Molecular Breeding **11:** 303–313, 2003.

Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.)

P. Masi[1,*], P.L. Spagnoletti Zeuli[1] and P. Donini[2]
[1]*Dept. Biologia Difesa e Biotecnologie Agro-forestali, Università degli Studi della Basilicata, 85100, Italy;*
[2]*NIAB (National Institute of Agricultural Botany), Molecular Research Group, Cambridge, CB3 0LE, UK;*
**Author for correspondence (e-mail: bio346@unibas.it; phone: +39.971.205533; fax: +39.971.205503)*

Received 8 March 2002; accepted in revised form 8 January 2003

Key words: Common bean, Fluorescent-based DNA detection, Multiplex-PCR, Simple sequence repeats

Abstract

Multiplexing involves the simultaneous amplification of several loci in a single PCR reaction, and subsequent analysis of multiple molecular markers in a single gel lane. This study focuses on the usefulness of SSR (simple sequence repeat) multiplex-PCR, and how this method can be both highly informative and amenable to automation using a fluorescence-based, semi-automated DNA sequencing platform. We also discuss the relevance of this approach to the issue of the multiplex ratio of SSR markers and show that it is effective as an economic procedure for germplasm evaluation and evolutionary studies of plants. Various conditions of multiplex-PCR were examined and a method is proposed for developing multiplex sets of SSR markers in common bean (*Phaseolus vulgaris* L.). Seven multiplex sets, based on 30 SSR markers selected from GenBank sequences, were developed. The method was tested using a large number of samples from three common bean landraces. Technical aspects of the application and a description of the allelic constitution at various loci of all genotypes analysed are described. From a sample of 264 genotypes, selected from three landraces, we detected a total of 135 alleles, equivalent to 4.3 alleles per SSR. Null alleles were observed in each of the three landraces analysed. The procedures used in this study are applicable for the study of genetic diversity in common bean germplasm collections consisting of a significant number of accessions, and should be transferable to similar analyses of any species.

Introduction

The genetic diversity of common bean (*Phaseolus vulgaris* L.), and its related and wild species, has been widely characterised using a range of molecular markers (Becerra and Gepts 1994; Tohme et al. 1996; Caicedo et al. 1999; Beebe et al. 2000). Current analyses of plant molecular diversity strongly emphasise the relevance and use of SSRs, but to date, their presence, abundance and variation in the bean genome have been poorly characterised. This reflects firstly the lack of significant amounts of publicly accessible genomic sequence as compared to what is available for other major crop species (only 341 *Phaseolus* and *Vigna* sequences were publically available in 2002 (Yu et al. 2000; Metais et al. 2002)); and secondly, the relatively minor status of *Phaseolus* spp. among commercial crops hinders any *de novo* development of SSR assays. Some effort has, however, been made to develop and use SSR assays to improve our understanding of bean genetics and evolution (Schlotterer and Tautz 1992; Yu et al. 1999, 2000).

The development and application of SSRs facilitates the acquisition of a large quantity of genetic information of relevance to genotype identification, and therefore provides an opportunity to characterise germplasm collections (Mitchell et al. 1997). SSR loci have numerous advantages for molecular genetic analyses. Specifically these are: a high discriminatory power, a high information content arising from their multi-allelic nature, co-dominant transmission, a robust and reproducible assay, their relative abundance with uniform genome coverage, the small amount of DNA template necessary, and their easy detection via

automated systems (Powell et al. 1996; Rafalsky and Tingey 1993). However, there are also some disadvantages, mainly that each individual SSR assay generates only one datum point per genotype.

Other technical problems relate to assay cost with respect to consumables and operator time, and to scoring problems when alleles differ only marginally in length from one another. Competing marker systems such as RAPDs and AFLPs, or those relying on the detection of repetitive sequences, allow for the simultaneous detection of multiple loci, resulting in a high multiplex ratio; but these generally are inherited as dominant markers and thus deliver a lower information content in the context of a genetic diversity index. To increase the information per SSR assay, multiplex-PCR approaches have been developed and combined with fluorescence-based DNA detection and semi-automated fragment analysis. Multiplexing increases the information content of each PCR reaction and thereby lowers the cost barriers to large scale use of SSRs as genetic markers.

In this paper we describe the development of an SSR-PCR approach, which combines multiplexing (the simultaneous amplification of several loci in a single PCR reaction) with multiloading (the pooling of PCR products from several individual reactions into a single gel track). The amplification of more than one SSR locus per reaction requires the optimisation of PCR conditions in order to satisfactorily amplify all the target amplicons. The development of an SSR multiplex also requires *a priori* knowledge of the allele size range of each amplified locus, so that primers can be rationally grouped into sets. We have tested the consistency of results over three large, genetically heterogeneous common bean populations, and we describe their allelic constitution at the detected loci. In summary, the objectives of this study were: 1) to develop multiplex SSR-PCR in common bean, 2) to evaluate the potential of multiplex and multiloading PCR and fluorescent labelling and semi-automatic detection systems for mass genotyping of common bean germplasm, and 3) to evaluate the level of polymorphism present in three common bean landraces.

Materials and methods

Plant materials

Three common bean landraces, with local names ″Poverella″, ″Marrozzo″ and ″Verdolino″, were collected in the Basilicata region of Southern Italy, and stored at the Centro Interdipartimentale per la Salvaguardia delle Risorse Vegetali ″P. Iannelli″, University of Basilicata. These landraces, still grown by traditional farmers for family use as dry seeds (Masi et al. 1999), are morphologically distinct and, based on seed morphology and phaseolin electrophoretic seed patterns (Masi 2001; Masi et al. 1999), were identified as originating from the Andean diversification centre. In all, 64 individuals of ″Poverella″, 86 of ″Marrozzo″ and 116 of ″Verdolino″ were analysed. These three landraces were used to develop multiplex-PCR approaches and to evaluate allelic variation at 30 SSR loci (Yu et al. 1999).

Genomic DNA isolation

Good quality DNA is necessary to guarantee reproducibility of data. Common bean DNA, after conventional extraction, is often contaminated with polysaccharides, which gives a brownish colour to the solution (probably due to the presence of oxidated polyphenolics (Rogers and Bendich 1988)). We therefore tested a range of DNA extraction protocols: a) modification of Dellaporta method (Dellaporta et al. 1983), b) the CTAB method of Saghai Maroof (Saghai-Marrof et al. 1984), c) the Doyle and Doyle method (Doyle and Doyle 1990), and d) the CTAB method of Yu and Pauls (1994). Best results were obtained from young leaves (first true leaf stage), using the latter method. For PCR, DNA samples were diluted to a concentration of 5 ng/ul.

SSR primers

Thirty primer pairs published by Yu et al. (1999, 2000) designed from bean genomic sequence (Table 1) were used to develop SSR multiplex sets. Twenty of these SSR markers have been previously mapped (Gepts 1999; Yu et al. 2000). Primers (20–30 bp) targeting the flanking regions of SSR sequences were selected to have a melting point (Tm) > 46 °C at a concentration of 50 mM Mg^{2+}. Forward primers, synthesised by MWG Biotech (Ebersberg, Germany),

had an IRD 700 or 800 dye bound to the 5′ end. The expected amplicon size was in the range 95–305 bp.

PCR conditions

The 10× PCR buffer contained 100 mM Tris-HCl (pH 8.3 at room temperature), 500 mM KCl, 15 mM $MgCl_2$ and 0.01% (w/v) gelatin. Forward primers were diluted to a stock concentration of 1.5 µmol/µl, and mixed in a 1:1 ratio with unlabelled primer. Reverse primers were unlabelled. PCR amplifications were performed in 96-well microtitre plates using a Perkin Elmer GeneAmp PCR 9700 system. Multiplex PCR reactions (20 ul) contained 25 ng of template, 0.15 uM of each primer, 1X PCR buffer, 200 uM dNTPs, 1 mM of additional $MgCl_2$ and 0.4 U of AmpliTaq Gold DNA polymerase (Applied Biosystems). Three different cycling conditions (Table 2) were initially tested, and the most effective (programme B) was then used for the analysis of diversity using multiplex-PCR. To optimise the multiplex sets, DNA from five plants from each of three common bean landraces was analysed.

Analysis of amplicons

The PCR products (2.5 ul) were denatured in 5 ul of denaturing dye (0.05% pararosaline, 10 mM EDTA and 95% formamide) at 94 °C for 3 min, and 0.6 ul of sample was loaded on a standard 6% polyacrylamide sequencing gel, and run at constant power (40 W) for 2–3 h. We used a LICOR sequencing gel apparatus (MWG Biotech – Ebersberg, Germany), and Gene ImagIR™ software for fragment analysis. Fragment sizing of amplicons was enabled by including a number of tracks containing size 15 IRD-labelled DNA fragments, covering a size range of 50–350 bp. For each electrophoretic run it was possible to visualise an image of the gel both at 700 IRD and 800 IRD. This allowed the scoring of four or five SSRs per sample per run.

Results

Identification of primers and optimisation of single locus PCR

Since longer high GC content primers allow for higher annealing temperatures, and thus yield fewer non-specific products, primers with a 35–60% GC content, corresponding to theoretical annealing temperature of 47–50 °C, and longer than 20 bp were selected from the GenBank sequences (Yu et al. 1999). A diagram of the optimisation procedure for both reaction and PCR conditions is presented in Figure 1.

Optimisation of a PCR is commonly carried out by a sequential investigation of each reaction variable. Initially, the cycling conditions proposed by Yu et al. (1999) were adopted, but the results were unsatisfactory. Thus, simplex PCR conditions were optimised for each primer pair by first testing different cycling conditions and then by varying (1) the amount of DNA template, (2) the concentration of primer, and (3) the concentration of $MgCl_2$ (step 3). To optimise the cycling conditions, three PCR programs (A, B and C) with different times for annealing and extension, and for variation in both temperature and number of cycles were tested (Table 2).

Changes in neither the annealing time (25–60 s) nor the extension time (25–120 s) visibly influenced amplification efficiency. Product yield was generally improved by raising the extension temperature from 68 °C to 72 °C, but the specificity and yield of PCR product were significantly affected by changes in the annealing temperature. To amplify the 30 SSR loci, the cycling conditions of programme B gave the optimal results. The robustness of the assay was confirmed by comparing amplification profiles of identical reactions obtained both from a single thermal cycler, and in independent machines (MWG Primus 96/384HPL, Biorad TCycler™ and Eppendorf Mastercycler system).

Assays of each simplex SSR were scored for complexity on a scale of 1 to 4, following the criteria suggested by Stephenson et al. (1998). Nine loci gave single band amplification products (scale 1), 14 gave a single band with some stuttering, but were easy to score (scale 2), five gave a single band with irregular stuttering, which was sometimes difficult to score (scale 3), and only two gave multiple amplification products (scale 4). The allele size range of each SSR marker was estimated by analysing 15 genotypes from the three landraces.

Optimisation of multiplex PCR reaction components

For multiplexing, an important requirement is that there is no overlap in product size between the individual components of the multiplex. Another issue surrounds any interaction between competing amplicons when multiple primers are present in the reac-

Table 1. List of thirty microsatellite loci of common bean (*Phaseolus vulgaris* L.) identified from the Gen Bank database, and allelic variation in the 264 genotypes of the three local populations: ″Poverella″, ″Marrozzo″ and ″Verdolino″.

Gen-Bank acces-sion number	Name in Genbank (symbol)	Core Motifs	Size (bp)	Primer sequences (forward, reverse)	Tm[b] (°C)	Linkage group	Alleles number
X74919	Endochitinase (PVGEC9)	$(AT)_5$	132	ccgttgcctgtatttccccatcgtgtgaagtcatctggagtggtc	50	B5	4
X04660	Phytohemagglutinin pseudogene (PVPDLEC1)	$(AG)_8$	201	ttgatgacgtggatgcattgcaaagggctagggagagtaagttgg	48	B4	3
X13329	5′ flanking sequence of glutamine syntetase beta subunit gene (PVGLNB)	$(GA)_8(A)_2(GA)_3$	139	gctcacgtacgagttgaatctcagatctgagagcagcgacatggtag	48	Unlinked	5*
M99497	Protein kinase in *Vigna* (VIRPK)	$(AG)_{12}(AAG)_2$	163	gggtagtaaaggaaagagaagaaagagccaccttcctcgtactgttccatg	48	B3	1
X80051	NADP-dependent malic enzyme (PVME1G)	$(AT)_{12}$	192	agttaaattatacgaggttagcctaaatccattcccttcacacattcaccg	49	B9	22
X60000	Small subunit of ribulose 1,5 biphosphate carbosylase/oxygenase (PVRBCOS)	$(AT)_4(T)_2(AT)_6$	139	acctagagcctaatccttctgcgtgaatgtgaatatcagaaagcaaatgg	49	B4	6*
X79722	Sn-glycerol-3-phosphate acyltransferase (PVPLB)	$(CCT)_7$	149	ccaaccacattcttccctacgtcgcggaggcagttatctttaggagtg	49	B2	3*
m75856	Phathogenesis-related protein 3 (PHVPVPR3A)	$(CT)_{11}$	157	caatcctctctctctcatttccaatcgaccttgaagtcggtgtcgttt	47	B11	7
X61293	Promotor region of gln-delta gene for plas-tid-located glutamine synthetase (PVGLND3)	$(AT)_{18}$	163	aatctgccgagagtggtcctgcgattgaaatatcaaagagaattgttacc	47	B6	4
U77935	DnaJ-like protein(PVU77935)	$(GCCACC)_5$	95	cgttagatcccgcccaatagtccgtccaggaagagcgagc	48	B2	5
U54703	Dehydrin (PVU54703)	$(TTA)_4$	106	cgaggaggaggagaagacgggagggttatcacaaggaagacacg	49		2
K03288	Erythroagglutinating phytohemagglutinin (PHVDLECA)	$(ATGC)_4$	126	tgccaccacagctttctcctctatgagagaagcggttggcacg	49	B4	2
U28645	Embryo-specific acidic transcriptional activator(PVU28645)	$(GGT)_5$	115	gcaagagaacactgaagaggatcggacattactcatttcatcatctactacacg	49		4*
K03289	Leucogglutinating phytohemagglutinin (PHVDLECB)	$(ATGC)_4$	144	agctttcacactatgacaccactggtgcgacatgagagaaagacacgg	49	B4	2
J04555	Kinase-1 protein (PHVPVPK)	$(CTT)_3(T)_3(CTT)_6$	152	gagggtgtttcactattgtcactgcttcatggatggtggaggaacag	48	B4	7
X96999	Ribonuclease-like pathogenesis-related pro-tein (PVNPV30G)	$(AT)_9$	161	agtcgccatagttgaaatttaggtgcttattaaaacgtgagcatatgtatcattc	49	B3	9
X57022	Small subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase (PVSS15BCO)	$(AATG)_6$	163	aaggatgggttccgtgcttgcacggtacacgaaaccatgctatc	49	B4	9
X04001	Glutamine syntetase (PVGSR1)	$(AG)_8$	164	tcacgtacgagttgaatctcaggatggtgtcggagaggttaaggttg	49		5

Table 1. Continued

Gen-Bank accession number	Name in Genbank (symbol)	Core Motifs	Size (bp)	Primer sequences (forward, reverse)	Tm[b] (°C)	Linkage group	Alleles number
X59469	Silencer region of chalcone synthase promoter (PVCHALCPO)	$(GAAT)_3$	167	aaacacacaaaaagttggacgcacttcgtgaggtaggagtttggtgg	50	B2	2
J01263	Beta-phaseolin (PHVBCSP)	$(ATCC)_3$ $(AG)_2$ $(TAC)_3$	171	atgcatgttccaaccaccttctcggagtggaacccttgctctctcatc	49	B7	3
X52626	Alpha-phaseolin (PVAPHASE)	$(ATCC)_3$	176	tctccatgcatgttccaaccacggagtggaacccttgctctcatc	49	B7	3
M18094	Hydroxyproline_rich glycoprotein (PHVHRGPB)	$(CCA)_5$	179	taatttctctcttcccatcccaaacgtagtaataaggaggaggcggtgag	49		2
M13968	Chitinase (PHVCHM)	$(GAA)_3$ TAG $(GAT)_5$	182	acaccttatcatttagaggaaaagagaacccgaactggctgcaacag	47		3*
U10419	Nitrite reductase (PVU10419)	$(AAAT)_3$	203	tggagccatctgtcctcttacccacgagcacgagtcacgtttgcaac	49		3
U18349	Phaseolin G-box binding protein (PVU18349)	$(GGC)_5$	238	ctgaagcccgaatcttgcgacggcgagaggtgaacgaaagc	49		5*
U34754	Cellulase (PVU34754)	$(AT)_8$	254	gtttcttccttatggttaggttgtttgtcacgttatcaccagcatcgtagta	49	B2	1 [a]
X63525	Lipoxygenase (PVULOXA)	(AT)7	305	gacgttcgagtattttgtgatatagacataataccatgctcctactaca	48		4
X02980	Alpha-phaseolin (PVVPHASAR)	$(ATCC)_3(AG)_2(TAC)_3T(CTA)_3$	192	acttctttcatcatccatccatcctatcttggctctcttcctcctcc	48	B7	3
M68913	Arcelin (PHVARC1A)	$(ATCT)_3$	193	caattaaaactcaaccaacccaaatatttcccgccatagaatatgtgaga	49	B4	5
X58274	Ferritin (PVPFE)	$(CTT)_6$	164	tcgccgggaaagttgccagttagaaggagcgaggggccatg	49		1[a]

- All primer sequences are given in the 5′ to 3′ direction
- The calculated T_m values are based on a salt concentration of 50 mM
- [a] = faint banding pattern indicating poor amplification;
- * = presence of nulli allele
- Italic accession numbers indicate non polymorphic loci with clearly defined bands

Table 2. Development of PCR Cycling conditions for optimisation of M-PCR.

	Program A	Program B*	Program C
First denaturing (including hot start)	94 °C, 5 min	94 °C, 10 min	94 °C, 8 min
Denaturing	94 °C, 25 sec	94 °C, 1 min	94 °C, 1 min
Annealing	55 °C, 25 sec	47 °C, 1 min	49 °C, 1 min
Extension	68 °C, 25 sec	72 °C, 1 min	72 °C, 2 min
	40 cycles	35 cycles	45 cycles
Final extension	72 °C, 5 min	72 °C, 20 min	72 °C, 10 min

* Program B used according to the type of PCR amplification

tion. The simplex PCR conditions therefore had to be re-optimised by testing different 1) primer concentrations, 2) dNTPs and $MgCl_2$ concentrations and 3) amounts of DNA template and Taq DNA polymerase.

A concentration of 0.15 uM for each primer pair was initially used, but for some primers pairs this proved to be non-optimal, generating an uneven amplification of target loci. The loci VIRPK, PVU34754 and PVPFE amplified poorly, while in contrast, PVGEC9 and PVRBCOS amplified strongly. To obtain even amplification in multiplex, we attempted increasing the concentration of the ″weak″ primers and decreasing those of the ″strong″ primers in the reaction. We tested primer concentration ranges of 0.1–0.3 uM. However, this approach proved to be ineffective, so equimolar primer concentrations of 0.15 μM were retained.

The relationship between the concentrations of Mg^{2+} and dNTPs was investigated in reactions that contained 200, 300, 400 and 500 uM of each dNTP, combined with 1.5, 2, 2.5 and 3 mM $MgCl_2$. From an analysis of 16 different combinations, we concluded that at 200 uM of each dNTP and 1.5 mM $MgCl_2$ was optimal. At DNA template quantities in the range 25–100 ng per 20 ul reaction, there was little variation in amplification efficiency. However, below 25 ng, the amount of some products was decreased, sometimes below the level of detection.

Different concentrations of Taq DNA polymerase (Applied Biosystem) were tested. The most efficient enzyme concentration was in the range 0.4–2.0 U per 25 ul reaction volume. An excess of enzyme resulted in unbalanced amplification of various loci and a slight increase in the background, possibly because of the high glycerol concentration present in the polymerase buffer solution; too little enzyme resulted in poor amplification efficiency. A concentration of 0.4 U per 20 ul reaction volume was used in subsequent experiments.

Optimisation of multiplex PCR

The simultaneous amplification of several loci using mixtures of different combinations of SSR primers usually requires a process of re-optimisation of the parameters of the reaction. We first performed multiplex reactions with primers present in equi-molar amounts. Since the simplex test showed that primer concentration had little effect on amplification efficiency, the simultaneous amplification of two target loci in a single reaction (duplex primer combination) was attempted by combining the 30 primer pairs in various combinations considering: 1) the type of amplification profile; 2) the allele size ranges; and 3) the type of forward primer label with an IRD 700 or 800 dye at their 5′ ends.

Although the PCR parameters adopted followed the recommendations of Henegariu et al. (1997) and Mitchell et al. (1997), our results were discouraging in that five of the 15 duplex primer combinations did not produce a clear pattern, either generating non-specific products and/or successfully amplifying only one of the two targets (presumably an effect of competition between targets). We postulate that a contributory cause of this failure relates to the elevated sensitivity of fluorescence-based detection, as compared to ethidium bromide staining in agarose gels. The quantity of template required for the former is four orders of magnitude less than is necessary for the latter. To avoid non-specific amplification products, due to mispriming and/or undesirable interactions between primers, the PCR conditions adopted were set to be stringent (lowest possible Mg^{2+} concentration and the minimum number of amplification cycles).

Multiplex PCR and multi-loading set development

Multiplex sets, involving at least two primer pairs, were constructed, based on consideration of the expected allele size range (Stephenson et al. 1998) and

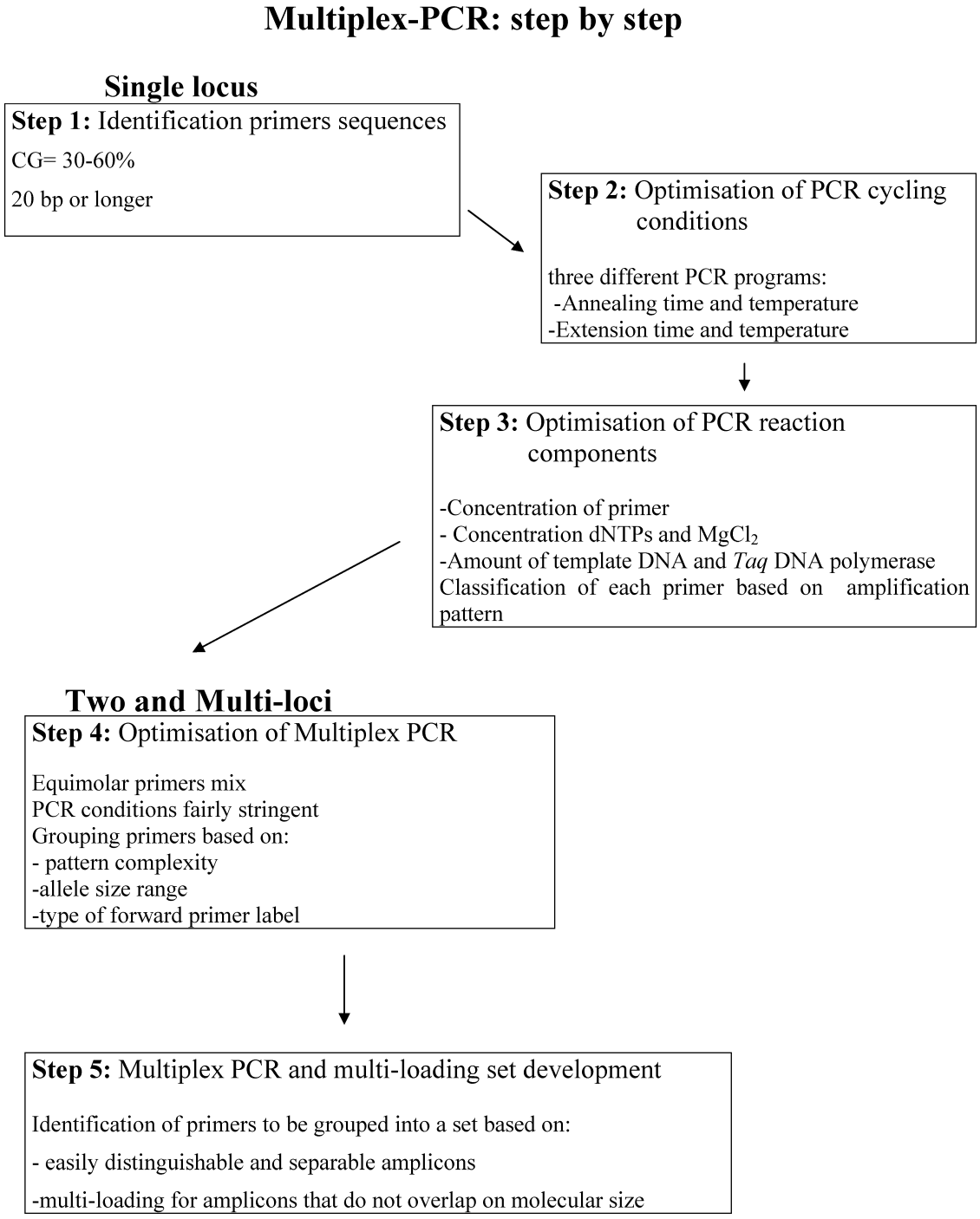

Figure 1. Procedure followed to optimize the efficiency of Multiplex-PCR for analysis of SSR polymorphism in common bean.

the type of forward primer label. Primer pairs that either generated overlapping allele size ranges but carried different labels, or were labelled identically but whose profiles were expected to be separated by at least 20 bp were also included within a set.

The primer combinations are detailed in Table 3, which also lists the number of alleles identified

Table 3. SSR markers included in each multiplex set.

Multiplex set	Total (no)	SSR markers[1]		Total alleles (no)
		IRD 700	800 IRD	
1	4	X79722, X58274	X04660, M99497	**9**
2	4	X61293, X02980	X59469, M68913	**15**
3	4	X96999, X6000	J04555, U10419	**27**
4	4	m75856, U77935	U54703, M13968	**17**
5	5	X80051, X13329	K03289, U28645, J01263	**36**
6	4	X74919, X57022	X04001, K03288	**21**
7	5	X52626, U18349, X63525	M18094, U34754	**15**

[1] SSR Genebank accession number

among the germplasm screened, from each constituent primer pair within the multiplex set.

To obtain completely unambiguous multiplex genotyping, it was important to optimise the combination of primers to be mixed. The best primer combinations were tested several times to ensure reproducibility. Particular attention was paid to loci which generated multiple band profiles, since this has an impact on the quality of the overall profile.

Because of a high allele size range and/or because of their inherent pattern complexity, two of the loci (X80051 and m75856) were amplified separately, and then combined in multiplex loading. Loading more than one PCR reaction in each track was also attempted for other combinations in an effort to reduce high background signal or to eliminate overlapping of bands. By loading in the same track the products of two PCR reactions, each including two to three primer pairs per single assay, the relative efficiency of each of the 30 primer pair combinations could be rapidly identified. When applied to the landrace collection, we obtained consistent multiplex amplification of SSRs with the 30 primer pairs for all 264genotypes (an example is shown in Figure 2a and 2b).

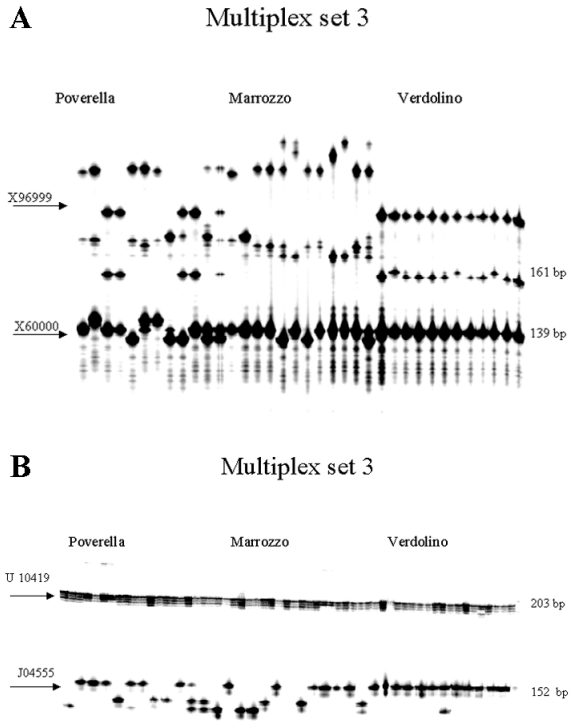

Figure 2. **a** An example of SSR polymorphism detected by 2 primers pairs labelled with 700 IRD dye (X96999 and X60000) in accessions of common bean: 15 genotypes of ″Poverella″, 15 genotypes of ″Poverella″, 15 genotypes of ″Marrozzo″ and 15 of ″Verdolino″). **b** An example of SSR polymorphism detected by 2 primers pairs labelled with 800 IRD dye (U10419 and J045550) in accessions of common bean: 15 genotypes of ″Poverella″, 15 genotypes of ″Marrozzo″ and 15 of ″Verdolino″).

Discussion

The importance of SSRs as genetic markers in plant is well established and documented (Morgante and Olivieri 1993; Akagi et al. 1996; Mitchell et al. 1997; Milbourne et al. 1998; Doyle and Doyle 1990), but the high cost for their development and the low number of loci per experiment that can be simultaneously analysed limit their application for the analysis of genetic diversity where a large number of individuals must be analysed, particularly in species of restricted commercial significance. In contrast, other molecular markers – in particular AFLPs – allow a significant number of loci to be scored per reaction, such that the amount of information recovered per experiment is of an order of a magnitude higher than is possible for SSRs (Powell et al. 1996).

In order to reduce this differential, we increased the efficiency of the SSR assay in bean by multiplex-

ing and multi-loading. For this purpose, the most time consuming step was the optimisation of the multiplex-PCR; but also critical was the determination of which markers could be grouped together in any given multiplex set. Finally, however, it proved possible to accurately and cost effectively fingerprint each genotype at 30 SSR loci, using only five gel tracks and 14 PCR runs. This reduced the cost of the PCR reagents by over 50%, and the gel costs by about 85%, compared with the conventional SSR procedures.

The high ability of simplex SSR to distinguish between genotypes, its robustness as an assay, and the ease of conversion into an allele-specific marker are all attractive features, but multiplex-PCR, coupled with semi-automated separation systems can significantly enhance genotyping efficiency. However, both multiplexing and multi-loading can only be effective if the range of electrophoretic profiles of each amplicon is known *a priori*, and if technical problems associated with non-specific amplification, band stuttering and high background signal can be eliminated by optimisation of the PCR parameters.

In this study, the multiplex method we developed was successfully applied to the genetic characterisation of three heterogeneous common bean landraces. The 264 genotypes were analysed with each multiplex set to ensure that the allele size range of multiplexed SSRs with the same label would not overlap. Additional alleles were detected, but none within any set overlapped. The maximum fractional difference between size estimates for the same fragment was around 20–30 bp for each set. This was in contrast with what has been reported for human, cow, *Citrus* and Brassica spp. (Kimpton et al. 1994; Glowatzki-Mullis et al. 1995; Kijas et al. 1995; Mitchell et al. 1997) where the maximum fractional difference between sizes was estimated to be around 5.0 nucleotide or less. Both multiplexing and multi-loading gave consistent results with respect to the identification of polymorphism and the definition of individual alleles. The 30 SSR primer pairs detected 1–20 alleles (mean of 4.3) per SSR locus. Twenty-seven loci were polymorphic among all three landraces: five (18.5%) were diallelic, seven (25.9%) were tri-allelic, and 15 had a higher number of alleles (Table 1). A particularly high level of polymorphism was observed for the loci PVME1G (22 alleles), PVNPV30G and PVSS15BCO (nine alleles), PHVPVPR3A (seven alleles), and PVRBCOS (six alleles), both between and within the landraces, and on more than 2 alleles per locus were found.

Null alleles were detected in all three landraces at loci PVRBCOS, PVGLNB, PVPLB, PVU28645, PHVDLECB, and PHVHRGPB, and these were confirmed through repeated PCR amplifications in order to exclude the possibility of failed PCR reactions (Bell and Ecker 1994). Nulls are thought to arise as a result of divergence in the SSR flanking sequences (Smulders et al. 1997), since deletions or point mutation(s) in the primer-annealing site would be expected to inhibit or prevent primer binding and hence amplification. Although null alleles are rare among human SSRs, they are relatively common in barley and in various polyploid species, particularly wheat (Devos et al. 1995; Donini et al. 1998).

Previous studies (Moore et al. 1991; Morgante and Olivieri 1993; Wang et al. 1994; Gupta et al. 1996) have shown that AT/TA and AG/CT repeats are relatively the more abundant in plant, as compared to mammalian genomes. This difference may in part be the result of different DNA methylation patterns or functional roles between plants and animals. In the present study, two SSRs with di-nucleotide repeats (AT/TA and CT/AG) were identified, and loci with $(AT)_n$ repeats were the most frequent (47%). Among the 30 SSRs, ten were composed of dinucleotide repeats, seven of trinucleotide repeats, 12 of tetranucleotide repeats and one a hexanucleotide repeat. Among the SSRs, nine (90%) of the ten dinucleotide-, five (71.9%) of the seven trinucleotide-, and five (41.6%) of the 12 tetranucleotide-carriers detected polymorphism among genotypes within and between the landraces. The total number of SSRs with di- and trinucleotide motifs found in this survey was higher than that with tetra- or higher nucleotide core motifs, confirming previous surveys in plants (Akagi et al. 1996; Chin et al. 1996). Compared to mammalian and soybean genomes (Levitt et al. 1994; Akkaya et al. 1995) the total length of repeat arrows appears to be much shorter in common bean. This is probably because most of sequences examined originated from within or adjacent to coding regions. In general, the number of alleles at an SSR locus is directly correlated with the number of repeats present, with longer repeats being more polymorphic than shorter ones (Weber 1990; Innan et al. 1997), but this is not a universal observation (Beckmann and Weber 1992; Schlotterer 1997). In our materials, loci X80051, X96999 and m75856 are highly polymorphic, but they carry rather small numbers of repeats.

Conclusions

An optimal annealing temperature is essential in any PCR to ensure the specificity of the amplificon. Mg^{2+} concentration needs only to be proportional to the amount of dNTPs, and this value can be constant for any reaction. In multiplex PCR, adjusting primer concentration for each locus is also essential. The optimisation of parameters, as presented in this study, provides a basic protocol to solve some of the common problems of multiplex PCR. Overall, this study has shown that: 1) the frequency of polymorphic SSRs in common bean genome is fairly high; 2) some common bean SSR loci are multi-allelic; 3) multiplex PCR enhances the efficiency of SSR markers, thus providing a marker system that allows the extent of heterozygosity to be assessed; 4) multiplex PCR and fluorescence-based, semi-automated sizing technology is an accurate and robust genotyping method; and 5) the multiplex approach has the advantage of reducing the unit cost of the assay. There is, however, a trade-off between the investment labour to establish this method and the time saved by the need to run fewer gels. The methodology developed will be useful to study levels and patterns of genetic diversity in bean collections, where a large number of accessions and genotypes needs to be fingerprinted.

References

Akagi H., Yokozeki Y., Inagaki A. and Fujimura T. 1996. Microsatellite DNA markes for rice chromosome. Theor. Appl. Genet. 93: 1071–1077.

Akkaya M.S., Shoemaker R.C., Specht J.E., Bhagwat A.A. and Cregan P.B. 1995. Integration of simple sequence repeats DNA markers into a soybean linkage map. Crop. Sci. 35: 1439–1445.

Becerra V.L. and Gepts P. 1994. RFLP diversity of common bean (*Phaseolus vulgaris* L.) in its centres of origin. Genome. 37: 256–263.

Beckmann J.S. and Weber J.L. 1992. Survey of human and rat microsatellites. Genomics 12: 627–631.

Beebe S., Skroch P.W., Tohme J., Duque M.C., Pedraza F. and Nienhuis J. 2000. Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. Crop. Sci. 40: 264–273.

Bell C.J. and Ecker J.R. 1994. Assigment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19: 137–144.

Caicedo A.L., Gaitan E., Duque M.C., Toro Chica O., Debouck D.G. and Tohme J. 1999. AFLP fingerprinting of *Phaseolus lunatus* L. and related wild species from South America. Crop. Sci. 39: 1497–1507.

Chin E.C.L., Senior M.L., Shu H. and Smith J.S.C. 1996. Maize simple repetitive DNA sequence Abundance and allele variation. Genome. 39: 866–873.

Cregan P.B. 1992. Simple sequence repeat DNA length polymorphisms. Probe. 2: 18–22.

Dellaporta S.L., Wood J. and Hicks J.B. 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1: 19–21.

Devos K.M., Bryan G.J., Collins A.J., Stephenson P. and Gale M.D. 1995. Application of two microsatellite sequences in wheat storage proteins as molecular markers. Theor. Appl. Genet. 90: 247–252.

Donini P., Stephenson P., Bryan G.R. and Koebner R.M.D. 1998. The potential of microsatellites for high throughput genetic diversity assessment in wheat and barley. Gen. Res. Crop. Evol. 45: 415–421.

Doyle J.J. and Doyle J.L. 1990. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11–15.

Gepts P. 1999. Development of an integrated linkage map in common bean (*Phaseolus vulgaris* L.) and its use. In: Singh S. (ed.), Bean Breeding for the 21st Century. Kluwer, Dordrecht, the Netherlands, pp. 53–91, 389–400.

Glowatzki-Mullis M.L., Gaillard G., Wigger G. and Fries R. 1995. Microsatellites-based parentage control in cattle. Animal Genetic 26: 7–12.

Gupta H.S., Balyan P.C. and Ramesh Sharma B. 1996. Microsatellites in plants: A new class of molecular markers. Current Science 70: 1–10.

Henegariu O., Heerema N.A., Dlouhy S.R., Vance G.H. and Vogt P.H. 1997. Multiplex PCR: Critical parameters and step-by step protocol. BioTechniques 23: 504–511.

Innan H., Terauchi R. and Miyashita T. 1997. Microsatellite polymorphism in natural populations on the wild plant *Arabidopsis thaliana*. Genetics 146: 1441–1452.

Kijas J.M.H., Fowler J.S.C. and Thomas M.R. 1995. An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. Genome. 38: 349–355.

Kimpton C., Fisher D., Watson S.K., Adams M., Urquahart A., Lugo J. et al. 1994. Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. Int. J. Leg. Med. 106: 302–311.

Levitt R.C., Kiser M.B., Dragwa C., Jedlika A.E., Xu J., Meyer M.A. et al. 1994. Fluorescence-based resources for semi-automated genomic analysis using microsatellite marker. Genomics 24: 361–365.

Masi P. 2001. Analysis of the genetic structure of *Phaseolus vulgaris L. landraces* collected in Basilicata, Italy. PhD Dissertation.

Masi P., Figliuolo G. and Spagnoletti Zeuli P.L. 1999. Landraces of beans (*Phaseolus vulgaris* L.) collected in Basilicata (Italy). Pl. Gen. Res. New. 119: 51–55.

Metais I., Hamon B., Jalouzot R. and Peltier D. 2002. Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. Theor. Appl. Genet. 104: 1346–1352.

Milbourne D., Meyer R.C., Collins A.J., Ramsay L.D., Gebhardt C. and Waugh R. 1998. Isolation, characterisation and mapping of simple sequence repeat loci in potato. Mol. Gen. Genet. 259: 233–245.

Mitchell S.E., Kresovich S., Jester C.A., Hernandez C.J. and Szewc-McFadden A.K. 1997. Application of multiplex-PCR and

fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop. Sci. 37: 617–624.

Moore S.S., Sargenat L.L., King T.J., Mattick J.S., Georges M. and Metzel D.J.S. 1991. The conservation of the di-nucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. Genomics 10: 654–660.

Morgante M. and Olivieri A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. The Plant Journal 3: 175–182.

Powell W., Gordon Marchay C. and Provan J. 1996. Polymorphism revealed by simple sequence repeats. (Reviews) Trends in Plant Science 1: 215–222.

Rafalsky J.A. and Tingey S.V. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet. 9: 275–279.

Rogers S.O. and Bendich A.J. 1988. Extraction of DNA from plant tissues. In: Gelvin S.B. and Schilperoot R.A. (eds), Plant Molecular Biology Manual. Vol. A6. Kluwer, Dordrecht, pp. 1–10.

Saghai-Marrof M.A., Soliman K.M., Jorgensen R.A. and Allard R.W. 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc. Natl. Acad. Sci. USA 81: 8014–8018.

Schlotterer C. 1997. Microsatellites. In: Hoelzel A.R. (ed.), Molecular Genetic Analysis of Populations. A Practical Approach. IRL Press.

Schlotterer C. and Tautz D. 1992. Slippage synthesis of simple sequence DNA. Nucl. Ac. Res. 20: 211–215.

Smulders M.J.M., Bredemeijer G., Rus-Kortekaas W., Arens P. and Vosman B. 1997. Use of short microsatellites from data base sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor. Appl. Genet. 97: 264–272.

Stephenson P., Bryan G., Kirby J., Collins A., Devos K., Busso C. et al. 1998. Fifty new microsatellite loci for the wheat genetic map. Theor. Appl. Genet. 97: 946–949.

Tautz D. and Renz M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucl. Acid Res. 12: 4127–4138.

Tohme J., Gonzàlez D.O., Beebe S. and Dunque M.C. 1996. AFLP analysis of gene pool a wild bean core collection. Crop. Sci. 36: 1375–1384.

Wang Z., Weber J.L., Zhong G. and Tanksley S.D. 1994. Survey of plant short tandem DNA repeats. Theor. Appl. Genet. 88: 1–6.

Weber J.L. 1990. Informativeness of human $(dC\text{-}dA)_n/(dG\text{-}dT)_n$ polymorphisms. Genomics 7: 524–530.

Yu K., Soon S.J. and Poysa V. 1999. Abundance and variation of microsatellite DNA sequences in beans *Phaseolus* and *Vigna*. Genome. 42: 27–34.

Yu K. and Pauls K.P. 1994. Optimisation of DNA extraction and PCR procedures for random amplified polymorphic DNA (RAPD) analysis in plants. In: Griffin H.G. and Griffin A.M. (eds), PCR Technology: Current Innovations. CRC press, Boca Raton, Florida, pp. 193–200.

Yu K., Park S.J., Poysa V. and Gepts P. 2000. Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). Journal of Heredity 91: 429–434.