Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.)

M.W. Blair, H.F. Buendía, L. Díaz, J.M. Díaz, M.C. Giraldo, E. Tovar, M.C. Duque, S.E. Beebe, D. Debouck

Bean Project and Genetic Resource Unit, CIAT - International Center for Tropical Agriculture, Apartado Aéreo 6713, Cali, Colombia

Centro Internacional de Agricultura Tropical
International Center for Tropical Agriculture

Introduction

Common bean is the third most important grain legume in the world produced over an area of 18 million hectares with large amounts of production in developing countries of Latin America and Eastern and Southern Africa (Broughton et al., 2003). Cultivated common bean germplasm is especially diverse due to the existence of two genepools in the Mesoamerican and Andean centers of diversity and these genepools can be distinguished into various races (Singh et al., 1991). In this poster, we describe the evaluation of a mini core collection of useful parental genotypes with 129 microsatellite markers. Allelic diversity and heterozygosity of these microsatellite markers is also described. Two types of microsatellites were evaluated, based respectively on gene coding and genomic sequences. Genetic diversity was evaluated by estimating the polymorphism information content (PIC), as well as the distribution and range of alleles sizes.

Materials and Methods

Genotypes: A total of 43 common bean (*P. vulgaris*) and one tepary bean (*P. acutifolius*) genotypes were used in this study (Table 1; Blair et al., 2006). The genotypes represented parents of genetic mapping populations being studied at CIAT for biotic or abiotic stress resistance and nutritional quality. The genotypes were grouped in three parental surveys with common controls run in each gel, namely the Mesoamerican DOR364 and the Andean G19833. Among the 43 common bean genotypes were a total of 12 Andean (10 cultivated and 2 wild) and 31 Mesoamerican (30 cultivated and 1 wild) genotypes. The growth habit of each genotype was classified from I (determinate bush) to IV (indeterminate climber) according to Singh et al. (1991). DNA extraction were from 10 seedlings selected at random from each accession with the method of Afanador et al. (1993). DNA was diluted to 10 ng/ml for further experiments.

Microsatellite Analysis: Markers included the 150 SSRs analyzed by Blair et al. (2003). PCR reactions were carried out in 12 µl reaction volumes containing 20 ng of genomic DNA, 0.15µM each of forward and reverse primers, 200 µM of total dNTP and 1 unit of Taq polymerase in 1X PCR buffer (10 mM of Tris-HCl (pH 7.2), 50 mM of KCl). Amplification products were evaluated on 4% denaturating polyacrylamide gels which were run in Sequi-Gen GT electrophoresis units (Bio-Rad, Hercules CA) at 100 constant Watts for approximately 1.5 hours and silver stained. Allele sizes were estimated based on comparison of microsatellite bands to a 10-pb molecular-weight ladder (Invitrogen, Carlsbad, CA) that was placed every twentieth lane in one hundred lane gels.

Data Analysis: Polymorphism information content (PIC) of each microsatellite was calculated using the formula: PIC = 1 -åpij 2, where pij is the frequency of the patterns (j) for each marker (i). The microsatellite dataset was also used for a multiple correspondence analysis (MCA) of the genotypes and for UPGMA clustering in SAS v. 9.1.3 (SAS Institute, Cary NC).

Tables and Figures

Table 1. Common bean genotypes used for assessment of microsatellite diversity and their accession number, phaseolin status, race and gene pool identity, origin and growth habit.

Genotype	Ph	Genepool	Race	Status	Origin	GH
G 11360	S	Mesoamerican	J	Cultiv	Mexico	IV
G 11350	S	Mesoamerican	M	Cultiv	Mexico	III
G 21657	C	Andean	P	Cultiv	Bulgaria	III
G 21078	T	Andean	P	Cultiv	Argentina	IV
G 21242	C	Andean	na	Cultiv	Colombia	IV
G 14519	S	Mesoamerican	M	Cultiv	USA	IV
G 4825	B	Mesoamerican	M	Cultiv	Brazil	III
G 19833	C	Andean	P	Cultiv	Peru	III
DOR 364	S	Mesoamerican	M	Cultiv	CIAT	II
BAT 477	S	Mesoamerican	M	Cultiv	CIAT	II
G 3513	S	Mesoamerican	M	Cultiv	Mexico	II
BAT 881	S	Mesoamerican	M	Cultiv	CIAT	II
G 21212	B	Mesoamerican	M	Cultiv	Colombia	II
G 24404	C	Andean	na	Wild	Colombia	IV
Radical Cerinza	T	Andean	P	Cultiv	Colombia	I
G 24390	M	Mesoamerican	na	Wild	Mexico	IV
DOR 390	S	Mesoamerican	M	Cultiv	CIAT	II
G19892	T	Andean	na	Wild	Argentina	IV
DOR 476	S	Mesoamerican	M	Cultiv	CIAT	II
SEL 1309	S	Mesoamerican	M	Cultiv	CIAT	II
BAT 93	S	Mesoamerican	M	Cultiv	CIAT	II
Jalo EEP558	T	Andean	NG	Cultiv	Brazil	III
ICA Pijao	B	Mesoamerican	M	Cultiv	Colombia	II
G 40001	na	Tepary Bean	na	Cultiv	Mexico	III
VAX 6	S	Mesoamerican	M	Cultiv	CIAT	II
MAR 1	S	Mesoamerican	M	Cultiv	CIAT	II
J 117	S	Mesoamerican	J	Cultiv	Mexico	III
JAMAPA	S	Mesoamerican	M	Cultiv	Mexico	II
G 2333	S	Mesoamerican	G	Cultiv	Mexico	IV
G 19839	T	Andean	P	Cultiv	Peru	III
G 855	Sb	Mesoamerican	J	Cultiv	Mexico	IV
BRB 191	T	Andean	NG	Cultiv	CIAT	II
MAM 49	S	Mesoamerican	D	Cultiv	CIAT	III
G 5273	T	Andean	NG	Cultiv	Mexico	II
MAM 38	S	Mesoamerican	D	Cultiv	CIAT	III
SEQ 1027	T	Andean	NG	Cultiv	CIAT	III
G 4090	Sd	Mesoamerican	M	Cultiv	El Salvador	II
Tio Canela	S	Mesoamerican	M	Cultiv	EAP	II
DOR 714	S	Mesoamerican	M	Cultiv	CIAT	II
SEA 5	S	Mesoamerican	D	Cultiv	CIAT	II
MD 23-24	S	Mesoamerican	M	Cultiv	EAP	II
SEA 15	S	Mesoamerican	D	Cultiv	CIAT	II
G 685	Sb	Mesoamerican	G	Cultiv	Guatemala	IV
SEA 21	S	Mesoamerican	M	Cultiv	CIAT	II

Table 2. Number of alleles and polymorphism information content (PIC) of 129 common bean microsatellite markers.

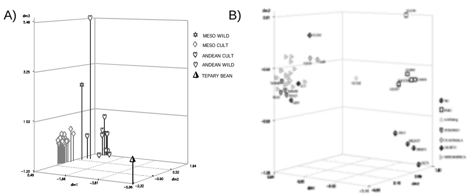
(57 gene-based and 72 genomic)

Marker	No. of alleles	PIC	Expected Allele
A) Gene-based			
BMd1	28	0.942	165
BMd2	6	0.629	106
BMd3	7	0.626	223
BMd4	3	0.129	146
BMd5	3	0.129	122
BMd6	2	0.087	122
BMd7	7	0.730	166
BMd8	9	0.724	176
BMd9	6	0.287	135
BMd10	9	0.761	139
BMd13	3	0.168	194
BMd14	3	0.206	186
BMd15	9	0.722	166
BMd16	6	0.526	136
BMd17	7	0.666	116
BMd18	5	0.757	156
BMd19	4	0.209	154
BMd20	10	0.793	123
BMd21	5	0.387	190
BMd22	6	0.596	121
BMd23	2	0.087	127
BMd25	3	0.400	118
BMd26	4	0.458	141
BMd27	2	0.236	109
BMd28	13	0.874	151
BMd30	4	0.170	134
BMd31	2	0.044	161
BMd32	3	0.206	150
BMd35	1	0.000	128
Gene-based (cont'd)			
BMd37	6	0.799	134
BMd45	2	0.416	129
BMd46	4	0.379	158
BMd47	4	0.573	150
BMd48	1	0.000	131
BMd49	2	0.201	95
BMd50	3	0.334	124
BMd51	4	0.407	116
BMd53	5	0.574	105
BMd55	3	0.088	188
PV-CTT001	14	0.822	152
PV-AG001	9	0.546	157
PV-AG003	7	0.721	164
PV-GAAT002	5	0.448	156
PV-TTTC001	7	0.448	161
PV-AT007	20	0.941	192
PV-AT001	24	0.943	170
PV-CTT002	3	0.305	218
PV-AG004	9	0.546	201
PV-AG004b	12	0.829	202
PV-AAAT001	1	0.000	205
PV-ATCC001	1	0.000	172
PV-ATCC002	2	0.044	192
PV-ATCC003	3	0.501	178
PV-ATCT001	3	0.206	196
PV-CCCT001	3	0.501	150
PV-CCT001	7	0.684	137
PV-TTTC001	6	0.638	143
B) Genomic			
AG1	7	0.483	132
BM003	2	0.325	193
BM006	2	0.044	153
BM048	2	0.087	232
BM053	21	0.927	287
BM068	4	0.493	170
BM098	5	0.601	247
BM114	13	0.818	234
BM137	23	0.937	155
BM138	8	0.779	203
BM139	13	0.632	107
BM140	8	0.459	190
BM141	16	0.813	218
BM142	6	0.537	157
BM143	18	0.887	143
BM146	4	0.170	281
BM147	3	0.088	178
BM148	2	0.044	295
BM149	6	0.320	273
BM151	8	0.819	153
BM152	16	0.897	127
BM153	15	0.876	226
BM154	23	0.936	218
BM155	5	0.387	114
BM156	16	0.818	267
BM157	7	0.510	113
BM159	6	0.590	198
BM160	11	0.580	211
BM161	13	0.873	185
BM164	12	0.495	182
BM165	11	0.826	177
BM166	2	0.044	151
BM167	12	0.693	165
BM170	14	0.877	179
BM172	11	0.580	107
BM175	16	0.749	170
B) Genomic (cont'd)			
BM181	9	0.571	192
BM183	11	0.839	149
BM184	10	0.684	160
BM185	11	0.833	105
BM187	18	0.854	191
BM188	18	0.880	177
BM189	6	0.647	114
BM195	2	0.044	138
BM197	4	0.557	201
BM199	18	0.913	304
BM200	16	0.893	221
BM201	11	0.820	102
BM202	7	0.796	156
BM205a	9	0.707	137
BM209	14	0.848	129
BM210	17	0.880	166
BM211	13	0.848	186
BM212	8	0.517	214
BM213	3	0.394	154
BMd11	3	0.088	161
BMd12	3	0.623	167
BMd33	7	0.690	110
BMd36	12	0.762	164
BMd38	2	0.044	178
BMd39	1	0.000	126
BMd40	7	0.801	197
BMd41	7	0.716	245
BMd42	10	0.764	149
BMd43	4	0.414	176
BMd44	4	0.210	135
BMd56	5	0.534	193
BMd57	2	0.044	140
GATS11	5	0.683	306
GATS11B	3	0.385	160
GATS54	4	0.563	114
GATS91	16	0.914	229

Table 3. Observed intra (Hs) and inter population (Hsi) diversity for genotypes belonging to wild and cultivated common beans, to Andean and Mesoamerican gene pools and to races within each gene pool.

Category	N	Observed Heterogeneity			Value
		cDNA based (57)	Genomic (72)	Total	
Total	44	0.444	0.593	0.527	H_t
Species/Status[1]	44	0.429	0.575	0.511	H_s
Cultivated *P. vulgaris*	40	0.432	0.583	0.516	H_{si}
Wild *P. vulgaris*	3	0.388	0.477	0.437	H_{si}
Tepary Bean *P. acutifolius*	1	0.000	0.000	0.000	H_{si}
Gene pools	40	0.343	0.486	0.422	H_s
Mesoamerican	30	0.319	0.481	0.410	H_{si}
Andean	10	0.412	0.500	0.461	H_{si}
Races	40	0.253	0.363	0.314	H_s
Nueva Granada	4	0.215	0.352	0.292	H_{si}
Peru	5	0.397	0.436	0.419	H_{si}
Introgressed	1	0.000	0.000	0.000	H_{si}
Durango	4	0.154	0.325	0.249	H_{si}
Guatemala	2	0.246	0.292	0.271	H_{si}
Jalisco	3	0.257	0.367	0.319	H_{si}
Mesoamerica	21	0.289	0.430	0.368	H_{si}

Figure 1. Multiple correspondance analysis using UPGMA clustering for 44 common bean genotypes based on 129 microsatellite markers where A) includes wild common bean (*Phaseolus vulgaris* L.) and tepary bean (*P. acutifolius* Gray) outgroups and B) includes only cultivated common bean classified according to races within Andean and Mesoamerican genepools.



References

1. Afanador L, Hadley S, Kelly JD (1993) Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean (*Phaseolus vulgaris* L). Bean Improv Coop 36:10-11
2. Blair MW, Pedraza F, Buendia H, Gaitan-Solis E, Beebe S, Gepts P,Tohme J (2003) Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L). Theor Appl Genet 107:1362-1374.
3. Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, Beebe SE (2006) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.) Theor Appl Genet 113: 100–109.
4. Broughton WJ, Hernandez G, Blair MW, Beebe SE, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.); Model Food Legumes. Plant & Soil 252: 55-128
5. Díaz LM, Blair MW (2006) Race structure within the Mesoamerican gene pool of common bean (*Phaseolus vulgaris* L.) as determined by microsatellite markers. Theor Appl Genet 114: 143-54.
6. Singh. S., Gepts. P. & Debouk. D (1991) Races of common bean (*Phaseolus vulgaris*, Fabaceae). Econ. Bot. 45(3): 379-396.

Results and Discussion

➢ Comparable allele sizes could be ascertained across the three parental surveys for a total of 129 SSRs (57 gene based and 72 genomic) and these were used for PIC calculations (Table 2) and diversity evaluation. Notably, gene-based microsatellites proved to be less polymorphic than genomic microsatellites for number of alleles (6.0 vs. 9.2) and PIC values (0.446 vs. 0.594).

➢ The microsatellites were useful for identifying Andean and Mesoamerican genotypes, for uncovering the races within each genepool and for separating wild accessions from cultivars in the multiple correspondence analyses (Figure 1).

➢ Principal inertia adjustments according to Benzecri were used to calculate the variance explained by each dimension in the MCA. In figure 1A, the tepary bean accession was differentiated from all common bean accessions in a first dimension which explained 24.3% of the variance. The wild accessions from Colombia and Mexico (G24404 and G24390) were distinguished from the domesticated common bean genotypes of each gene pool in a second dimension which explained 19.3% of the variance.

➢ Within the cultivated genotypes of common beans there were two principal clusters in the multiple correspondence analysis corresponding to the Andean and Mesoamerican gene pools. In Figure 1A, these were predominantly separated by the third dimension which explained 7.0% of the variance.

➢ Separation of races is shown in Figure 1B. Within the Andean group there was evidence for two subgroups; the first of which corresponded to the Peru race (G21078, G19833, G19839, G21657 and Radical Cerinza) and the second to the Nueva Granada race (BRB191, G5273, Jalo EEP558 and SEQ1027). Meanwhile, one of the cultivated Andean genotypes, G21242, showed signs of introgression from the Mesoamerican gene pool and was found half way between these two gene pools. Within the Mesoamerican group there was less distinction or race structure although the Guatemala race genotypes (G685 and G2333) were associated and were separate from both Jalisco and Mesoamerica race genotypes.

➢ Values for total diversity (Ht), intra population diversity (Hs) and inter population diversity (Hsi) as well as population differentiation (Gst) coefficients were also calculated for the races and genepools (Table 3). Within the Andean gene pool, race Peru had higher diversity compared to race Nueva Granada, while within the Mesoamerican gene pool, the races Durango, Guatemala and Jalisco had comparable levels of diversity which were below that of race Mesoamerica.

Conclusions

Microsatellites are a priority tool for genetic diversity assessment in common beans and are leading to a better understanding of race structure and crop genetic diversity for common beans, especially for the races or genepool where previous marker systems were less polymorphic (Blair et al., 2006; Díaz and Blair, 2006).