highly transferable between populations (Gupta et al. 2003).

SSRs occur frequently within or near genes in plants (Morgante et al. 2002). Comparing coding and non-coding regions in different plant species, it was observed that tri and tetra-nucleotide microsatellite motives are more common within introns, whereas other types of motives are found within exons (Tóth et al. 2000). Morgante et al. (2002) reported a selective pressure acting over untranslated 3’ and 5’ regions (UTR) of genes, favoring a higher pressure of SSRs in these regions than in the rest of the genome. The 5’-UTR presents a higher CT and CTT number of repetitions, while the 3’ UTR showed to be rich in AG and AAG. Regions of DNA involved in gene regulation are expected to exhibit sequence conservation between related species over evolutionary time due to functional constraints. Microsatellite transferability amongst related species is allowed by the homologous nature of the DNA sequence in microsatellite flanking regions (Varshney et al. 2005).

**Microsatellite enriched libraries development**

The development of microsatellite markers involves several distinct steps from obtaining the library to developing a working set of primers that can amplify polymorphic microsatellite loci.

Traditional methods for developing microsatellites have a low efficiency and can be time-consuming, especially in the selection stage and clone screening (Zane et al. 2002). Ostrander et al. (1992) proposed the enrichment methodology for microsatellites achieving enrichment 50 times superior to genomic libraries. The selective hybridization protocols are very commonly used in microsatellite enrichment (Kijas et al. 1994, Billotte et al. 1999).

**Example 1**

1.1. Construction of an enriched microsatellite library for common beans

An enriched library was constructed according to Billotte et al. (1999) for IAC-UNA according to Benchimol et al. (2007). The first step was to quantify the genomic DNA (200 ng mL) extracted using CTAB method (Hoisington et al. 1994) on agarose gel. The selection of microsatellite sequences was performed using biotin-labeled microsatellite oligoprobe and streptavidin-coated magnetic beads following a hybridization-based capture methodology adapted from Kijas et al. (1994).

Plasmids form positive clones containing microsatellites were extracted by plasmid DNA minipreparation. All clones were sequenced in both directions. Reads were processed by the Phred version 0.000925.c base calling program (Ewing et al. 1998) and vector sequences, poly-A tail, and adapters were trimmed after cross-match analysis. Only perfect and/or imperfect sequences containing to five or more repeated units were considered. Complementary primers to single sequences, flanking the microsatellites, were designed. Stringency criteria adopted was based on GC content between 40 and 60 %, melting temperature between 46 and 60 °C, a salt concentration of 50 mM, and a product length between 150 and 300 pb.

1.2. Genic microsatellites from EST database

The major constraint of using SSR markers from genomic libraries is the high development cost and effort required to obtain working primers for a given study species. With the establishment of expressed sequence tag (EST) sequence projects for many plant species, a great amount of information and DNA sequences were generated and deposited in online databases. The development of EST-based SSR markers through data mining has become a fast, efficient, and relatively inexpensive compared with the development of genomic SSRs (Gupta et al. 2003). Genbank (www.ncbi.nlm.nih.gov/ genbank/), for instance, gathers a great amount of EST sequences, gene and cDNA clones. The TIGR Institute (The Institute for Genomic Research – <http://www.tigr.org/ tdb/e2k1/plant.repeats/> ) possesses an EST database for many cultivated plant species of economic importance and therefore, it is hard to find sequences of wild species.

Many programs that have been developed for recognition of SSR patterns in the sequence files, such as
the ‘Tandem Repeats Finder’ - TRF (Benson 1999), the ‘Simple Sequence Repeat Identification Tool’ - SSRIT (Temnykh et al. 2001, <http://www.gramene.org/db/searches/ssritool>), and program MISA (Micro Satellite) identification tool (<http://pgrc.ipk-gatersleben.de/misa/>). A major disadvantage of EST-derived microsatellites is the sequence redundancy that yields multiple set of markers at the same locus. Therefore it is very important to verify the existence of redundancy among sequences and BLASTN tool (NCBI; <http://blast.ncbi.nlm.nih.gov/blast.cgi>) could be used for that purpose.

**Application of microsatellites in plant breeding**

Microsatellites have been used for plant genetic analysis such as to measure the effects of natural selection (Rodrigues and Santos 2006), to unveil the genetic diversity (Vieira et al. 2009, Guan et al. 2010), to measure population structure (Ribeiro et al. 2010, Albertini et al. 2011), to integrate the genetic, physical and sequence-based physical maps (McCleaf et al. 2010, Garcia et al. 2011) and for marker assisted selection (Benchimol et al. 2005, Chen et al. 2011).

**Using microsatellite in genetic diversity and population structure in common bean**

Molecular marker analysis has contributed for the understanding of common bean genetic diversity and phylogenetic analysis (Asfaw et al. 2009). For molecular markers, such as RAPD, SSR, and RFLP, there were major findings such as plant population continuous variation and high level of dissimilarity within population and among germplasm collections. Microsatellite markers are involved not only in genetic diversity studies, population genetics and evolutionary studies, but are also being used in fundamental research like genome analysis, gene mapping and marker-assisted selection (Kalia et al. 2011).

Given the high polymorphism level of microsatellites when compared to other markers, many worldwide researchers have preferred using these markers on the study of common bean population structure, genetic divergence and genetic mapping (Diaz and Blair 2006, Hanai et al. 2007, Blair et al. 2009, Campos et al. 2011).

**Example 2**

**Genetic divergence and population structure in common bean using microsatellite**

The assessment of genetic diversity and population structure of 122 accessions of common bean from Brazil and Africa was performed using microsatellites, genetic distance estimates and probabilistic models to analyze population structure.

Sixteen microsatellite markers from all the 11 linkage groups of *P. vulgaris* L. gene map were selected for analyses based on their dispersed map location. Markers originated from genic and genomic sequences were chosen in equal proportions (Yu et al. 2000, Gaitán-Solís et al. 2002, Blair et al. 2003, Grisán et al. 2007). Microsatellite analysis was conducted as described by Kwak and Gepts (2009), including an economic method of microsatellite fluorescent labeling of fragments amplified by PCR (Schuelke 2000).

Microsatellite methods employing fluorescent labeling and automated band calling with precise software-based allele detection are considered the most accurate way of genotyping (McCouch et al. 1997, Blair et al. 2002). Additionally, this sort of labeling organizes markers into distinct dye color panels allowing multiplexing during band separation, with advantages of high throughput genotyping and simultaneous analysis of multiple loci (Coburn et al. 2002, Oblessuc et al. 2009).

The methodology of multiplexing can also be carried out in PCR amplification step with mixtures of the appropriate primers (Hayden et al. 2008). Current technology allows capillary-based separation of microsatellite bands with four different color panels, one for each individual marker, which can be evaluated through a single capillary with a separate size standard, allowing precise band size estimates (Coburn et al. 2002).

The amplified fragments were multiplexed depending on their size variation and analyzed in an ABI 3730 (Applied Biosystems). Genotypes of markers were determined using the GeneMarker program (version1.51, SoftGenetics). The genetic relationship among entire accessions was analyzed by principal coordinate analysis performed using the GenAlex 6 Program (Peakall and Smouse 2006).

The genetic distance among 122 Andean and Mesoamerican accessions was calculated using the C.S. Chord distance (Cavalli-Sforza and Edwards 1967). The C.S. Chord distance does not require any mutation model for microsatellite evolution and is free from bottleneck effect (Takezaki and Nei 1996). Based on this genetic distance, an unrooted neighbor-joining (NJ) tree was constructed in PowerMarker (Liu and Muse 2005). Distinct clusters were apparent with the SSR markers unambiguously assigning accessions to the Andean and Mesoamerican gene pools with the 2D plot of the PCoA based on pairwise genetic distances.
The obtained results demonstrated that Andean gene pool accessions from Brazil and Africa were clustered together showing a great genetic similarity. Likewise, the Mesoamerican accessions from both countries formed a second group. These results indicate the possibility of introgression between countries. Some African accessions with low genetic similarity could be use in breeding programs to broaden the genetic base of Brazilian cultivars.

Example 3

Potential application of SSR markers for mapping and tagging disease resistance genes common bean

Anthracnose (ANT) caused by Colletotrichum lindemuthianum (Sacc.) Scrib and angular leaf spot (ALS) caused by Pseudocercospora griseola (Sacc.) Crous e U. Braun, are the most widespread, recurrent and devastating diseases of common bean in Latin America and Africa. Use of disease resistance genes is the most practical, cost-effective and environmentally friendly strategy for the control of ANT and ALS. Resistance to various diseases in common bean is conferred mostly by single, dominant genes with race-specific resistance (R-genes). Resistance to C. lindemuthianum is conditioned by some 13 reported genes identified by the Co symbol (Kelly and Vallejo 2004). Six independent dominant genes identified by the Phg symbol condition resistance to P. griseola (Caixeta et al. 2005).

Recent studies were conducted at the Laboratório de Biotecnologia do Núcleo de Pesquisa Aplicada a Agricultura (NUPAGRI) of the Universidade Estadual de Maringá (Paraná, Brazil) to study the co-segregation between ANT and ALS resistance genes that confer resistance to C. lindemuthianum and P. griseola, respectively. A total of 112 individual F2 derived from a cross between resistant and susceptible parents were inoculated with races 73 of C. lindemuthianum and 63-39 of P. griseola. The analysis was performed using 22 SSR markers. Two contrasting DNA bulks (Michelmore et al. 1991) were constructed by pooling equal volumes of fluorometrically standardized DNA from four to six resistant and susceptible individuals, respectively, of the F2 population. Of these 22 SSR markers, the g2303 showed contrasting amplification patterns in parental materials and in resistant vs. susceptible bulks or individuals. Segregation analyses of the disease reaction from F2 plants were performed by the chi-square (χ²) test, according to a Mendelian segregation hypothesis of 3 R_ (resistant) to 1 rr (susceptible). Linkage analyses were performed using the computer software Mapmaker/exp 3.0 (Lincoln and Lander 1993). The results of molecular mapping revealed that the SSR marker g2303 was linked at 0.9 cM from ANT resistance gene and 1.8 cM from ALS resistance gene in coupling phase with both genes, and mapped at linkage group Pv04 in the common bean linkage map. The co-segregation between these genes revealed the possibility of monitoring those genes indirectly through the use of g2303 molecular marker. Similar results were obtained by Gonçalves-Vidigal et al. 2011). In this study, the findings might help plant breeding programs to reduce time and cost associated with pyramiding resistance genes to different pathogens.

ACKNOWLEDGEMENTS

We would like to thank to SBMP, CNPq, Capes, FAPESP for the financial support. MC Gonçalves-Vidigal is sponsored by grants from CNPq.
Desenvolvimento e aplicação dos microsatélites no melhoramento de plantas

RESUMO - Marcadores moleculares são ferramentas auxiliares nas análises da diversidade genômica dentro de uma espécie e para avaliar as relashões entre indivíduos e populações. Dentre eles, os microsatélites (SSRs) é uma das classes de marcadores mais importante que pode ser usada com eficiência para distinguir acessos provenientes de um banco de germoplasma. Estes marcadores apresentam elevado grau de informação devido a sua herança co-dominante, multialelismo, padrão mendeliano e boa cobertura do genoma. A metodologia de enriquecimento para o desenvolvimento de microsatélites apresenta uma eficiência superior em plantas, especialmente quando se utiliza sondas marcadas com biotina e esferas magnéticas revestidas com estreptavidina.

O desenvolvimento de EST-SSRs tem se tornado uma forma rápida para obtenção de marcadores moleculares, porém limitada a espécies para as quais existam bancos de sequências. Dado o elevado nível de polimorfismo dos microsatélites quando comparados com outros marcadores, estes têm sido utilizados em estudos da estrutura genética de populações, em análises de diversidade genética, em mapeamento genético e em seleção assistida por marcadores.

Palavras-chave: diversidade genética, feijoeiro comum, melhoramento de plantas, microsatélites.

REFERENCES


Development and application of microsatellites in plant breeding


MC Gonçalves-Vidigal and LB Rubiano


