Mol Breeding (2012) 30:1207–1217
DOI 10.1007/s11032-012-9709-8

Marker-assisted selection for low phytic acid (*lpa1-1*) with single nucleotide polymorphism marker and amplified fragment length polymorphisms for background selection in a maize backcross breeding programme

R. Naidoo · G. M. F. Watson · J. Derera · P. Tongoona · M. D. Laing

Received: 13 September 2011 / Accepted: 7 February 2012 / Published online: 28 February 2012

Abstract The level of phytic acid is difficult to assess in a maize breeding programme, therefore a co-dominant single nucleotide polymorphism (SNP) marker was used to detect the single recessive low phytic acid (*lpa1-1*) gene in a BC2F1 population developed from a locally adapted tropical normal inbred line (P 16) and CM 32 (*lpa1-1* donor). High-resolution melt analysis of the *lpa1-1* SNP marker was able to identify 11 homozygous recessive and 17 heterozygote genotypes for the *lpa1-1* mutation. The SNP R^2 values for the heterozygotes were higher (90.95–99.59%) than the *lpa1-1* recessives (82.81–99.58%). The selected BC2F1 lines were fingerprinted with six amplified fragment length polymorphism (AFLP) *Eco*RI/*Mse*I primer combinations to determine the amount of recurrent parent genome present. The 277 AFLP markers were clearly able to differentiate all the BC2F1 lines from each other and the parental controls with a similarity range from 62.12 to 92.15%. It is expected in the BC2 generation to find 87.5% similarity to the recurrent parent, however in this study higher levels of similarity in 13 BC2F1 lines (six heterozygotes and seven homozygous recessive) with 92.15–83.33% similarity were observed. The use of marker-assisted selection for foreground and background selection greatly increased the efficiency of detection of the homozygous recessive (99.58%) and heterozygous (99.59%) genotypes as well as improving the recovery of the recurrent parent (92.15%) in the BC2F1 generation of the maize backcross breeding programme.

Keywords Phytic acid · Maize breeding · Single nucleotide polymorphism · Amplified fragment length polymorphism

R. Naidoo (✉) · J. Derera · P. Tongoona · M. D. Laing
School of Agricultural Sciences and Agribusiness, African Centre for Crop Improvement, University of KwaZulu-Natal, Scottsville 3209, South Africa
e-mail: naidooro@gmail.com

G. M. F. Watson
School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal, Scottsville 3209, South Africa

Introduction

Phytic acid is an anti-nutritional compound found in the seeds of cereals and legumes. It chelates essential minerals such as iron, zinc, magnesium, calcium and potassium as it passes through the digestive system of monogastric consumers such as humans and pigs. There is a considerable amount of evidence to support the fact that dietary phytate has a negative effect on the bioavailability of dietary minerals in humans, as the substitution of low phytic acid (LPA) grain in a maize-based diet is associated with a substantial increase in zinc (Adams et al. 2002; Hambidge et al. 2004), iron (Mendoza et al. 1998, 2001; Hurrell et al. 2003), calcium (Hambidge et al. 2005) and magnesium (Bohn

et al. 2004) absorption. In animal feeding studies using low phytate maize, there was an increase of 2–5 times the amount of bioavailable phosphorus observed (Douglas et al. 2000; Li et al. 2000; Spencer et al. 2000a, b; Yan et al. 2000, 2003; Veum et al. 2001; Peter and Baker 2002) which helps reduce unwanted phosphorus in livestock manure (Ertl et al. 1998; Veum et al. 2001).

There have been temperate maize lines developed that contains lower levels of phytic acid but these are not adapted to local tropical and subtropical conditions. There is a need, therefore, to have the *lpa1-1* gene introgressed into locally adapted lines to improve their nutritional benefit. The low phytic acid (*lpa1-1*) mutation, which shows 66% reduction in phytic acid phosphorus, was generated by ethylmethane sulphonate-induced (EMS) pollen grain mutation (Raboy 2000). The *lpa1-1* gene was first developed by Raboy et al. (2000) using a synthetic population (early ACR) as the seed parent and mutagenised pollen from inbred lines A632 and Mo17. The *lpa1-1* gene has been mapped onto chromosome 1 distal region and is reported to be due to a single recessive mutation (Raboy et al. 2000). Shi et al. (2007) sequenced the *lpa1-1* gene (5,149 bp) and reported the *lpa1-1* phenotype to be probably due to an amino acid change of alanine to valine that occurred within the myo-inositol-1-phosphate synthase. The co-dominant *lpa1-1* SNP marker is able to detect homozygous dominant, homozygous recessive and heterozygous genotypes using PCR and high-resolution melt (HRM) analysis. The gene sequence used to design the primers for the SNP marker was published by Shi et al. (2007). The marker was validated by forward and reverse sequencing of parental PCR products (150 bp) and a single base change from C to T was found which confirmed the amino acid change (Naidoo 2010).

Marker-assisted selection (MAS) that entails the use of molecular markers in the selection process in plant breeding programmes has been shown to be cost-effective (Abalo et al. 2009) and is most useful for the introgression of recessive traits. SNP markers are biallelic, abundant in genomes, have relatively low mutation rates, have an even distribution in the genome, and are relatively easy to detect. Although there is a high cost in developing SNPs detection assays (Batley et al. 2003; Kim and Misra 2007), the advancement in technology leading to specific low-density SNP chips and other affordable technologies, such as HRM, means that SNPs are being used more in plant studies (Gupta et al. 2001; Rafalski 2002; Barker and Edwards 2009). HRM is the most inexpensive, simple and rapid of these technologies for detecting SNPs (Gundry et al. 2003; Wittwer et al. 2003).

In HRM analysis, the polymerase chain reaction (PCR) amplification and melt analysis can be performed in one closed tube reaction (Montgomery et al. 2007) with no post-PCR analyses required (Ririe et al. 1997; Montgomery et al. 2007). The PCR products are differentiated from each other by melting curve profiles (Ririe et al. 1997) with heterozygous and homozygous genotypes being identified (Gundry et al. 2003; Wittwer et al. 2003; Reed and Wittwer 2004; Montgomery et al. 2007). The shape and position of the melting curves are due to the GC/AT ratio, length and sequence and can be used to differentiate PCR products (Ririe et al. 1997).

The main aim of marker-assisted backcrossing (MABC) is to introgress a trait of interest into the genetic background of a recipient genotype by recurrent backcrossing and also to recover the recurrent parent genome as rapidly and completely as possible. In each backcross generation there is variation around the mean of a trait which can be exploited to gain advantage in each generation. AFLPs are a universal, multi-locus marker technique that can be applied to genomes of any source or complexity and shows high levels of polymorphism, effective multiplex ratio, marker index and genotype index. A number of studies have used AFLPs in backcross plant breeding programmes to determine the percentage of recurrent parent genome, including cotton (Zhong et al. 2002), rice (Chen et al. 2000, 2001; Zhou et al. 2003; Gopalkrishnan et al. 2008), soybean (VanToai et al. 1996) and beans (Mũnoz et al. 2004).

A backcross breeding programme was developed at the African Centre for Crop Improvement (ACCI) in South Africa to introgress the *lpa1-1* gene into a wild-type tropical and subtropical adapted maize inbred line using molecular markers for both foreground and background selection. The objectives of this study were to use the *lpa1-1* SNP marker as foreground selection to detect homozygous recessive and heterozygote individual genotypes from the BC2F1 population using HRM analysis, and to use AFLP markers as background selection to determine the amount of recurrent parent genome (P16) present in each BC2F1 line.

Materials and methods

Germplasm

Two inbred maize lines were used in the backcross breeding programme to introgress the *lpa1-1* gene into tropical and subtropical adapted germplasm. The normal (wild-type) tropical locally adapted line P 16 and the temperate *lpa1-1* source CM 32 were used in this study. The normal line was crossed with the LPA line to produce the F1 generation. The F1 was backcrossed to the recurrent parent (P 16) to produce the BC1F1 generation. The BC1F1 generation was planted in pots in the greenhouse and backcrossed to the recurrent parent, with no selection for the *lpa1-1* gene, to generate the BC2F1 generation.

Crude DNA extraction

Two hundred and fifty BC2F1 progeny were grown in pots in a greenhouse and maize leaves were sampled for crude DNA extraction. A crude DNA extraction method of Edwards et al. (1991) was used to extract DNA for the *lpa1-1* SNP marker screening. Leaf samples for PCR analysis were collected using the lid of a sterile 1.5 ml microcentrifuge tube to punch out a disc of leaf material into the tube. The tissue was macerated with extraction buffer (200 mM Tris–HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% (w/v) SDS) and vortexed for a few seconds. The sample was heated for 10 min at 65°C and centrifuged for 2 min at 12,000*g*. The supernatant was removed, an equal volume of ice-cold isopropanol added and mixed before incubation at −20°C for 30 min. The sample was centrifuged for 5 min at 4°C at 12,000*g* and the pellet dried before re-suspending in 20 µl of TE (100 mM Tris–HCl, 1 mM EDTA, pH 8) buffer overnight.

PCR and HRM analysis

The Rotor-Gene 6000 real-time rotary analyser (Corbett Research, Australia) was used for the PCR amplification and HRM analysis. The PCR amplifications were performed in 20 µl reaction volumes consisting of approximately 15 ng of genomic DNA template, SensiMixdT for the PCR reaction components (Celtic Diagnostics, Cape Town), 1× SYBR Green I dye and 200 nM of forward and reverse PCR primers. The sequence of the forward and reverse primers for the SNP marker is listed in Table 1. The PCR amplification was initiated with a 10 min hold at 95°C as an initial denaturation step, followed by 40 cycles of 95°C for 10 s, 55°C for 15 s and 72°C for 20 s. A negative control was added in each set of PCR reactions with no DNA included to ensure non-contamination of PCR reagents. The HRM analysis was performed automatically after the PCR and

Table 1 Sequences of adapters and primers used for the AFLP process and SNP analysis

Name	Type	Sequence (5′–3′)
*Eco*RI-AAG	Primer + 3	GACTGCGTACCAATTCAAG
*Eco*RI-ACG	Primer + 3	GACTGCGTACCAATTCACG
*Eco*RI-ACC	Primer + 3	GACTGCGTACCAATTCACC
*Eco*RI-ACA	Primer + 3	GACTGCGTACCAATTCACA
*Mse*I-CAA	Primer + 3	GATGAGTCCTGAGTAACAA
*Mse*I-CAC	Primer + 3	GATGAGTCCTGAGTAACAC
*Mse*I-CAG	Primer + 3	GATGAGTCCTGAGTAACAG
*Mse*I-CAT	Primer + 3	GATGAGTCCTGAGTAACAT
*Eco*RI-A	Pre-amplification	AGACTGCGTACCAATTCA
*Mse*I-C	Pre-amplification	GACGATGAGTCCTGAGTAAC
*Eco*RI	Adapter	5′-CTCGTAGACTGCGTACC
		CATCTGACGCATGGTTAA-5′
*Mse*I	Adapter	5′-GACGATGAGTCCTGAG
		TACTCAGGACTCAT-5′
lpa1-1 SNP *fwd*	Forward primer	5′-ATAACTGGAGCGTGGGACAG-3′
lpa1-1 SNP *rev*	Reverse primer	5′-CTGCGGATGATCTTTTGGAT-3′

programmed to ramp temperature from 72 to 95°C, raised by 0.1 degree/step after the final extension step. The Rotor-Gene 6000 real-time rotary analyser created the melting curves and Rotor-Gene 6000 software version 1.7 (Corbett Research, Australia) was used to discriminate genotypes by difference plots.

Classification of BC2 lines

The samples were classified as heterozygous or homozygous recessive genotypes after analysis with the *lpa1-1* SNP marker by the R^2 values from the difference plots. The parental line, P 16, was used as the wild-type standard genotype, CM 32 as the LPA recessive standard genotype and the F1 cross between the two parental lines as the heterozygote genotype. The R^2 values were set at a threshold of 80%. The R^2 values provide an estimate of the confidence of the marker assay. Lines were selected for fingerprinting with AFLPs to determine the amount of recurrent parent present if $R^2 \geq 80\%$ for heterozygous and homozygous recessive genotypes. Their leaf material was sampled and ground in liquid nitrogen for high-quality genomic DNA extraction using Wizard genomic DNA purification kit (Promega, Whitehead Scientific, Cape Town, South Africa), as higher quality DNA is required for AFLP analysis. The DNA was screened on a 1% (w/v) agarose gel in 1× TAE buffer (40 mM Tris–acetate, 1 mM EDTA, pH 8.0) and visualized under UV light after staining with ethidium bromide (0.5 μg ml^{-1}).

AFLP process

The three main steps in the AFLP process included template preparation (restriction digestion of genomic DNA and ligation of oligonucleotide adaptors), fragment amplification (preselective and selective amplification) and identification of amplified products by capillary electrophoresis. The AFLP reactions were performed using the AFLP plant kit modules according to the manufacturer's recommendations (Applied Biosystems, South Africa).

Genomic DNA was double-digested by two restriction enzymes, *Eco*RI (rare cutter) and *Mse*I (frequent cutter). Double-stranded oligonucleotide adapters, with overhangs compatible with either 5′- or 3′-end generated fragments produced during the restriction digestion, were ligated to the DNA fragments. The ligated DNA fragments were pre-amplified by PCR using primers complementary to the adapter and restriction site sequence with only one additional selective nucleotide at their 3′ end (*Eco*-A, *Mse*-C). Ten microlitres of the pre-amplification reaction was screened by electrophoresis on a 1.5% (w/v) agarose gel in 1× TAE buffer (40 mM Tris–acetate, 1 mM EDTA, pH 8.0) and visualized under UV light after staining with ethidium bromide (0.5 μg ml^{-1}), to ensure quality and quantity of pre-amplification products. Selective amplification was performed with the use of selective primers with three nucleotides for *Eco*RI- and *Mse*I-specific primers according to the cycling conditions of the manufacturer. The sequences of the AFLP primers and SNP primers are listed in Table 1. The amplified fragments of the parental profiles were initially screened on the Rotor-Gene 6000. The 10 primer combinations showing the greatest differences between the parents were used for fingerprinting the BC lines. The pre-amplification and amplification PCR reactions for the BC2 lines and parents were performed in tubes and 96-well microtitre plates (Applied Biosystems, South Africa) on GeneAmp PCR system 2700 (Applied Biosystems) according to the manufacturer's instructions.

The amplified samples were sent to the Central Analytical Facility at the University of Stellenbosch (Stellenbosch, Cape Town, South Africa) for analysis of products on the 3100 DNA sequencer. A ROX 500 (Applied Biosystems) size standard was included in each sample for sizing of fragments amplified. The F filter set was used with a 36-cm capillary for the screening of the amplified products on the 3100 DNA sequencer.

Data analysis

There were only eight primer combinations selected for scoring due to incomplete data and poor amplification of the other two primer combinations. The eight *Eco*RI/*Mse*I primer combinations selected run files were viewed using Peak Scanner software v1.0 (Applied Biosystems, freeware, http://www.appliedbiosystems.com) and raw data tables of the product sizes exported to Microsoft Excel for manual scoring. The parental amplification reactions were repeated in each sequencer run and only consistent bands were

selected for scoring the individual BC2 lines. Bands were conservatively manually scored in the size range of 100–500 bp.

Similarity matrices and dendrograms were generated for each primer combination separately and the data combined for the six primer combinations. For the similarity matrices construction, bands were scored as present (1) or absent (0). The fragment data was converted to binary data by generating a binary matrix (0,1) using the AFLP banding patterns of each individual. NTSYS v2.1 software (Numerical Taxonomy and Multivariate Analysis for personal computers, Exeter Software, Setauket, NY, USA) was used to evaluate the genetic similarities between the maize BC2F1 lines and the parental lines (P 16 and CM 32). Pair-wise comparisons were made between the genotypes based on Dice similarity coefficients (Dice 1945). The resultant distance matrix data was used to construct a dendrogram using the agglomerative hierarchical un-weighted pair-group method with an arithmetic average (UPGMA) sub-program of NYSTS (Rohlf 1998).

Results and discussion

There were 250 BC2F1 plants screened with the co-dominant *lpa1-1* SNP marker, of which 28 plants were selected that carried at least one copy of the *lpa1-1* gene and were therefore selected for fingerprint analysis with AFLP markers. A much lower number of homozygous recessive and heterozygous plants than expected were identified (11%) using the SNP marker. This could be due to the single recessive *lpa1-1* gene and also that no selection for the *lpa1-1* gene was practiced in the BC1 generation before advancement to the BC2 generation. Selection in the BC1 generation would have provided a higher number of homozygous recessive and heterozygous genotypes in the BC2 generation. The *lpa1-1* SNP marker was able to detect homozygous dominant (wild-type/normal), homozygous recessive (LPA) and heterozygous genotypes by difference plots using HRM analysis (Fig. 1). The SNP marker was successfully validated by forward and reverse DNA sequencing in a previous study (Naidoo 2010).

In Fig. 1, the difference plot graph with LPA normalized and the normal parental inbred line (P 16), heterozygote and three BC2F1 lines is more informative than the melt profiles, as the shape of the BC2F1 lines are more clearly defined with BC2F1 line 260 very closely resembling the CM 32 baseline, and BC2F1 lines 246 and 257 resembling the heterozygote curve shape. There were 11 homozygous recessive (LPA) and 17 heterozygotes identified with the *lpa1-1* SNP marker according to their R^2 values and curve shape in difference plots. The heterozygotes showed a higher range of R^2 values of the *lpa1-1* SNP marker, from 90.95 to 99.59%, than the LPA recessives, which ranged from 82.81 to 99.58% (Table 2). The HRM technology has been successfully used in other studies to detect homozygous and heterozygous genotypes

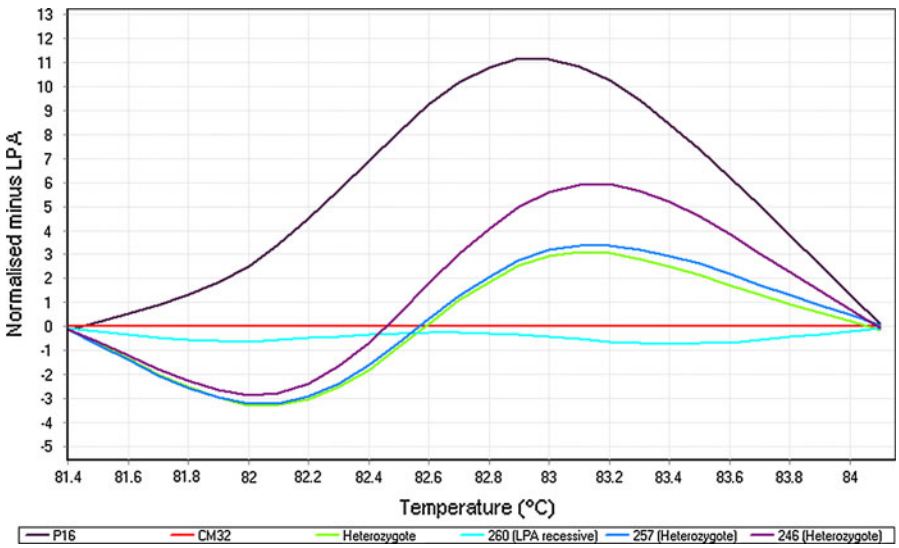

Fig. 1 High-resolution melt difference plots obtained using a SNP assay for maize *lpa1-1* gene in P 16 (recurrent parent), heterozygote (F1) and three of the selected BC2F1 lines (260, 257, 246) with CM 32 (LPA donor) normalized

Table 2 R^2 values of SNP marker analysis and similarity percentage to recurrent parent (P 16) for selected 28 BC2F1 individuals with six *Eco*RI/*Mse*I primer combinations of AFLP analysis

BC2F1 line	Genotype classification	R^2 (SNP) %	Similarity % (recurrent parent P 16)
119	Heterozygote	91.04	92.15
257	Heterozygote	99.55	90.65
235	LPA recessive	92.95	90.32
258	Heterozygote	96.55	89.80
246	Heterozygote	91.81	87.14
215	LPA recessive	90.15	86.54
264	LPA recessive	82.81	86.13
253	Heterozygote	96.04	85.79
117	LPA recessive	95.08	84.26
256	LPA recessive	97.20	83.94
136	LPA recessive	89.31	83.76
158	LPA recessive	93.12	83.37
271	Heterozygote	96.96	83.33
286	Heterozygote	93.77	82.07
176	Heterozygote	98.99	81.66
260	LPA recessive	99.58	81.05
138	LPA recessive	92.92	80.27
142	Heterozygote	91.89	80.09
252	LPA recessive	94.63	79.91
92	Heterozygote	91.31	79.82
242	Heterozygote	95.90	79.65
249	LPA recessive	92.88	79.19
78	Heterozygote	93.62	75.69
164	Heterozygote	99.59	73.73
145	Heterozygote	95.94	73.66
75	Heterozygote	98.60	72.90
265	Heterozygote	94.12	67.96
150	Heterozygote	90.95	62.12

based on melt curves and difference plots (Gundry et al. 2003; Wittwer et al. 2003; Reed and Wittwer 2004; Montgomery et al. 2007).

AFLP markers were chosen for this study over other types of molecular markers to determine the amount of recurrent parent genome in BC2 lines due to financial constraints and availability of primers and enzymes. In the initial AFLP screening of primer combinations on the two parental lines of the backcross breeding programme, there were ten *Eco*RI/*Mse*I primer combinations screened. Only eight *Eco*RI/*Mse*I primer combinations were selected for the 28 BC2F1 lines and the two parental lines of the backcross breeding programme. Due to incomplete data and poor amplification, two of the primer combinations were not included in the combined data set used for final analysis. The two primer combinations that were excluded from the combined data also showed very high similarity percentage values between P 16 and CM 32 (*Eco*-AAG *Mse*-CAT and *Eco*-ACC *Mse*-CAC) and lower similarity vales for the BC lines than the donor parent. Each individual primer combination was analysed separately and similarity values obtained.

For the individual primer combinations, similarity values were obtained for each of the 28 BC2F1 lines and CM 32 (Table 3). The primer combination *Eco*-ACA *Mse*-CAT showed BC line 75 as 0% similarity to the recurrent parent due to absence of similar bands between P 16 and 75. The amplification reaction of BC line 75 was repeated with the same results. It could be explained by the primer combination targeting an area that was dissimilar to the recurrent parent. Also with this primer combination, there were two other BC lines that showed 100% similarity to the recurrent parent, having amplified exactly the same bands as the recurrent parent.

For the combined data, there was a total of 277 data points/bands scored with 43 (16%) non-polymorphic and 234 (84%) polymorphic bands scored between the parental lines of the backcross breeding programme (Table 4). The number of bands scored in each primer combination ranged from 34 (*Eco*-ACA *Mse*-CAT) to 59 (*Eco*-AGG *Mse*-CAT). The primer combinations *Eco*-ACA *Mse*-CAA and *Eco*-ACA *Mse*-CAC showed the highest percentage of polymorphic bands (94%) with *Eco*-ACG *Mse*-CAA (90%) and *Eco*-ACG *Mse*-CAA (90%) and *Eco*-ACA *Mse*-CAT (88%) also showing high polymorphic rates. The primer combinations had a range of 6–28% of non-polymorphic bands and range of 72–94% of polymorphic bands. Positive polymorphic bands are defined as those present only in the recurrent parent (P 16) (coupling markers) since the objective is to determine the amount of recurrent parent genome present and whether negative polymorphic bands are present in the donor parent (CM 32) only. There was a total of 224 (81%) positive polymorphic bands in the combined data with 10 (4%) negative polymorphic bands (Table 4). All the primer combinations showed high positive polymorphic band percentages (>91%)

Table 3 Similarity percentage values to the recurrent parent (P 16) for the individual *Eco*RI/*Mse*I selective nucleotide primer combinations used for AFLP analysis of the 28 BC2F1 maize lines and CM 32 (LPA donor)

Line	*Eco*RI-AGG *Mse*I-CAA	*Eco*RI-AGG *Mse*I-CAT	*Eco*RI-AAG *Mse*I-CAT	*Eco*RI-ACA *Mse*I-CAA	*Eco*RI-ACA *Mse*I-CAC	*Eco*RI-ACA *Mse*I-CAT	*Eco*RI-ACC *Mse*I-CAC	*Eco*RI-ACG *Mse*I-CAA
CM32	28.26	27.12	70.49	5.56	6.00	11.76	37.50	7.69
75	84.78	67.80	54.10	55.56	42.00	0.00	58.33	78.85
78	76.09	69.49	60.66	58.33	50.00	55.88	66.67	57.69
92	71.74	74.58	62.30	52.78	60.00	55.88	75.00	78.85
117	80.43	71.19	67.21	63.89	60.00	88.24	79.17	78.85
119	97.83	88.14	75.41	72.22	74.00	94.12	62.50	88.46
136	80.43	76.27	47.54	61.11	54.00	88.24	87.50	76.92
138	84.78	69.49	55.74	50.00	50.00	82.35	79.17	71.15
142	69.57	72.88	50.82	66.67	54.00	76.47	70.83	69.23
145	63.04	54.24	47.54	55.56	50.00	79.41	62.50	59.62
150	30.43	52.54	47.54	47.22	34.00	58.82	50.00	53.85
158	78.26	72.88	62.30	58.33	58.00	94.12	50.00	75.00
164	63.04	57.63	37.70	47.22	56.00	73.53	37.50	57.69
176	78.26	71.19	57.38	63.89	48.00	85.29	66.67	75.00
215	84.78	77.97	59.02	75.00	64.00	97.06	75.00	67.31
235	91.30	76.27	63.93	69.44	74.00	94.12	79.17	92.31
242	84.78	76.27	54.10	47.22	28.00	97.06	70.83	71.15
246	86.96	74.58	49.18	77.78	66.00	100.00	75.00	69.23
249	78.26	64.41	65.57	58.33	50.00	91.18	79.17	63.46
252	78.26	71.19	62.30	58.33	50.00	79.41	70.83	65.38
253	82.61	77.97	62.30	55.56	0.00	91.18	79.17	71.15
256	78.26	77.97	4.92	61.11	64.00	91.18	79.17	67.31
257	93.48	81.36	4.92	72.22	66.00	100.00	83.33	90.38
258	97.83	77.97	4.92	61.11	72.00	97.06	83.33	86.54
260	80.43	76.27	4.92	61.11	40.00	97.06	45.83	63.46
264	86.96	76.27	4.92	69.44	64.00	79.41	54.17	80.77
265	78.26	54.24	34.43	41.67	24.00	58.82	4.17	57.69
271	91.30	69.49	67.21	63.89	48.00	94.12	4.17	71.15
286	80.43	77.97	62.30	50.00	54.00	85.29	79.17	71.15

showing that the recurrent parent was highly polymorphic and generally amplified more fragments/bands compared to the donor parent.

There are numerous genetic diversity studies with AFLP markers on maize using different enzyme primer combinations showing differing polymorphic rates, average number of markers per primer combination and the range of number of markers amplified in each primer combination. The AFLP polymorphism rate differs due to the number of primer combinations used, the number of polymorphic bands produced and the genetic background and number of inbred maize lines used. Examples of results using different *Eco*RI/*Mse*I primer combinations and bands produced include 284 polymorphic bands with 10 primer pairs (Hartings et al. 2008), 621 (59.8%) polymorphic bands with 16 primer combinations (Wu 2000), 209 polymorphic bands (41.8%) with six primer combinations (Ajmone-Marsan et al. 1998), 261 polymorphic bands (83%) with four primer combinations (Lübberstedt et al. 2000), 232 polymorphic bands with six primer combinations (Pejic et al. 1998) and 408 (81.7%) polymorphic bands with seven primer combinations (Legesse et al. 2007). These studies used more than two inbred lines at a time and none of these studies compared temperate × tropical cross material. This

Table 4 Levels of polymorphism and non-polymorphism among six *Eco*RI/*Mse*I selective nucleotide primer combinations used for AFLP analysis of the maize BC2F1 lines

Primer combination	Scored amplicon No.	Non-polymorphic amplicon No. (%)	Polymorphic amplicon No. (%)	Positive polymorphic amplicon No. (%)	Negative polymorphic amplicon No. (%)
*Eco*RI-AGG *Mse*I-CAT	59	16 (27)	43 (73)	39 (91)	4 (9)
*Eco*RI-AGG *Mse*I-CAA	46	13 (28)	33 (72)	32 (97)	1 (3)
*Eco*RI-ACG *Mse*I-CAA	52	5 (10)	47 (90)	45 (96)	2 (4)
*Eco*RI-ACA *Mse*I-CAT	34	4 (12)	30 (88)	30 (100)	0 (0)
*Eco*RI-ACA *Mse*I-CAA	36	2 (6)	34 (94)	33 (97)	1 (3)
*Eco*RI-ACA *Mse*I-CAC	50	3 (6)	47 (94)	45 (96)	2 (4)
Overall total	277	43 (16)	234 (84)	224 (81)	10 (4)
Overall mean	46.2	7.2	39.0	37.3	1.7

study found 84% polymorphic bands with only six primer combinations tested, which is relatively high considering only two inbred lines were compared. The type of restriction enzyme influences the polymorphism rate, as *Eco*RI restriction sites are randomly located over the chromosome while *Pst*I is clustered at the methylation-specific sites, and therefore *Pst*I primer combinations show a lower polymorphism ratio (Lübberstedt et al. 2000).

Both polymorphic and non-polymorphic bands were selected for data points to determine the degree of similarity between the parental lines. The *lpa1-1* source CM 32 showed a baseline similarity of 26.33% to P 16, the recurrent parent with the combined data (unpublished data). This is to be expected as they are both maize inbred lines although from different pedigrees and adaptation (tropical and temperate).

All the BC2F1 selected lines were clearly differentiated from each other and the parental controls in the combined data dendrogram, based on the Dice coefficient of similarity (Fig. 2). The combined similarity matrix generated by NTSYS was able to successfully differentiate between all 28 BC2F1 lines and donor parental line based on the similarity values to the recurrent parent (data not shown). It is generally expected in a BC2 generation to find 87.5% similarity to the recurrent parent. However, in this study, the similarity percentage values for the heterozygotes ranged from 62.12 to 92.15%, while the LPA homozygous recessives ranged from 79.19 to 90.32% (Table 2). This validates the variation around the mean that can be exploited to increase the genetic gain in each BC generation and reduce the number of BC generations required to reach a high percentage of the recurrent parent. The BC lines were very similar to the recurrent parent and no groups were identified in the dendrogram due to the close clustering seen between the lines and parental lines.

For the combined data of all six primer combinations, there were 13 BC2F1 lines: seven LPA recessives [235 (90.32%), 215 (86.54%), 264 (86.13%), 117 (84.26%), 256 (83.94%), 136 (83.76%). 158 (83.37%)] and six heterozygous lines [119 (92.15%), 257 (90.65%), 258 (89.80), 246 (87.14%), 253 (85.79%), 271 (83.33%)] showing the highest similarity percentages and therefore containing the highest percentage of the recurrent parent genome. These can be selected for further improvement in the breeding programme.

According to the clustering on the dendrogram, CM 32 was closely related to three BC2F1 lines, 150 (38%), 265 (32%) and 75 (27%), therefore these lines have the lowest percentage of recurrent parent genome (P 16). Due to the high level of similarity between the BC2F1 lines, it was not possible to group the lines in the combined data dendrogram (Fig. 2). This is due to the BC2F1 generation which would have a high percentage of the recurrent parent genome (expected 87.5%) by this stage.

Studies on rice using AFLPs for background selection in the BC1F1 generation produced 364 polymorphic bands to obtain 81.4% of the recurrent parent genome (Jeung et al. 2005), while a BC3F2 population had 98.8% recurrent parent recovery with 129 polymorphic bands (Chen et al. 2001), and selection in the BC3F1 generation resulted in 84.2–100% recovery of the recurrent parent with 118 polymorphic bands (Zhou et al. 2003). In this study for

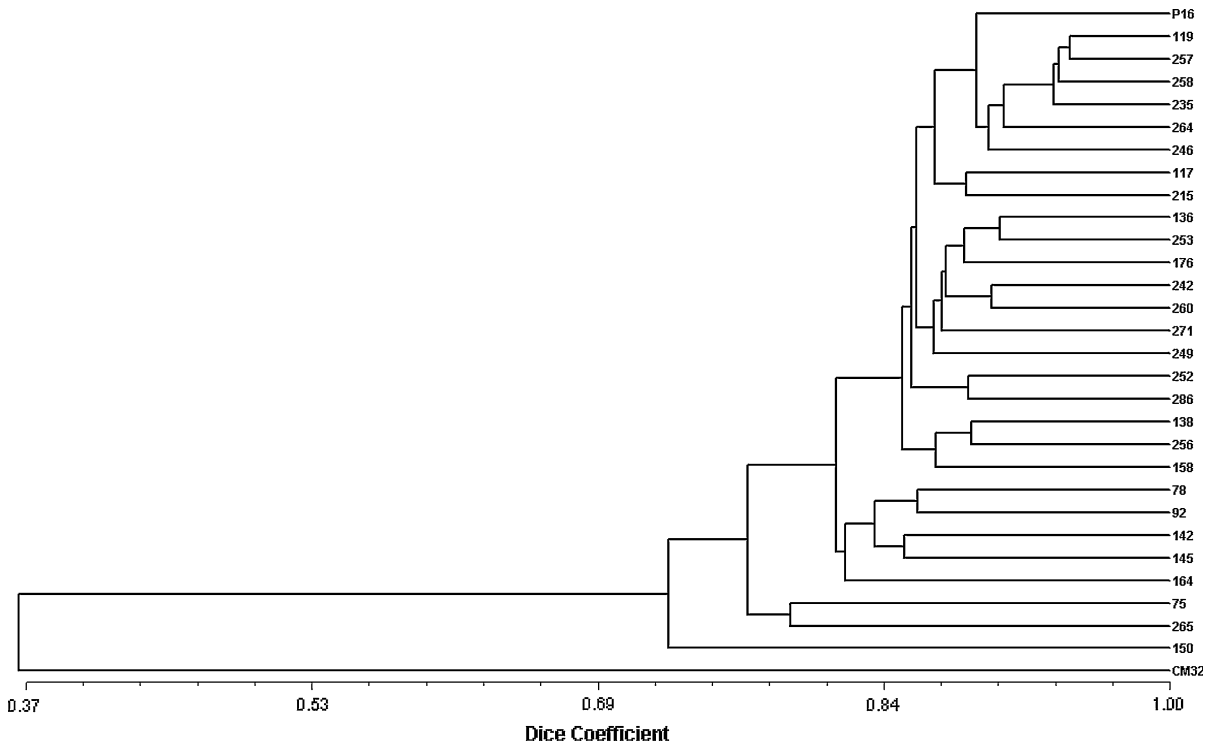

Fig. 2 Dendrogram of fingerprinting 28 BC2F1 maize lines and parental lines, P 16 (recurrent parent) and CM 32 (LPA donor) based on six primer combinations of *Eco*RI/*Mse*I AFLP data generated using NTSYS v2.1 and Dice coefficient of similarity

introgression of the *lpa1-1* gene into the recurrent parent, selection was in the BC2F1 generation with 234 polymorphic bands resulting in a maximum recovery of 92.15% of the recurrent parent genome. This was higher than the expected BC2 recovery of 87.5% in the second backcross generation.

Selection of suitable BC2F1 lines would be in the range of high R^2 SNP values with high similarity percentage values. The choice of heterozygous or homozygous recessive lines would also be part of the selection process, with homozygous recessive lines being preferred due to both alleles of the *lpa1-1* gene being present and fixed. All 28 BC2F1 lines will have to undergo further testing under field conditions to determine yield potential, adaptability to local environmental conditions and disease assessment.

A promising candidate for MAS for LPA mutant detection is the use of SNP markers accompanied by a high-throughput analysis method such as HRM analysis. This will lead to an improved effective breeding programme enhanced with MAS to transfer genes of interest and accelerate the recovery of the elite parent genome in backcross breeding programmes.

Conclusions

The backcross breeding programme to introgress the *lpa1-1* gene into tropical and subtropical adapted inbred line (P 16) was successful with both foreground and background selection phases with molecular markers. The foreground selection for the *lpa1-1* gene was efficient and rapid with use of the *lpa1-1* SNP marker with HRM analysis for detection of the different alleles of the gene. Due to the co-dominant nature of the *lpa1-1* SNP, it was possible to correctly identify the homozygous dominant (wild-type), homozygous recessive (LPA) and heterozygote lines of the BC2F1 population. The marker assay with the SNP marker and HRM proved to be effective in identifying all genotypes and would be less time-consuming than the traditional approach of detecting the *lpa1-1* gene. There were 250 BC2F1 lines screened for the *lpa1-1* gene with 11 homozygous recessive (LPA) and 17 heterozygous lines being selected for background selection using AFLP markers. The AFLP technique used six *Eco*RI/*Mse*I primer combinations to produce 84% of polymorphic bands between the parental

inbred lines. The percentage of recurrent parent genome recovered ranged from 62.12 to 92.15%. There were 13 lines showing >83% of the recurrent parent genome and these are recommended for further advancement and field assessment.

Acknowledgments The study was funded by a grant from The Rockefeller Foundation/Alliance for a Green Revolution in Africa. Additional funding was obtained from the Generation Challenge Programme (GCP).

References

Abalo G, Tongoona P, Derera J, Edema R (2009) A comparative analysis of conventional and marker-assisted selection methods in breeding maize streak virus resistance in maize. Crop Sci 49:509–520

Adams CL, Hambidge M, Raboy V, Dorsch JA, Sian L, Westcott JL, Krebs NF (2002) Zinc absorption from a low-phytic acid maize1-3. Am J Clin Nutr 76:556–559

Ajmone-Marsan P, Casiglioni P, Fusari F, Kuiper M, Motto M (1998) Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor Appl Genet 96:219–227

Barker GLA, Edwards KJ (2009) A genome-wide analysis of single nucleotide polymorphism diversity in the world's major cereal crops. Plant Biotechnol J 7:318–325

Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D (2003) Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol 132:84–91

Bohn T, Davidsson L, Walczyk T, Hurrell RF (2004) Phytic acid added to white-wheat bread inhibits fractional magnesium absorption in humans. Am J Clin Nutr 79:418–423

Chen S, Xu Lin XH, Xu CG, Zhang Q (2000) Improvement of bacterial blight resistance of 'Minghui 63', an elite restorer line of hybrid rice, by molecular marker-assisted selection. Crop Sci 40:239–244

Chen S, Xu CG, Lin XH, Zhang Q (2001) Improving bacterial blight resistance of '6078', an elite restorer line of hybrid rice, by molecular marker-assisted selection. Plant Breed 120:133–137

Dice LR (1945) Measures of the amount of ecological association between species. Ecology 26:297–302

Douglas MW, Peter CM, Boling SD, Parsons CM, Baker DH (2000) Nutritional evaluation of low phytate and high protein corns. Poult Sci 79:1586–1591

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res 19:1349

Ertl DS, Young KA, Raboy V (1998) Plant genetic approaches to phosphorus management in agricultural production. J Exp Qual 27:299–304

Gopalkrishnan S, Sharma RK, Rajkumar KA, Joseph M, Singh VP, Singh AK, Bhat KV, Singh NV, Mohapatra T (2008) Integrating marker assisted background analysis with foreground selection for identification of superior bacterial blight resistance recombinants in Basmati rice. Plant Breed 127:131–139

Gundry CN, Vandersteen JG, Reed GH, Pryor R, Chen J, Wittwer CT (2003) Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. Clin Chem 49:396–406

Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr Sci 80:524–535

Hambidge KM, Huffer JW, Raboy V, Grunwald GK, Westcott JL, Sian L, Miller LV, Dorsch JA, Krebs NF (2004) Zinc absorption from low-phytate hybrids of maize and their wild-type isohybrids. Am J Clin Nutr 79:1053–1059

Hambidge KM, Krebs N, Westcott JL, Sian L, Miller LV, Peterson KL, Raboy V (2005) Absorption of calcium from tortilla meals prepared from low-phytate maize. Am J Clin Nutr 82:84–87

Hartings H, Berardo N, Mazzinelli GF, Valoti P, Verderio A, Motto M (2008) Assessment of genetic diversity and relationships among maize (*Zea mays* L.) Italian landraces by morphological traits and AFLP profiling. Theor Appl Genet 117:831–842

Hurrell RF, Reddy MB, Juillerat M-A, Cook JD (2003) Degradation of phytic acid in cereal porridges improves iron absorption by human subjects. Am J Clin Nutr 77:1213–1219

Jeung JU, Hwang HG, Moon HP, Jena KK (2005) Fingerprinting temperate *japonica* and tropical *indica* rice genotypes by comparative analysis of DNA markers. Euphytica 146:239–251

Kim S, Misra A (2007) SNP genotyping: technologies and biomedical applications. Annu Rev Biomed Eng 9:289–320

Legesse BW, Myburg AA, Pixley K, Twumasi-Afriyie S, Botha AM (2007) Genetic diversity of maize inbred lines revealed by AFLP markers. Afr Crop Sci Conf Proc 8:649–653

Li YC, Ledoux DR, Veum TL, Raboy V, Ertl DS (2000) Effects of low phytic acid corn on phosphorus utilisation, performance, and bone mineralisation in broiler chicks. Poult Sci 79:1444–1450

Lübberstedt T, Melchinger AE, Dußle C, Vuylsteke M, Kuiper M (2000) Relationships among early European maize inbreds: IV. Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD, and pedigree data. Crop Sci 40:783–791

Mendoza C, Viteri FE, Lönnerdal B, Young KA, Raboy V, Brown KH (1998) Effect of genetically modified, low-phytic maize on absorption of iron from tortillas. Am J Clin Nutr 68:1123–1127

Mendoza C, Viteri FE, Lönnerdal B, Raboy V, Young KA, Brown KH (2001) Absorption of iron from unmodified maize and genetically altered, low phytate maize fortified with ferrous sulphate or sodium iron EDTA. Am J Clin Nutr 73:80–85

Montgomery J, Wittwer CT, Palais R, Zhou L (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nat Protoc 2:59–66

Mũnoz LC, Blair MW, Duque MC, Tohme J, Roca W (2004) Introgression in common bean x tepary bean interspecific

congruity-backcross lines as measured by AFLP markers. Crop Sci 44:637–645

Naidoo R (2010) Development and application of SNP marker for low phytic acid gene (lpa1-1) with studies on the effect of low phytic acid on seed germination, vigour and yield in maize. PhD thesis. University of KwaZulu-Natal, Pietermaritzburg, South Africa

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. Theor Appl Genet 97:1248–1255

Peter CM, Baker DH (2002) Bioavailability of phosphorus in corn gluten feed derived from conventional and low-phytate maize. Anim Feed Sci Technol 95:63–71

Raboy V (2000) Low-phytic-acid grains. Food Nutr Bull 21(4):423–427

Raboy V, Gerbasi PF, Young KA, Stoneberg SD, Pickett SG, Bauman AT, Murthy PPN, Sheridan WF, Ertl DS (2000) Origin and seed phenotype of maize low phytic acid 1–1 and low phytic acid 2–1. Plant Physiol 124:355–368

Rafalski JA (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

Reed GH, Wittwer CT (2004) Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. Clin Chem 50:1748–1754

Ririe KM, Rasmussen RP, Wittwer CT (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal Biochem 245:154–160

Rohlf FJ (1998) NTSYSpc numerical taxonomy and multivariate analysis system version 2.0. User guide. Applied Biostatistics Inc.

Shi J, Wang H, Schellin K, Li B, Faller M, Stoop JM, Meeley RB, Ertl DS, Ranch JP, Glassman K (2007) Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. Nat Biotechnol 25:930–937

Spencer JD, Allee GL, Sauber TE (2000a) Phosphorus bioavailability and digestibility of normal and genetically modified low-phytate corn for pigs. J Anim Sci 78:675–681

Spencer JD, Allee GL, Sauber TE (2000b) Growing-finishing performance and carcass characteristics of pigs fed normal and genetically modified low-phytate corn. J Anim Sci 78:1529–1536

VanToai TT, Peng J, St. Martin S (1996) Using AFLP markers to determine the contribution of parental genomes during recurrent selection. Soybean Genet Newsl 23:214–216

Veum TL, Ledoux DR, Raboy V, Ertl DS (2001) Low-phytic acid corn improves nutrient utilization for growing pigs. J Anim Sci 79:2873–2880

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High resolution genotyping by amplicon melting analysis using LCGreen. Clin Chem 49:853–860

Wu M (2000) Genetic diversity and its relationship to hybrid performance and heterosis in maize as revealed by AFLPs and RAPDs. Maize Genet Coop News 74:62–63

Yan F, Kersey JH, Fritts CA, Waldroup PW, Stilborn HL, Crum RC Jr, Rice DW, Raboy V (2000) Evaluation of normal yellow dent corn and high available phosphorus corn in combination with reduced dietary phosphorus and phytase supplementation for broilers grown to market weights in litter pens. Poult Sci 79:1282–1289

Yan F, Fritts CA, Waldroup PW, Stilborn HL, Rice DW, Crum RC Jr, Raboy V (2003) Comparison of normal and high available phosphorus corn with and without phytase supplementation in diets for male large white turkeys grown to market weights. Int J Poult Sci 2:83–90

Zhong M, McCarty JC, Jenkins JN, Saha S (2002) Assessment of day-neutral backcross populations of cotton using AFLP markers. J Cotton Sci 6:97–103

Zhou PH, Tan YF, He YQ, Xu CG, Zhang Q (2003) Simultaneous improvement for four quality traits of Zhenshan 97, an elite parent of hybrid rice, by molecular marker-assisted selection. Theor Appl Genet 106:326–331