



Contents lists available at ScienceDirect

## Electronic Journal of Biotechnology

journal homepage:



## Research Article

## Exon based amplified polymorphism (EBAP): A novel and universal molecular marker for plants



Faqian Xiong<sup>a,b</sup>, Jing Liu<sup>b</sup>, Ronghua Tang<sup>b</sup>, Taiyi Yang<sup>b</sup>, Xinghai Yang<sup>d</sup>, Liangqiong He<sup>b</sup>, Zhuqiang Han<sup>b</sup>, Lihang Qiu<sup>c</sup>, Chenglin Zou<sup>b</sup>, Xiumei Tang<sup>b</sup>, Cong Luo<sup>a</sup>, Ruichun Zhong<sup>b</sup>, Jing Jiang<sup>b</sup>, Zhipeng Huang<sup>b</sup>, Haining Wu<sup>b</sup>, Junxian Liu<sup>c,\*</sup>, Xinhua He<sup>a,\*</sup>

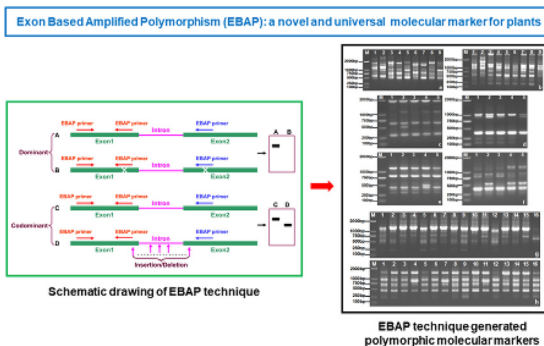
<sup>a</sup> College of Agriculture, State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangxi University, Nanning 530004, China

<sup>b</sup> Cash Crops Research Institute, Guangxi Academy of Agricultural Sciences, Nanning 530007, China

<sup>c</sup> Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences/Key Laboratory of Sugarcane Biotechnology and Genetic Improvement (Guangxi), Ministry of Agriculture and Rural Affairs/Guangxi Key Laboratory of Sugarcane Genetic Improvement, Nanning, 530007, China

<sup>d</sup> Rice Research Institute, Guangxi Academy of Agricultural Sciences, Nanning 530007, China

## G R A P H I C A L A B S T R A C T



## A R T I C L E I N F O

## Article history:

Received 25 May 2021

Accepted 6 January 2022

Available online 12 January 2022

## Keywords:

Exon based amplified polymorphism (EBAP)

Molecular marker

Exon

Fingerprinting

## A B S T R A C T

**Background:** Exons are an important part of genes. However, there has been no molecular marker technique based on single primer amplification developed for exons region.

**Results:** A novel molecular marker technique called exon based amplified polymorphism (EBAP) which used single primer to perform the PCR was developed. Single primers were designed based on the rich GC bases in the exon region of genes. The single primer consists of filling sequence, intermediate core region sequence and selective bases sequence. The filling sequence has a total of six bases, which can be 2A and 2T or 3A and 3T. The intermediate core region sequence consists of eight bases, all of which are G and C, but at least 2G or 2C are required. The first of the three selective bases must be either A or T, and the second and third are arbitrary bases. Due to the principle of primer design, these single primers should be universal. We applied it to the DNA polymorphisms detection of maize, sugarcane, potato, cassava, cabbage, and peanut. The results showed that it detected more abundant DNA polymorphisms in the first five crops, but not in the cultivated peanut. In addition, different band patterns were obtained by amplifying peanut cDNA with different single primers.

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso

\* Corresponding authors.

E-mail addresses: xfq2002@126.com (F. Xiong), liujunxian868@163.com (J. Liu), honest66222@163.com (X. He).

<https://doi.org/10.1016/j.ejbt.2022.01.001>

0717-3458/© 2022 Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Conclusions:** We have developed a novel and universal molecular marker technique. It is simple, fast and efficient. It does not require genomic sequence information. It can be widely used in germplasm identification, genetic diversity analysis, and molecular fingerprinting.

**How to cite:** Xiong F, Liu J, Tang R, et al. Exon based amplified polymorphism (EBAP): A novel and universal molecular marker for plants. *Electron J Biotechnol* 2022;56. <https://doi.org/10.1016/j.ejbt.2022.01.001>. © 2022 Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Molecular markers can reveal the differences between individuals. They can directly reflect the genetic nature and genetic variation of organisms at the DNA level. Because of the advantages of abundant quantity, a large amount of information, high polymorphisms, simple and rapid detection methods that are not affected by seasons, environment and individual development stages, they have been widely used in genetic diversity analysis, molecular fingerprinting, genetic linkage map construction, gene/QTL mapping, and molecular marker-assisted breeding [1].

Since the concept of molecular marker was first put forward in 1980, especially since 1990, molecular marker techniques have developed rapidly. Currently, the most widely applied molecular marker techniques include restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), inter-simple sequence repeats (ISSR), amplified fragment length polymorphism (AFLP), simple sequence repeat (SSR), sequence-related amplified polymorphism (SRAP), and start codon targeted polymorphism (SCoT) [2,3,4,5,6,7,8]. According to the different detection methods used, they can be divided into molecular marker techniques based on molecular hybridization technology and molecular marker techniques based on PCR technology. According to the different properties of the primers used, they can be partitioned into random primer based amplification and specific primer pair based amplification. According to the number of primers used, they can be delimited into those based on single primer amplification and those based on double primer amplification [1]. According to the different functions of the generated molecular markers, they can be classified into random type, gene-targeted type and functional type [9].

Conventional molecular marker techniques use primer pair for PCR amplification, while a large group only use single primer for PCR amplification (Table 1). Single primer amplification means the use of only one primer in a PCR reaction and the single primer acts as both upstream and downstream primers. When the target location distances between the two binding sites of single primer and the time allowed for the extension step were appropriate, effective amplification can be achieved. Since 1990s, molecular marker techniques based on single primer amplification have been continuously developed. Primers design are the core of these molecular marker techniques. Single primers used in RAPD were designed completely random [3]. Single primers of DAMD, ISSR and URP were designed based on the core sequences of minisatellite DNA, simple repeat sequence, and pKRD with high copy respectively [4,10,11]. Single primers of IRAP, IMP and iPBS were designed according to the conserved LTR sequences of LTR retro-

transposons, the TIR sequences of MITE transposons, and the conservative tRNA binding sites adjacent to the 5'-LTR sequences in the LTR retrotransposons respectively [12,13,14]. Single primers of ISJ, SCoT, BPS, and CDBP were designed according to the conserved sequences of intron–exon splicing sites, the conserved sequences flanking the translation initiation sites of plant genes, the short conservative and consistent branch point signal sequences, and the CAAT box sequences in the promoters of plant genes respectively [8,15,16,17]. Single primers of CDDP and HFO-TAG were designed based on the sequences of protein families in plant genomes that have common conservative consistency and a large number of high frequency short sequences with a length of 8~10 bp existing in watermelon unigenes [18,19].

Exons are an important part of genes in eukaryotic genomes. They are rich in GC bases and have coding function [7]. However, up to now, there has been no molecular marker technique based on single primer amplification developed for exons region. In this study, single primers were designed according to the GC bases that are rich in the exons region of genes in eukaryotic genomes. Using single primer for PCR amplification, a novel molecular marker technique termed exon based amplified polymorphism was developed. Then it was applied to the DNA polymorphisms detection and genetic diversity analysis of maize, sugarcane, potato, cassava, cabbage, and peanut.

## 2. Materials and methods

### 2.1. Plant materials and DNA/RNA extraction

Six crops including maize, sugarcane, potato, cassava, cabbage, and peanut were used in this study (Table 2). Of sixteen accessions of genus *Arachis*, No. 1~12 represent cultivated peanut, No. 13~16 represent wild peanut of section *Arachis*. Genomic DNA was extracted using the established modified CTAB method [20]. The purity and concentration of DNA were detected by 1.2% agarose gel electrophoresis and micro-spectrophotometer (NanoDrop) respectively. Finally, DNA concentrations were uniformly adjusted to 50 ng/μl and stored under –20°C condition. The established mCTAB-dLiCl method was used to extract the RNA from peanut leaves. 6 μl (about 200 μg) RNA and 2 μl Oligo(dT)<sub>18</sub> were used as template and reverse transcription primer respectively. They were incubated at 70°C for 10 min, then quickly cooled on ice for more than 2 min, and centrifuged for several seconds to make the denatured solution of RNA template and primer gather at the bottom of the tube. Then 4 μl of 5 × M-MLV buffer, 0.5 μl of dNTP (10 mmol/l), 0.25 μl of RNase Inhibitor (40 U/μl), and 0.25 μl of RTase M-MLV (RNase H-) were added to the above solution. The

**Table 1**

The summarized PCR-based molecular marker techniques.

Type	Reference	Type	Reference	Type	Reference	Type	Reference
RAPD	Williams et al. [3]	ISSR	Zietkiewicz et al. [4]	AFLP	Vos et al. [5]	SSR	Tautz and Renz [6]
SRAP	Li and Quiros [7]	SCoT	Collard and Mackill [8]	DAMD	Heath et al. [10]	URP	Kang et al. [11]
IRAP	Kalendar et al. [12]	IMP	Chang et al. [13]	iPBS	Kalendar et al. [14]	ISJ	Weining and Langridge [15]
BPS	Xiong et al. [16]	CDBP	Singh et al. [17]	CDDP	Collard and Mackill [18]	HFO-TAG	Levi et al. [19]

**Table 2**  
Six crops used in the present study.

Code	Variety	Code	Variety
<b>Maize</b>			
1	Zhaofeng788	6	Guidan589
2	Guidan688	7	Yunrui88
3	Guidan591	8	Yumeitou105
4	Zhaofeng688	9	Nanxiao9665
5	Guidan0810	10	Guidan901
<b>Sugarcane</b>			
1	Guitang21	5	Guitang32
2	ROC22	6	Guitang34
3	Guitang30	7	Guitang35
4	Guitang31	8	Badila
<b>Potato</b>			
1	Zhongshu 11 hao	6	Jizhangshu 8 hao
2	Zhongshu 7 hao	7	Zhongshu 13 hao
3	Zhongshu 14 hao	8	Zhongshu 6 hao
4	FAVORITA	9	Zhongshu 17 hao
5	Zhongshu 8 hao		
<b>Cassava</b>			
1	Huanan205	4	Xinxuan048
2	Huanan 9 hao	5	Col22
3	Nanzhi199		
<b>Cabbage</b>			
1	Teqing 60tian cutiao cabbage	9	Lvbao701
2	Australia008 quannian youlv sweet cabbage	10	3-1cabbage
3	Australia super608	11	3-6-1cabbage
4	Hongkong 45tian youqing sweet cabbage	12	5-3cabbage
5	Mingyou308 chaoguan sweet cabbage king	13	5-4cabbage
6	Australia50 duanbing cutiao cabbage	14	5-14cabbage
7	Guiliu October liuye cabbage	15	6-6cabbage
8	Caixingli guoji xinyidai bilvyou cabbage49,	16	6-6-1cabbage
<b>Peanut</b>			
1	Lengshui damake	9	Guihua17
2	Yaoshang xiaomake	10	Guihua771
3	Lingui makechangyao	11	Guihua836
4	Pinglezi	12	Guihua1026
5	Lipu fanguidou	13	<i>A.monticola</i> (PI468199)
6	Ziyuan huasheng	14	<i>A.duranensis</i> (PI262133)
7	Quanxian fanguidou	15	<i>A.duranensis</i> (PI219823)
8	Quanzhou fankezi	16	<i>A.ipaensis</i> (PI468322)

first strand of cDNA was obtained by keeping the above mixed solution at 42°C for 1 h and 70°C for 15 min and then cooled on ice. The first strand of cDNA was diluted quantitatively (50.0 ng/μl) and used as PCR template [21].

## 2.2. Primers design

According to the fact that exons are normally rich in GC-rich regions, single primer of EBAP technique was designed to target exons to open reading frame (ORF) regions [7] (Fig. 1). The length of single primer is 17 bp and has no merged bases. It consists of the filling sequence, the intermediate core region sequence and the selective bases sequence from the 5' to the 3'. For the filling sequence, the length is six bases, the AT content is 66.67% or more, the ratio of A to T is 1:1. For the intermediate core region sequence, the total length is eight bases, the GC content is 100%, the ratio of G to C is 1:3 or more. The last three selective bases at the 3' require that the first base must be A or T, and the second and third bases can be arbitrary bases (Table 3).

## 2.3. PCR amplification and electrophoresis

The EBAP-PCR amplification reaction was carried out in a 20 μl volume containing 200 μM of each dNTP, 1.5 mM MgCl<sub>2</sub>, 1 U Taq DNA polymerase, 10 pmol single primer, 1×PCR buffer, and 50 ng DNA/cDNA. The EBAP-PCR amplification procedure was performed as follows: initial denaturation at 94 °C for 4 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 1 min, and extension at 72 °C for 1 min 30 s, with a final extension at 72 °C for 10 min. It should be noted that the annealing temperature should not be lower than 50 °C to reduce the specificity and repeatability of PCR amplification. After amplification, 6 μl loading buffer was added to the PCR products, then 5 μl PCR products were separated by 1.5% agarose gel electrophoresis (1 × TAE buffer). After electrophoresis, the agarose gel was stained with ethidium bromide (0.5 μg/ml) for 10 min and photographed on the UV gel imaging system. The repeatability of PCR was checked by at least two PCRs. Only clear and reproducible bands were used for statistical analysis.

## 2.4. Band scoring and data analysis

EBAP is regarded as a multi-loci dominant molecular marker technique. And each amplified band is regarded as single locus. Only the clearly distinguishable bands were counted with band marked as 1 and no band marked as 0, thus forming a binary data matrix. According to the binary data matrix, total number of bands (TNB), number of polymorphic bands (NPB) and percentage of polymorphic bands (PPB) were counted. Genetic similarity was calculated by the formula  $GS_{ij} = 2N_{ij}/(N_i + N_j)$ , wherein  $N_{ij}$  was the number of common amplified bands between accessions i and j, and  $N_i$  and  $N_j$  were the number of amplified bands in accessions i and j respectively. Genetic distance was calculated by the formula  $GD_{ij} = 1 - GS_{ij}$ . Polymorphic index (PI) was calculated by the previous reported method [22]. Observed number of alleles ( $N_a$ ), effective number of alleles ( $N_e$ ), Nei's gene diversity (H) and Shannon's information index (I) were calculated using POPGENE32. Resolving power was calculated by the formula  $R_p = \sum I_b$ , where  $I_b = 1 - (2 \times |0.5 - p|)$ , where p is the portion of the samples containing the observed band [23]. Effective multiplex ratio (EMR), that is, the average number of polymorphic bands is equal to the number of polymorphic bands generated by the amplification of all polymorphic primers divided by the number of used polymorphic primers [24]. Polymorphism information content (PIC) was calculated by the formula  $PIC = 1 - \sum P_{ij}^2$ , where  $P_{ij}$  represents the frequency of the i-th site that appears in the j-th gene. Marker index (MI), a parameter that can reflect the overall efficiency of a molecular marker technique or single primer, was calculated by multiplying the effective multiple ratio by the average polymorphism information content or the polymorphism information content of single primer [24].  $R_p$ , EMR, PIC and MI were automatically calculated by our new developed software (unpublished). Clustering analysis was conducted by MEGA6.06 using the genetic distance matrix computed from DPS software.

## 3. Results

### 3.1. EBAP markers in maize and sugarcane (Monocotyledons)

For ten maize and eight sugarcane varieties, thirty-three and fifteen single primers were screened, of which eighteen and ten single primers could successfully amplify clear and repeatable bands. A total of one hundred and thirteen and eighty-seven bands were amplified, of which sixty-four and fifty-seven were polymorphic bands, accounting for 56.64% and 65.52% of the total bands. The

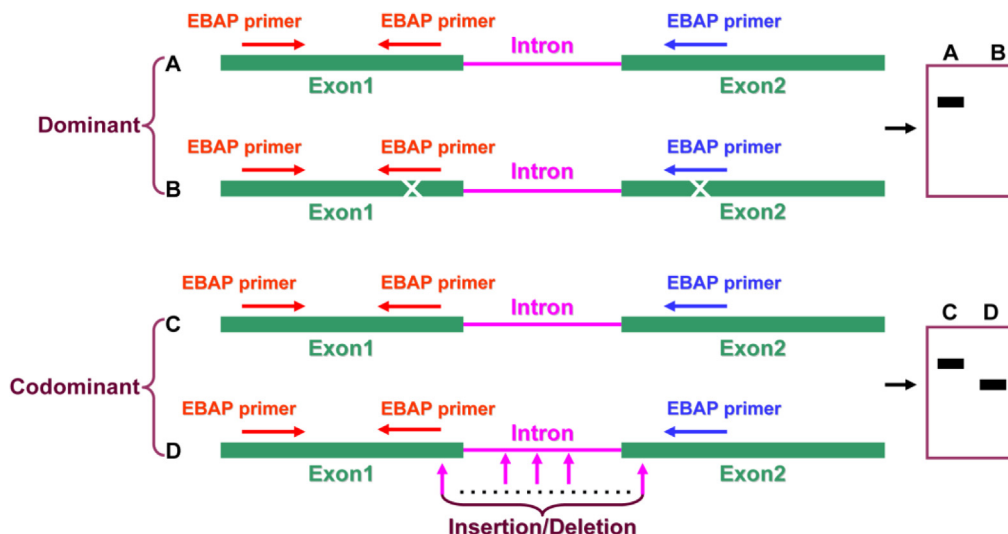


Fig. 1. The schematic drawing describes the principle of EBAP molecular marker technique. The drawing was not scaled.

Table 3  
EBAP primers used in the present study.

Primer ID	Primer sequence (5'→3')	GC content	Primer ID	Primer sequence (5'→3')	GC content
EBAP1	GAATTCGGCGGCGATA	58.82%	EBAP19	GAATTCGGCGGCGACG	70.59%
EBAP2	GAATTCGGCGGCGATC	64.71%	EBAP20	GAATTCGGCGGCGAGC	70.59%
EBAP3	GAATTCGGCGGCGGATG	64.71%	EBAP21	GAATTCGGCGGCGGAGT	64.71%
EBAP4	GAATTCGGCGGCGGAAC	64.71%	EBAP22	GAATTCGGCGGCGGAAT	58.82%
EBAP5	GAATTCGGCGGCGGTGC	70.59%	EBAP23	GAATTCGGCGGCGGATT	58.82%
EBAP6	GAATTCGGCGGCGGTAC	64.71%	EBAP24	GAATTCGGCGGCGGACT	64.71%
EBAP7	GAATTCGGCGGCGGTCC	70.59%	EBAP25	GAATTCGGCGGCGGTAG	64.71%
EBAP8	GAATTCGGCGGCGGTGA	64.71%	EBAP26	GAATTCGGCGGCGGTGA	64.71%
EBAP9	GAATTCGGCGGCGGATA	58.82%	EBAP27	ATATATCGGCGGCGACT	52.94%
EBAP10	GAATTCGGCGGCGGATC	64.71%	EBAP28	ATATATCGGCGGCGTCC	58.82%
EBAP11	ATCGATCGGCGGCGATC	64.71%	EBAP29	AGTACTCGGCGGCGATA	58.82%
EBAP12	ATATATCGGCGGCGATA	47.06%	EBAP30	AGTACTCGGCGGCGTAC	64.71%
EBAP13	ATATATCGGCGGCGATC	52.94%	EBAP31	TGATCAGCGGCGGATA	58.82%
EBAP14	ATATATCGGCGGCGATG	52.94%	EBAP32	TGATCAGCGGCGGTAC	64.71%
EBAP15	ATATATCGGCGGCGTCC	58.82%	EBAP33	AATATTCGGCGGCGATA	47.06%
EBAP16	GAATTCGGCGGCGGACT	64.71%	EBAP34	AATATTCGGCGGCGTAC	52.94%
EBAP17	GAATTCGGCGGCGGAAA	58.82%	EBAP35	TTTAAACGGCGGCGATA	47.06%
EBAP18	GAATTCGGCGGCGGTTT	58.82%	EBAP36	TTTAAACGGCGGCGTAC	52.94%

indexes of TNB, NPB, PPB, Na, Ne, H, I, PI, Rp, PIC, MI for each primer were summarized in Table 4. The amplification results of EBAP9 and IBAP10 in ten maize varieties and EBAP9 in eight sugarcane varieties were shown in Fig. 2.

In maize, the genetic distance ranged from 0.04 to 0.24, with an average of 0.16. In sugarcane, the genetic distance varied from 0.1 to 0.35, with a mean of 0.2. According to the genetic distance, the cluster analysis revealed that ten maize varieties and eight sugarcane varieties were divided into four and three classes, respectively (Fig. 3a).

### 3.2. EBAP markers in potato, cassava and cabbage (Dicotyledons)

Initially, thirteen, thirty-six, and fifteen single primers were used among nine potato, five cassava, and sixteen cabbage varieties respectively. After screening, seven, twenty-one, and ten single primers could successfully produce unambiguous and reproducible band patterns. Using these primers, eighty-one, one hundred and twenty-three, and seventy-nine bands were amplified, of which sixty-nine, sixty-seven, and fifty-five were polymorphic bands, accounting for 85.19%, 54.47%, and 69.62% respectively. The indexes of TNB, NPB, PPB, Na, Ne, H, I, PI, Rp, PIC, MI for each

primer were summarized in Table 5. The agarose gel electrophoresis of EBAP4 and EBAP7 in nine potato varieties, EBAP6, EBAP38, EBAP39 and EBAP42 in five cassava varieties, and EBAP2 and EBAP8 in sixteen cabbage accessions were shown in Fig. 4.

Among nine potato varieties, the genetic distance varied from 0.13 to 0.55, averaging 0.32. The genetic distance of five cassava varieties ranged from 0.1 to 0.22, with a mean of 0.19. As for sixteen cabbage accessions, the genetic distance varied from 0.06 to 0.42, with an average of 0.17. Based on the genetic distance, phylogenetic tree showed that nine potato varieties, five cassava varieties and sixteen cabbage accessions were all partitioned into three groups (I, II and III), respectively (Fig. 5).

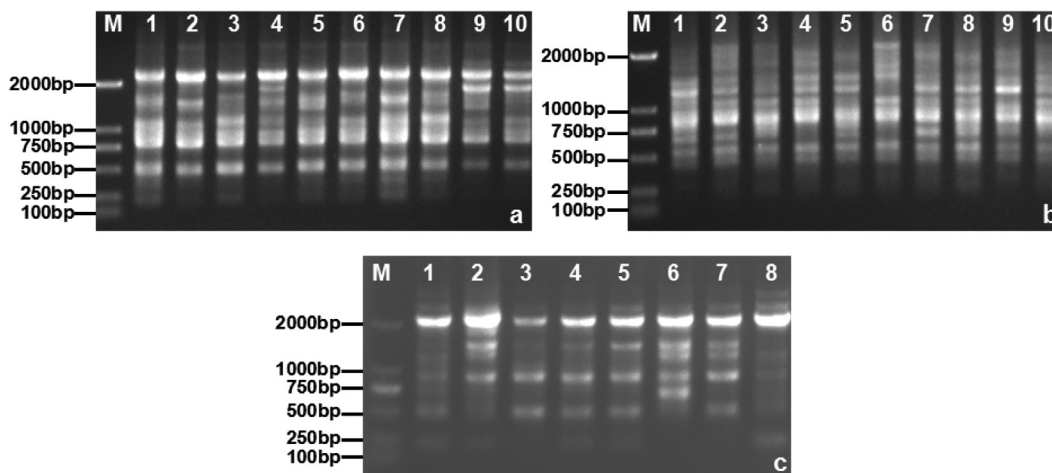
### 3.3. EBAP markers in peanut

In this study, fifteen single primers were used to detect the DNA polymorphism of twelve cultivated peanut varieties and four wild peanut accessions of section *Arachis*. It was found that except for four wild peanut accessions with different band patterns, the remaining twelve cultivated peanut varieties had the same band patterns. These results preliminary demonstrated that the DNA polymorphism of cultivated peanut varieties in genus *Arachis* could

**Table 4**  
DNA polymorphisms among ten maize varieties and eight sugarcane varieties revealed by EBAP primers.

Primer ID	TNB	NPB	PPB (%)	Na	Ne	H	I	PI	Rp	PIC	MI
Ten maize varieties											
EBAP1	7	1	14.29	1.1428	1.0332	0.0269	0.0481	0.32	0.4	0.8637	3.0708
EBAP2	7	6	85.71	1.8571	1.4957	0.2807	0.4196	1.66	2.2	0.7616	2.708
EBAP3	5	1	20	1.2	1.0215	0.0194	0.0404	0.18	0.2	0.8072	2.87
EBAP4	7	4	57.14	1.5714	1.335	0.2024	0.3047	1.78	3	0.8143	2.8952
EBAP5	4	1	25	1.25	1.0582	0.0472	0.0843	0.32	0.4	0.77	2.7378
EBAP6	3	1	33.33	1.3333	1.1788	0.1164	0.1779	0.48	0.8	0.72	2.56
EBAP7	5	2	40	1.4	1.163	0.1023	0.1614	0.68	1.2	0.7912	2.8132
EBAP9	9	4	44.44	1.4444	1.2099	0.1316	0.2043	1.56	2.4	0.8472	3.0121
EBAP10	8	6	75	1.75	1.502	0.2923	0.4312	1.7	2.2	0.8616	3.0633
EBAP13	5	3	60	1.6	1.3494	0.2065	0.3117	1.12	1.6	0.784	2.7876
EBAP17	7	5	71.43	1.7142	1.27	0.1785	0.2868	1.48	2	0.7282	2.589
EBAP18	10	10	100	2	1.5361	0.3246	0.4937	3.74	5.8	0.6508	2.314
EBAP21	7	6	85.71	1.8571	1.5311	0.3073	0.4587	2.4	3.6	0.738	2.6239
EBAP25	8	4	50	1.5	1.2319	0.1412	0.2201	1.3	1.8	0.8266	2.9389
EBAP27	4	1	25	1.25	1.094	0.0683	0.1113	0.42	0.6	0.7762	2.76
EBAP31	4	1	25	1.25	1.094	0.0683	0.1113	0.42	0.6	0.7762	2.76
EBAP33	9	7	77.78	1.7777	1.492	0.2819	0.4196	2.36	3.2	0.8165	2.9033
EBAP34	4	1	25	1.25	1.094	0.0683	0.1113	0.42	0.6	0.7762	2.76
Mean	6.28	3.56	50.82	1.5664	1.2989	0.1806	0.2758	1.24	1.8111	0.7839	2.7871
Total	113	64	56.64	-	-	-	-	-	-	-	-
Eight sugarcane varieties											
EBAP1	11	7	63.64	1.6363	1.3838	0.2198	0.3279	2.48	3.75	0.8407	4.7917
EBAP2	9	6	66.67	1.6666	1.2909	0.1743	0.2747	1.89	2.5	0.7523	4.2882
EBAP4	13	12	92.31	1.923	1.482	0.2873	0.4401	4.22	6	0.6869	3.9156
EBAP5	6	1	16.67	1.1666	1.0229	0.0201	0.0399	0.22	0.25	0.8394	4.7846
EBAP6	11	8	72.73	1.7272	1.3875	0.2241	0.3411	2.52	3.5	0.7893	4.4988
EBAP7	8	5	62.5	1.625	1.2917	0.1862	0.2909	2.01	3	0.7764	4.4253
EBAP8	11	7	63.64	1.6363	1.3558	0.2056	0.311	2.27	3	0.8332	4.749
EBAP9	10	7	70	1.7	1.4295	0.2466	0.3679	2.61	3.75	0.8178	4.6615
EBAP10	5	3	60	1.6	1.3958	0.2155	0.3164	0.82	1	0.8037	4.5814
EBAP13	3	1	33.33	1.3333	1.0458	0.0402	0.0798	0.22	0.25	0.691	3.9385
Mean	8.7	5.7	60.15	1.6552	1.3468	0.2038	0.3109	1.93	2.7	0.7831	4.4635
Total	87	57	65.52	-	-	-	-	-	-	-	-

Note: TNB: Total Number of Bands; NPB: Number of Polymorphic Bands; PPB: Percentage of Polymorphic Bands; Na: Observed Number of Alleles; Ne: Effective Number of Alleles; H: Nei's Gene Diversity; I: Shannon's Information Index; PI: Polymorphic Index; Rp: Resolving Power; PIC: Polymorphism Information Content; MI: Marker Index.



**Fig. 2.** The amplification results of ten maize varieties with EBAP9 (a) and EBAP10 (b), and eight sugarcane varieties with EBAP9 (c) respectively.

not be detected using EBAP technique. The amplification band patterns of EBAP1 and EBAP9 in sixteen peanut accessions were shown in Fig. 6.

**3.4. Display of differentially expressed genes in plants by EBAP technique**

Since the single primer of EBAP technique is designed according to the GC bases rich in exon region of genes in

eukaryotic genomes, it is speculated that EBAP technique can be used not only for molecular marker development at the DNA level, but also for differentially expressed genes display at the RNA/cDNA level. When single primers (EBAP1~EBAP8) of EBAP technique were used to amplify the cDNA of peanut, the eight single primers obtained different amplification band patterns (Fig. 7). The results preliminarily indicated that EBAP technique could be used to display the differentially expressed genes in plants.



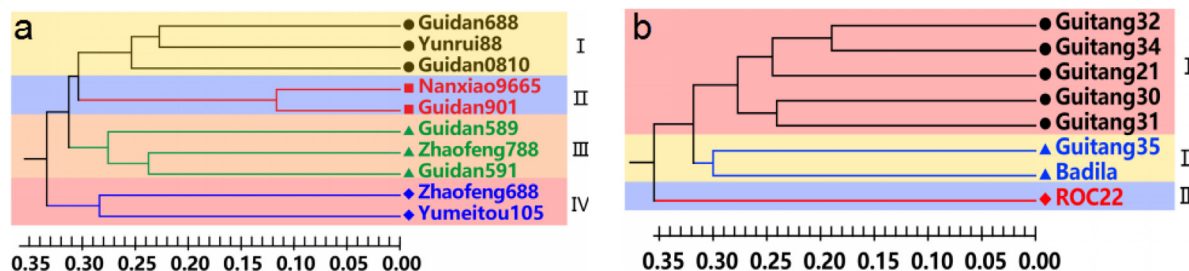


Fig. 3. Cluster analysis of phylogenetic relationships among ten maize varieties (a) and eight sugarcane varieties (b) based on EBAP markers.

Table 5

DNA polymorphisms in nine potato varieties, five cassava varieties and sixteen cabbage accessions revealed by EBAP primers.

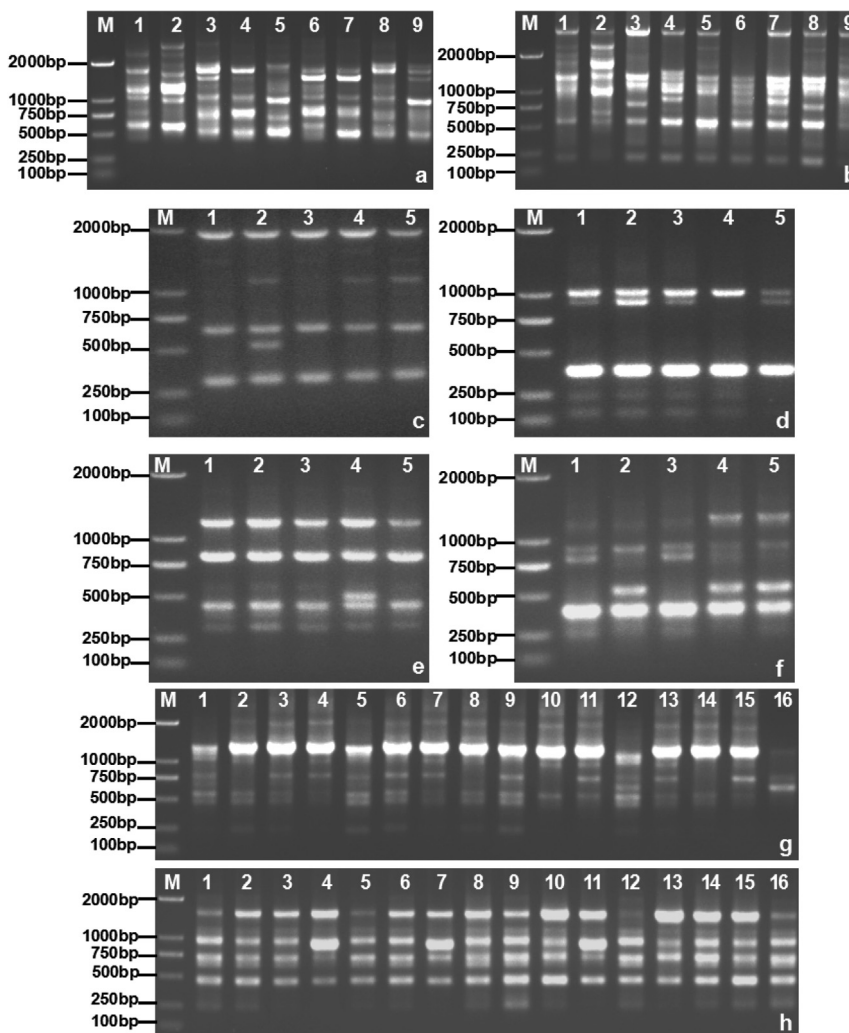
Primer ID	TNB	NPB	PPB (%)	Na	Ne	H	I	PI	Rp	PIC	MI
Nine potato varieties											
EBAP2	7	6	85.71	1.8571	1.4661	0.2717	0.4113	2.17	3.3333	0.6939	6.8397
EBAP4	14	14	100	2	1.4583	0.2843	0.4444	5.03	7.3333	0.5644	5.5631
EBAP6	15	11	73.33	1.7333	1.4136	0.241	0.3643	3.43	4.4444	0.8284	8.1659
EBAP7	12	10	83.33	1.8333	1.5028	0.2936	0.4393	3.8	5.7778	0.7767	7.6565
EBAP8	14	13	92.86	1.9285	1.615	0.3506	0.518	5.05	7.5556	0.7756	7.6456
EBAP9	8	7	87.5	1.875	1.6119	0.3408	0.4972	2.32	3.3333	0.816	8.0432
EBAP13	11	8	72.73	1.7272	1.2711	0.1844	0.3005	2.92	4	0.7026	6.9254
Mean	11.57	9.86	85.07	1.8519	1.4741	0.2801	0.4244	3.53	5.1111	0.7368	7.2628
Total	81	69	85.19	-	-	-	-	-	-	-	-
Five cassava varieties											
EBAP1	4	2	50	1.5	1.2754	0.1634	0.2487	0.8	1.2	0.76	2.4248
EBAP3	7	6	85.71	1.8571	1.4992	0.295	0.4452	2.24	3.2	0.7363	2.3492
EBAP6	5	2	40	1.4	1.2203	0.1307	0.1989	0.8	1.2	0.8064	2.5728
EBAP7	4	4	100	2	1.6469	0.3855	0.5721	1.44	2.8	0.68	2.1695
EBAP15	3	2	66.67	1.6666	1.2564	0.1793	0.2903	0.8	1.2	0.6489	2.0703
EBAP16	3	2	66.67	1.6666	1.2564	0.1793	0.2903	0.8	1.2	0.6489	2.0703
EBAP18	10	5	50	1.5	1.1467	0.1104	0.1882	1.76	2.4	0.8024	2.56
EBAP19	10	6	60	1.6	1.3053	0.1919	0.2963	2.4	3.6	0.8184	2.6111
EBAP20	7	2	28.57	1.2857	1.2794	0.1412	0.1964	0.64	0.8	0.8686	2.7712
EBAP21	5	2	40	1.4	1.2421	0.1366	0.2049	0.64	0.8	0.8128	2.5932
EBAP22	8	5	62.5	1.625	1.3181	0.1907	0.2946	1.76	2.4	0.795	2.5364
EBAP23	6	3	50	1.5	1.379	0.2131	0.3081	1.44	2.4	0.8378	2.6729
EBAP24	5	4	80	1.8	1.5232	0.2994	0.4432	1.6	2.4	0.752	2.3992
EBAP26	7	4	57.14	1.5714	1.2632	0.1766	0.2769	1.76	2.8	0.7902	2.5211
EBAP27	5	1	20	1.2	1.0465	0.0377	0.0674	0.32	0.4	0.8128	2.5932
EBAP30	5	3	60	1.6	1.5649	0.2907	0.4065	0.32	1.6	0.8288	2.6443
EBAP32	6	3	50	1.5	1.2176	0.1478	0.2341	1.28	2	0.7956	2.5382
EBAP33	5	2	40	1.4	1.2203	0.1307	0.1989	0.8	1.2	0.8064	2.5728
EBAP34	5	1	20	1.2	1.0465	0.0377	0.0674	0.32	0.4	0.8128	2.5932
EBAP35	7	4	57.14	1.5714	1.4056	0.2262	0.33	1.76	2.8	0.8343	2.6618
EBAP36	6	4	66.67	1.6666	1.436	0.2495	0.3693	1.6	2.4	0.8	2.5524
Mean	5.86	3.19	54.81	1.5447	1.3085	0.1844	0.2795	1.20	1.8667	0.7833	2.4989
Total	123	67	54.47	-	-	-	-	-	-	-	-
Sixteen cabbage accessions											
EBAP1	11	5	45.45	1.4545	1.3401	0.19	0.2754	1.99	3	0.9008	4.9542
EBAP2	6	5	83.33	1.8333	1.6022	0.3272	0.4747	1.62	2.25	0.8101	4.4556
EBAP3	7	3	42.86	1.4285	1.0957	0.0751	0.1326	0.88	1.125	0.7945	4.3697
EBAP6	6	3	50	1.5	1.3491	0.201	0.2947	1.34	2.125	0.8344	4.5893
EBAP7	13	10	76.92	1.7692	1.4307	0.2547	0.3838	3.27	4.875	0.8177	4.4972
EBAP8	9	8	88.89	1.8888	1.6554	0.3551	0.5153	2.43	3.125	0.841	4.6258
EBAP9	7	4	57.14	1.5714	1.3972	0.2205	0.323	1.1	1.375	0.86	4.7301
EBAP13	5	5	100	2	1.6976	0.3916	0.5732	1.6	2.375	0.7731	4.2522
EBAP14	8	5	62.5	1.625	1.4686	0.2651	0.384	1.71	2.75	0.8766	4.8212
EBAP15	7	7	100	2	1.5815	0.3075	0.4518	1.82	2.375	0.6569	3.6129
Mean	7.9	5.5	70.71	1.6962	1.452	0.254	0.3743	1.78	2.5375	0.8165	4.4908
Total	79	55	69.62	-	-	-	-	-	-	-	-

#### 4. Discussion

##### 4.1. The principle of EBAP technique

Primer design is the core of EBAP technique. The single primer is designed based on the GC bases which are rich in exon region of genes in eukaryote genome. The single primer acts as the upstream and downstream primers and plays the same role as the single pri-

mer of RAPD, ISSR, SCoT, and CDDP. The main difference is that the single primer of EBAP can probably simultaneously anchor two sites which are not far apart in exon region of genes in eukaryote genome. Sometimes the two binding sites are located in the same exon, sometimes located in two different exons spanning introns. The resulting amplified products probably include the exons region and even the adjacent introns region. To determine whether the amplified products include the exons region or even the adjacent



**Fig. 4.** The agarose gel electrophoresis of nine potato varieties with EBAP4 (a) and EBAP7 (b), five cassava varieties with EBAP6 (c), EBAP38 (d), EBAP39 (e), and EBAP42 (f), and sixteen cabbage accessions with EBAP2 (g) and EBAP8 (h) respectively.

introns region or not, the next step is to clone and sequence the amplified products. Although most of molecular markers produced by EBAP technique are dominant (polymorphisms caused by point mutations at single primer binding sites), perhaps very few co-dominant molecular markers caused by insertion/deletion and intron length variation between single primer binding sites may also be produced. In addition, EBAP also can be used to display differentially expressed genes in plants.

#### 4.2. The universality and validity of EBAP primers

In the present study, the universality of single primers of EBAP was verified by the band patterns generated by the amplification of the same single primer in different crops. The effectiveness of single primers was verified by the band patterns produced by the different single primers in the same crop or in different crops.

As shown in Fig. 8, the same single primer produced different band patterns in four different crops, indicating that single primers of EBAP could amplify among different crops and had strong universality. As shown in Fig. 9, the different single primers of EBAP2~EBAP9 for cassava, EBAP1~EBAP8 for sugarcane, and EBAP1~EBAP8 for peanut all produced different band patterns, indicating that single primers of EBAP could bind at the different sites of exons of genes in the genome for effective amplification.

#### 4.3. Separation and staining of EBAP-PCR products

The PCR products could be separated by agarose gel electrophoresis, high resolution non-denatured/denatured polyacrylamide gel and higher resolution capillary electrophoresis. The gels could be stained with ethidium bromide (EB) or silver nitrate ( $\text{AgNO}_3$ ). Researchers could choose the best method according to the specific experimental conditions.

#### 4.4. The application effects of EBAP molecular marker technique in six crops

To assess the application potential of EBAP technique, we first applied it to the DNA polymorphisms detection of maize, sugarcane, potato, cassava, and cabbage. The results showed that it detected relatively rich DNA polymorphisms in these five crops, and each crop also obtained a good molecular fingerprinting. Also all the accessions of these five crops could be distinguished. In addition to these five crops, cultivated peanut varieties were also used to verify the effectiveness of EBAP technique. The results demonstrated that cultivated peanut could be well amplified by EBAP technique. However, the DNA polymorphism of cultivated peanut could not be detected. The most reasonable explanation is that the genetic basis of cultivated

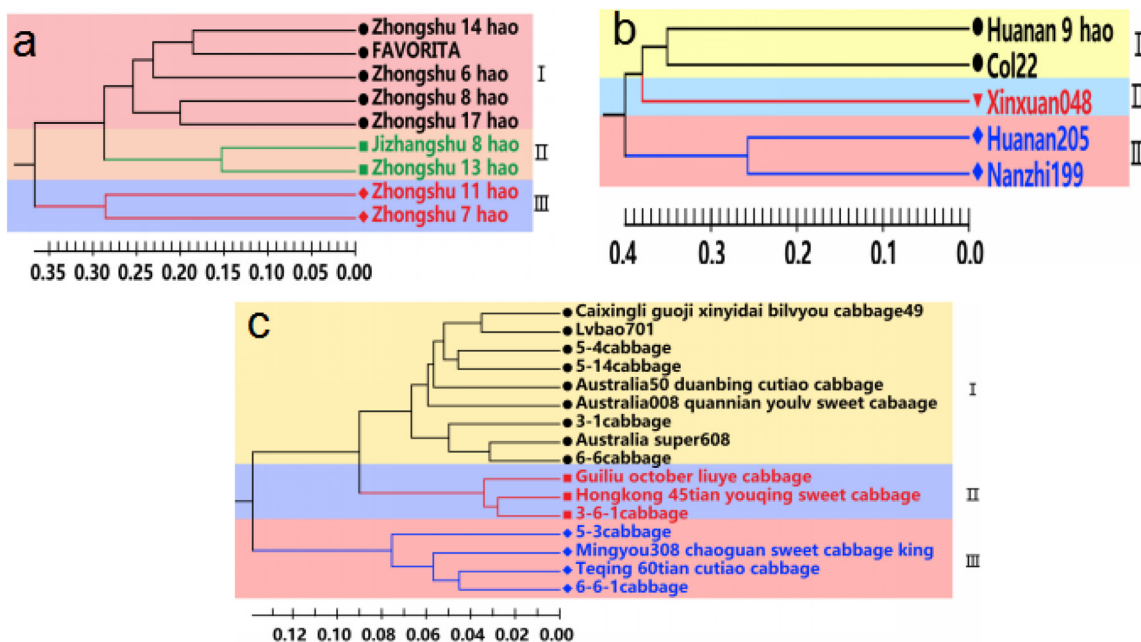


Fig. 5. Cluster analysis of nine potato varieties (a), five cassava varieties (b) and sixteen cabbage accessions (c) based on EBAP markers.

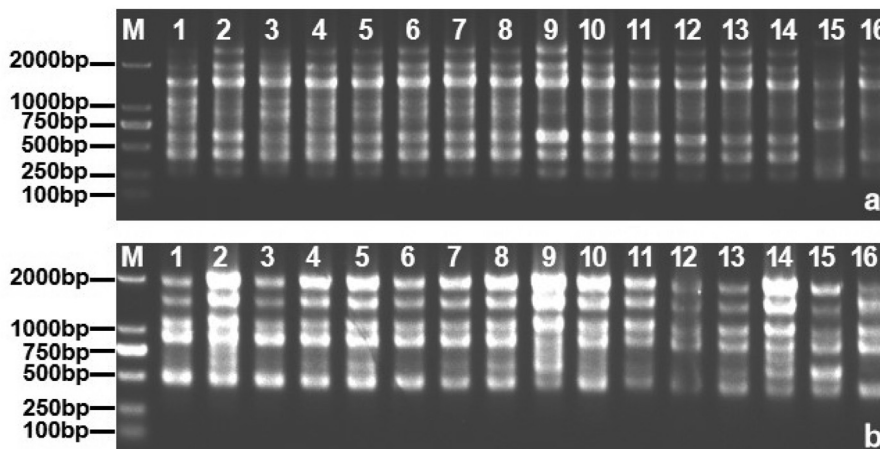


Fig. 6. The amplification band patterns of sixteen peanut varieties with EBAP1 (a) and EBAP9 (b) respectively.

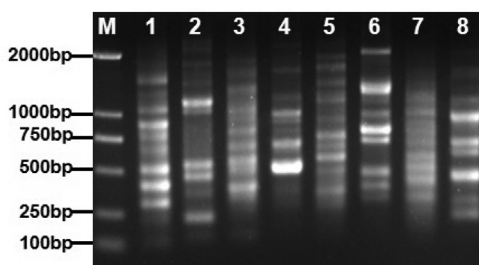


Fig. 7. The amplification band patterns of peanut cDNA exploiting different primers. Lane M represents DNA ladder. Lanes 1 to 8 represent EBAP1~EBAP8.

peanut is quite narrow and the genetic diversity is extremely low [16].

In order to comprehensively evaluate the ability of EBAP to amplify the DNA polymorphisms in different crops, the application results of EBAP in maize, sugarcane, potato, cassava, and cabbage

were comprehensively compared (Table 6). The results showed that: a) The DNA polymorphisms were all over 50% in five crops, which were relatively rich. b) For the average total number of bands, the number of polymorphic bands and the average polymorphic index, the results were as follows: potato>sugarcane>cabbage>maize>cassava. c) For the detected DNA polymorphisms, the average observed number of alleles and the average marker index, the results were as follows: potato>cabbage>sugarcane>maize>cassava. d) For the average effective number of alleles, the average Nei's gene diversity and the average Shannon's information index, the results were as follows: potato>cabbage>sugarcane>cassava>maize. e) For the average resolving power, the result was as follows: potato>sugarcane>cabbage>cassava>maize. f) For the average polymorphism information content, the result was as follows: cabbage>maize>cassava>sugarcane>potato. All of the above results indicated that EBAP technique had good application effects in both monocotyledon and dicotyledon crops.



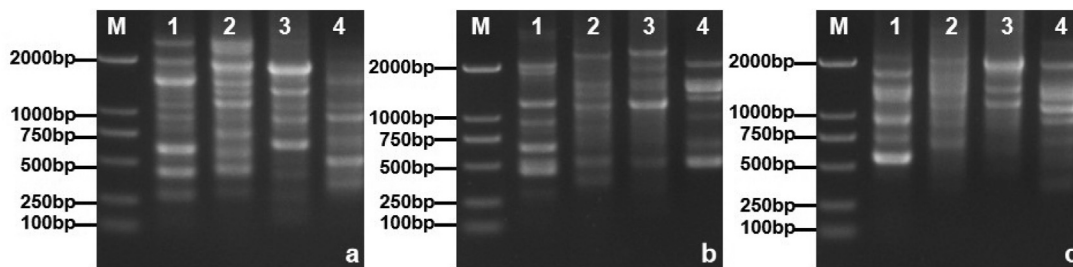


Fig. 8. The amplification band patterns of different plant materials with EBAP1 (a), EBAP4 (b), and EBAP14 (c) respectively. Lane M represents DNA ladder. Lanes 1 to 4 represent peanut (Guihua1026), fruit sugarcane (Badila), maize (Guidan688), and cabbage (Australia super608) respectively.

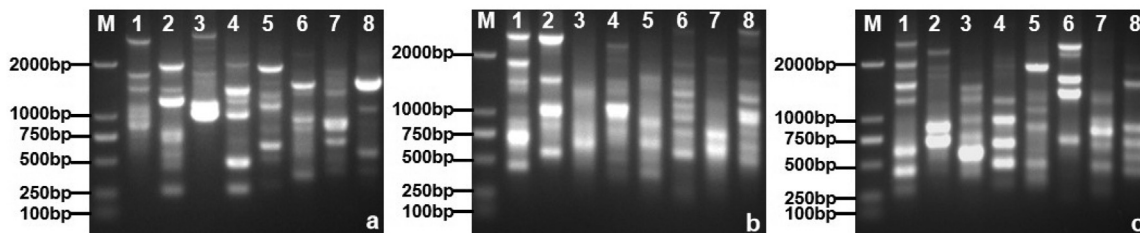


Fig. 9. The band patterns of cassava variety “Huanan205” (a), sugarcane variety “ROC22” (b), and peanut variety “Guihua1026” (c) produced by different primers. For Fig. 9a, lane M represents DNA ladder, lanes 1 to 8 represent EBAP2~EBAP9. For Fig. 9b, c, lane M represents DNA ladder, lanes 1 to 8 represent EBAP1~EBAP8.

Table 6  
Comparison of the application results of EBAP markers in five crops.

Parameters	Crops				
	Maize	Sugarcane	Potato	Cassava	Cabbage
Number of genotypes	10	8	9	5	16
Total assay units	33	15	13	36	15
Number of polymorphic primers	18	10	7	21	10
Percentage of polymorphic primers	54.55%	66.67%	53.85%	58.33%	66.67%
Total number of bands	113	87	81	123	79
Number of polymorphic bands	64	57	69	67	55
Average bands	6.28	8.70	11.57	5.86	7.90
Average polymorphic bands	3.56	5.70	9.86	3.19	5.50
Percentage of polymorphic bands	56.64%	65.52%	85.19%	54.47%	69.62%
Average observed number of alleles	1.5664	1.6552	1.8519	1.5447	1.6962
Average effective number of alleles	1.2989	1.3468	1.4741	1.3085	1.4520
Average Nei’s gene diversity	0.1806	0.2038	0.2801	0.1844	0.2540
Average Shannon’s information index	0.2758	0.3109	0.4244	0.2795	0.3743
Average polymorphic index	1.24	1.93	3.53	1.20	1.78
Average resolving power	1.8111	2.7000	5.1111	1.8667	2.5375
Average polymorphism information content	0.7839	0.7831	0.7368	0.7833	0.8165
Average genetic distance	0.16	0.20	0.32	0.21	0.17
Average marker index	2.7871	4.4635	7.2628	2.4989	4.4908

#### 4.5. Characteristics and advantages of EBAP technique

As a new type of molecular marker technique, EBAP has seven major characteristics and advantages. a) Primers design are simple. Single primers can be designed in large quantities as long as the principle of primer design is kept. The designed single primers can be universally used across eukaryotes, thus greatly reducing the cost of primer synthesis, greatly improving the utilization rate of primers and facilitating the rapid establishment of EBAP technique in various eukaryotes. b) Although EBAP technique is based on single primer amplification like RAPD and ISSR, it is likely to produce molecular markers related to traits. c) The PCR amplification of EBAP is carried out at a higher annealing temperature (50°C), which overcomes the disadvantages of SRAP and TRAP techniques that are easy to produce false positive bands due to the lower initial annealing temperature (35°C). It has advantages of strong amplification specificity and good amplification repeatability. d) EBAP uses rela-

tively long single primers, which is theoretically more reproducible and stable than traditional molecular marker technique such as RAPD. e) EBAP is a molecular marker system based on PCR technology. It is simple and reliable in operation, high in amplification efficiency and relatively rich in amplified bands.

With the rapid development of high-throughput sequencing technology and bioinformatics, the cost and price are also falling rapidly. Many ordinary researchers can use high-throughput sequencing technology and bioinformatics to quickly and efficiently develop large quantities of SSR, InDel, and SNP markers. However, molecular marker techniques using single primers will not be replaced by the large-scale SSR, InDel and SNP markers obtained through high-throughput sequencing technology and bioinformatics. EBAP technique has the advantages of simple operation, high amplification efficiency, more amplified bands, relatively rich polymorphism, simple primer design, strong primer versatility, low cost, low requirements for operators, and are

suitable for many common laboratory. In terms of identification of germplasm resources, variety purity and hybrid authenticity, there is no need to use high-throughput sequencing technology and bioinformatics. So EBAP technique can be used as an effective supplement to high-throughput sequencing technology to produce molecular markers.

### Financial support

This work was funded by grants from the National Natural Science Foundation of China (31960409; 31960416; 31660428), Guangxi Natural Science Foundation Program (2018GXNSFDA281027; 2018GXNSFDA294004; 2020GXNSFAA297081) and Guangxi Academy of Agricultural Sciences Fund Project (GNK2017JZ13; GNK2018YM06; GNK31960409).

### Conflict of Interest

The authors declare no competing financial interest.

### References

- [1] Xiong FQ, Jiang J, Zhong RC, et al. Two novel classifications on molecular marker techniques and proposal of targeted molecular marker technique. *Chin Agri Sci Bull* 2010;26(10):60–4.
- [2] Botstein D, White RL, Skolnick M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980;32(3):314–31.
- [3] Williams JGK, Kubelik AR, Livak KJ, et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 1990;18(22):6531–5. <https://doi.org/10.1093/nar/18.22.6531>.
- [4] Zietkiewicz E, Rafalski A, Labuda D. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 1994;20:176–83. <https://doi.org/10.1006/geno.1994.1151>.
- [5] Vos P, Hogers R, Bleeker M, et al. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 1995;23(21):4407–14. <https://doi.org/10.1093/nar/23.21.4407>.
- [6] Tautz D, Renz M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res* 1984;12(10):4127–38. <https://doi.org/10.1093/nar/12.10.4127>.
- [7] Li G, Quiros CF. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. *Theor Appl Genet* 2001;103(2–3):455–61. <https://doi.org/10.1007/s001220100570>.
- [8] Collard BCY, Mackill DJ. Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Mol Biol Rep* 2009;27(1):86–93. <https://doi.org/10.1007/s11105-008-0060-5>.
- [9] Andersen JR, Lübberstedt T. Functional markers in plants. *Trends Plant Sci* 2003;8(11):554–60. <https://doi.org/10.1016/j.tplants.2003.09.010>.
- [10] Heath DD, Iwama GK, Devlin RH. PCR primed with VNTR core sequence yields species specific patterns and hypervariable probes. *Nucleic Acids Res* 1993;21(24):5782–5. <https://doi.org/10.1093/nar/21.24.5782>.
- [11] Kang HW, Park DS, Go SJ, et al. Fingerprinting of diverse genomes using PCR with universal rice primers generated from repetitive sequence of Korean weedy rice. *Mol Cells* 2002;13(2):281–7.
- [12] Kalendar R, Grob T, Regina M, et al. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theor Appl Genet* 1999;98(5):704–11. <https://doi.org/10.1007/s001220051124>.
- [13] Chang RY, O'Donoghue LS, Bureau TE. Inter-MITE polymorphisms (IMP): a high throughput transposon-based genome mapping and fingerprinting approach. *Theor Appl Genet* 2001;102(5):773–81. <https://doi.org/10.1007/s001220051709>.
- [14] Kalendar R, Antonius K, Smýkal P, et al. iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. *Theor Appl Genet* 2010;121(8):1419–30. <https://doi.org/10.1007/s00122-010-1398-2>.
- [15] Weining S, Langridge P. Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. *Theor Appl Genet* 1991;82(2):209–16. <https://doi.org/10.1007/BF00226215>.
- [16] Xiong FQ, Jiang J, Han ZQ, et al. Molecular characterization of high plant species using PCR with primers designed from consensus branch point signal sequences. *Biochem Genet* 2011;49(5–6):352–63. <https://doi.org/10.1007/s10528-010-9412-6>.
- [17] Singh AK, Rana MK, Singh S, et al. CAAT box-derived polymorphism (CBDP): a novel promoter-targeted molecular marker for plants. *J Plant Biochem Biot* 2014;23(2):175–83. <https://doi.org/10.1007/s13562-013-0199-5>.
- [18] Collard BCY, Mackill DJ. Conserved DNA-derived polymorphism (CDDP): a simple and novel method for generating DNA markers in plants. *Plant Mol Biol Rep* 2009;27(4):558–62. <https://doi.org/10.1007/s11105-009-0118-z>.
- [19] Levi A, Wechter WP, Harris KR, et al. High-frequency oligonucleotides in watermelon expressed sequenced tag-unigenes are useful in producing polymorphic polymerase chain reaction markers among watermelon genotypes. *J Am Soc Hortic Sci* 2010;135(4):369–78. <https://doi.org/10.21273/JASHS.135.4.369>.
- [20] Xiong FQ, Liu JX, Liu J, et al. Comparative analysis and application of five improved CTAB extraction methods for peanut DNA. *Mol Plant Breeding* 2019;17(7):2207–16.
- [21] Xiong FQ, Liu JX, Wang CC, et al. mCTAB-dLiCl method for efficiently extracting RNA from various tissues of cultivated peanuts and its verification. *J Southern Agri* 2013;44(11):1781–4.
- [22] Raina SN, Rani V, Kojima T, et al. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome* 2001;44(5):763–72. <https://doi.org/10.1139/gen-44-5-763>.
- [23] Amiryousefi A, Hyvönen J, Poczaí P. iMEC: Online marker efficiency calculator. *Appl Plant Sci* 2018;6(6):e01159. <https://doi.org/10.1002/aps3.1159>.
- [24] Powell W, Morgante M, Andre C, et al. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) marker for germplasm analysis. *Mol Breeding* 1996;2(3):225–38. <https://doi.org/10.1007/BF00564200>.