Identification of massive molecular markers in *Echinochloa phyllopogon* using a restriction-site associated DNA approach

Guoqi Chen a, b, Wei Zhang a, b, Jiapeng Fang a, b, Liyao Dong a, b, *

a College of Plant Protection, Nanjing Agricultural University, Nanjing 210095, China
b Key Laboratory of Integrated Pest Management on Crops in East China (Nanjing Agricultural University), Ministry of Agriculture, Nanjing 210095, China

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

Article history:
Received 22 February 2017
Received in revised form 23 August 2017
Accepted 28 August 2017
Available online 14 September 2017

(Editor: Lianming Gao)

**Keywords:**
*Echinochloa phyllopogon*  
Polymorphic  
RAD sequencing  
SNP  
SSR

**A B S T R A C T**

*Echinochloa phyllopogon* proliferation seriously threatens rice production worldwide. We combined a restriction-site associated DNA (RAD) approach with Illumina DNA sequencing for rapid and mass discovery of simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers for *E. phyllopogon*. RAD tags were generated from the genomic DNA of two *E. phyllopogon* plants, and sequenced to produce 5197.7 Mb and 5242.9 Mb high quality sequences, respectively. The GC content of *E. phyllopogon* was 45.8%, which is high for monocots. In total, 4710 putative SSRs were identified in 4132 contigs, which permitted the design of PCR primers for *E. phyllopogon*. Most repeat motifs among the SSRs identified were dinucleotide (>82%), and most of these SSRs were four motif-repeats (>75%). The most frequent motif was AT, accounting for 36.3–37.2%, followed by AG and AC. In total, 78 putative polymorphic SSR loci were found. A total of 49,179 SNPs were discovered between the two samples of *E. phyllopogon*, 67.1% of which were transversions and 32.9% were transitions. We used eight SSRs to study the genetic diversity of four *E. phyllopogon* populations collected from rice fields in China and all eight loci tested were polymorphic.

@2017 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
2. Material and methods

2.1. DNA isolation

Seeds from *E. phyllopogon* individuals were collected and cultivated to fruiting stage in a greenhouse at Nanjing Agricultural University. Two *E. phyllopogon* plants with typical characteristics were used for SSR identification. Total genomic DNA was extracted from young leaves using DNeasy Plant Mini Kits (Qiagen, USA) according to the manufacturer's protocol.

The RAD library was constructed at Hengchuang Inc. (China), according to the protocol described by Baird et al. (2008). Briefly, genomic DNA (300 ng) was digested for 60 min at 37 °C in a 50 μL reaction containing 20 U each of SgrAI and PsiI (New England Biolabs, Beverly, MA, USA). Reactions were stopped by incubating at 65 °C for 20 min. The P1 adapter (a modified illumina adapter, see Baird et al., 2008) was ligated to the products of the restriction reaction, and the “barcoding” of the various samples was achieved with a set of index nucleotides in the P1 adapter sequence. A 2.5 μL aliquot of 100 mM P1 adapter was added to each sample, along with 1 μL 10 mM ATP (Promega), 1 μL 10× NEBuffer 4, 1 μL (equivalent to 1000 U) T4 DNA ligase (Enzymatics, Inc) and 5 μL water, then incubated at room temperature for 20 min, before heat-inactivated (20 min at 65 °C). The reactions were then pooled and the products randomly sheared to a mean size of 500 bp using a Bio-Rad (Diagenode). The material was electrophoresed through a 1.5% agarose gel, and the DNA in the range 300–800 bp isolated was used in a MinElute Gel Extraction Kit (Qiagen). dsDNA ends were treated with end blunting enzymes (Enzymatics, Inc) to remove overhangs, and the samples purified using a MinElute column (Qiagen). 3’-adenine overhangs were then added by the addition of 15 U Klenow exo-(Enzymatics), followed by incubation at 37 °C for 10 min. Following re-purification, 1 μL 10 μM P2 adapter (a modified illumina adapter, see Baird et al., 2008) was ligated, as described above for P1. The samples were then purified as before, and eluted in a volume of 50 μL. Following quantification (Qubit fluorimeter), 20 ng were taken as the template for a 100 μL PCR containing 20 μL Phusion Master Mix (NEB), 5 μL 10 μM P1 adapter primer (Illumina), 5 μL 10 μM P2 adapter primer (Illumina) and water. The Phusion PCR settings followed product guidelines (NEB) over 18 cycles. The amplicons were gel purified, the size range 300–700 bp was excised from the gel, with the DNA content adjusted to 3 ng/μL. The constructed RAD libraries were sequenced on the NGS illumina platform PE150 at Hengchuang Inc. (China), following the manufacturer's protocol.

To obtain clean, high quality reads, we discarded low quality raw sequences with adapter contamination or N content >10%. We used Stacks software for RAD tag clustering for each sample (ustacks). The Reads group (Read1 and Read2) at a same enzyme loci RAD were assembled by using the ABYSS software (Catchen et al., 2011).

2.3. SSR identification

SSR motifs were identified by SSRIT software (http://www.gramene.org/db/markers/ssritool) using default parameters (Temnykh et al., 2001). Both perfect and imperfect di- , tri-, tetra-, penta- and hexa-nucleotide motifs were targeted. Di-nucleotide motifs with at least 4 repeats and other motifs with at least 3 repeats were selected. We used Primer3 software (http://sourceforge.net/projects/primer3/) to design primers in the flank regions of SSR sequences (SSR sequences were not contained in the primers), the replicated primers were removed and unique primers and relative loci were retained.

To analyze the frequency of SSR motifs, SSRs were first standardized (Wang et al., 2015). For example, SSRs with motifs of AT and TA were analyzed as AT, and motifs of ATG, TGA, GAT, TAC, ACT and CAT are analyzed as ATG.

2.4. Sequence annotation

For the contigs with SSR loci, sequence annotation and Gene Ontology analyses were further conducted. BlastN searches were performed against the Gene Ontology database (http://www.geneontology.org/), using 90% identity and a minimum alignment of 100 bp as cut-off parameters. A threshold E-value of e−15 was adopted for each annotation. The annotated sequences were assigned a function based on the Gene Ontology database (http://www.geneontology.org/); GO terms were determined with respect to cellular component, biological process and molecular function (Barchi et al., 2011).

2.5. SNP discovery

SNPs were detected by Stacks pipeline, ustacks software was used to build loci, cstacks software was used to create a catalog of loci, and stacks software was used to match samples back against the catalog (Catchen et al., 2011). Default settings were used in Stacks.

2.6. Microsatellites amplification

To test the validity of the SSRs identified by RAD sequencing here, we used eight SSRs (Table 1) to study the genetic diversity of four *E. phyllopogon* populations collected from rice fields in China. We extracted total genomic DNA from four-leaf stage plants using a DNeasy Plant Mini Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's instructions. Isolated DNA concentration and relative purity were checked using Nanodrop ND-1000 (Thermo Scientific), and adjusted to 30–40 ng/μL. Forward primers of SSRs were labeled with fluorescent tags (Table 1). PCR amplification was conducted in a total volume of 10 μL. The PCR mixture contained 0.2 μL of DNA, 0.4 μL of each primer (10 μM), 5 μL of 2× PCR Taq Mix (Dongsheng Biotech, China), and ddH2O to a final volume of 10 μL. The amplifications were performed using the following cycling program: initial denaturation at 94 °C for 4 min, followed by 35 cycles of 94 °C for 30 s, relative annealing temperatures for 30 s, and 72 °C for 1 min, with a final extension step at 72 °C for 10 min. The amplification products were combined with formamide and a size standard GeneScan-500 LIZ (Applied Biosystems, Foster City, California, USA), and separated on a 3730 ABI automated sequencer ( Applied Biosystems). Sample profiles were scored manually using GeneMarker v. 2.4 ( Applied Biosystems).

2.7. Data analysis

The multilocus data were transformed to a binary matrix of presence/absence of each allele for each individual, which was used for further analysis with GenAlex 6.5 (Peakall and Smouse, 2012; Teixeira et al., 2014). Total number of alleles and the number of private alleles for each population were determined using GenAlex 6.5, and genetic diversity was determined using GenoDive 2.0b23 (Teixeira et al., 2014), according to the tutorials (www.patrickmeirmans.com/software/GenoDive.html). GenoDive allows analyzing polyploids for each allele, which was used for further analysis with Genalex 6.5, and genetic diversity was determined using GenoDive2.0b23 (Teixeira et al., 2014), according to the tutorials (www.patrickmeirmans.com/software/GenoDive.html). GenoDive allows analyzing polyploids with unknown dosage of alleles (Meirmans and Van Tienderen, 2004).

3. Results

3.1. Sequencing and contig assembly

The sequencing procedure generated 71.45 million reads for the two *E. phyllopogon* samples (Table 2). After editing/trimming, 10.440.6 Mb of high quality sequences were available, which were assembled into 37,662 contigs. Average contig lengths for the two...
The abundance of SSRs decreased significantly for the two samples tested for *E. phyllopogon* genotyping: locus name, forward (F) and reverse (R) primer sequences, motif, annealing temperature (Tm), fluorescent dye used (Fl. dye), allele size range (ASR), number of alleles amplified per sample, and number of alleles amplified among the four populations sampled (Allele. total).

| Marker | Sequence | Motif | ASR (bp) | Fl. dye | Tm (°C) | No. of alleles per sample | Allele. total |
|--------|----------|-------|----------|---------|---------|---------------------------|---------------|
| EG_1   | F: GCTCCTGAACCTGATACATTCCTGC  
R: TCGATCTACCCGATCTCTCTC | TG | 123–153 | TAM | 49 | 0 | 0.7 | 2 | 5 |
| EG_2   | F: CAETCAATGCAATGAAAGGG  
R: GCTCCTGAACCTGATACATTCCTGC | TA | 131–159 | FAM | 51.5 | 1 | 1.7 | 3 | 7 |
| EG_301 | F: GGCCTGCTCAAGCTCTTCTCA  
R: TGATCTACCCGATCTCTCTC | AT | 147–173 | TAM | 57 | 0 | 2.4 | 3 | 8 |
| EG_302 | F: ATTTGAACCCCTACCAACACCAAC  
R: GAAACAGAAGGGAGGTGTGCTG | ATTT | 133–293 | TAM | 57 | 1 | 2.8 | 5 | 12 |
| EG_305 | F: AGCCTGCTCCCTAGCTGGATTCTCT  
R: TATCCAGCTCGGTGATCAGTA | AT | 100–162 | ROX | 57 | 3 | 4.1 | 6 | 14 |
| EG_306 | F: TAACAAACCGGCCGCGTGA  
R: TCATCTTACCGCTCTCGAT | CT | 146–167 | HEX | 57 | 1 | 1.25 | 2 | 7 |
| EG_307 | F: AACACCTTGTACCAAAATATCACCACATCA  
R: AACACATTAAAGACCCCTCTCCTC | ATC | 108–134 | TAM | 57 | 2 | 3.5 | 5 | 8 |
| EG_320 | F: CACCTAATACACATCAACAGGTTT  
R: GCATATTTAAGCATCAAAATGACA | TA | 136–153 | FAM | 57 | 2 | 3.0 | 4 | 5 |

FAM: 6-carboxyfluorescein, HEX: hexachloro-fluoresceine, ROX: carboxy-X-rhodamine, and TAM: 5-TAMRA (5-Carboxytetramethylrhodamine).

**Fig. 1.** SSR motifs with different repeat numbers for the two samples of *E. phyllopogon*.

The majority of motifs among the RAD SSRs were dinucleotide (>82%) for both samples, and 14%–15% of the SSR motifs were trinucleotide (Table 3). The majority of SSRs were four motif-repeats. The abundance of SSRs decreased significantly (P < 0.01) with increasing motif-repeats for *E. phyllopogon* (Fig. 1).

Nearly all (97.3%) *E. phyllopogon* SSR motifs consisted of dinucleotide plus trinucleotide repeats. Thus, we further analyzed dinucleotide and trinucleotide repeats. Before the analysis, SSRs were standardized. For example, SSRs with motifs of AT and TA were analyzed as AT, and motifs of ATG, TGA, GAT, TAC, ACT and CAT were analyzed as ATG. ATG was the most frequent, accounting for 36.3%–37.5%, followed by AG and AC (Table 4). Among the four kinds of dinucleotide motifs, CG dinucleotide repeats represented the lowest percentage of all SSRs (<6%). CCG was the most frequent

**Table 4**

SSR motifs with a frequency > 0.5% and the ranges of PCR product length (mean length) of the relative motifs for the two samples tested for *E. phyllopogon*.

| Motif | Count (% of total SSRs) | PCR product length (average length, bp) |
|-------|-------------------------|----------------------------------------|
| AT    | 854 (37.2) 880 (36.3) 80–234 (133.5) 80–239 (131.0) |
| AG    | 562 (24.5) 617 (25.5) 80–208 (126.5) 80–208 (127.1) |
| AC    | 372 (16.2) 395 (16.3) 80–225 (130.3) 80–234 (126.6) |
| CG    | 120 (5.2) 106 (4.4) 80–204 (131.6) 80–237 (124.1) |
| CGG   | 99 (4.3) 103 (4.2) 80–172 (132.0) 80–160 (126.9) |
| AAG   | 45 (2.0) 43 (1.9) 85–159 (128.8) 81–153 (121.0) |
| AAT   | 28 (1.2) 30 (1.2) 80–160 (130.3) 80–160 (127.3) |
| ACC   | 27 (1.2) 14 (0.6) 80–157 (122.9) 80–220 (134.2) |
| AAC   | 25 (1.1) 47 (1.9) 85–155 (120.3) 122–155 (136.9) |
| AGG   | 24 (1.0) 25 (1.0) 81–188 (128.3) 80–159 (132.3) |
| AGC   | 23 (1.0) 29 (1.2) 80–157 (122.6) 89–159 (134.2) |
| ACT   | 22 (1.0) 15 (0.6) 86–160 (136.1) 83–159 (121.4) |
| ATG   | 14 (0.6) 22 (0.9) 91–160 (134.5) 87–159 (127.5) |

Note: motifs with dinucleotide plus trinucleotide contributed to 97.3% of the total SSRs for both samples. Thus motifs with length >3 were not shown in this table.
kind of trinucleotide motif for both samples (Table 4), accounting for about 4% of the total SSRs for *E. phyllopogon*. The predicted length of PCR products amplified by SSR primers designed in this study are shown in Table 4.

In total, 78 putative polymorphic SSR loci were found by RAD sequencing (Table 5). These 78 SSRs include 65 SSRs with dinucleotide motifs, 10 SSRs with trinucleotide motifs, two with tetranucleotide motifs and one with a pentanucleotide motif. The AT dinucleotide repeat, which accounts for 49.4% of all motifs, was the most frequent kind.

To test the validity of the SSRs identified by RAD sequencing here, we used eight SSRs to study the genetic diversity of four *E. phyllopogon* populations collected from rice fields in China. We amplified 66 alleles from the eight microsatellite loci. The primer sequence EG_305 amplified 14 alleles, EG_302 amplified 12 alleles, and EG_320 and EG_1 amplified five alleles (Table 1). EG_305 amplified three to six alleles per sample, while EG_307 and EG_320 amplified two to four alleles per sample, respectively. Moreover, EG_305 amplified the most alleles on average (4.1). On average, 3.1–4.8 alleles were amplified from one locus per population (Table 6). All four populations showed private alleles, among which the populations EP13 and EP50 showed 13 and eight private alleles, respectively. Moreover, EG_305 amplified by SSR primers designed in this study produced 14 alleles, EG_302 amplified 12 alleles, EG_38 and EG_37 amplified 10 alleles, EG_36 amplified 8 alleles, EG_35 amplified 7 alleles, EG_34 amplified 6 alleles, EG_33 amplified 5 alleles, and EG_32 amplified 3 alleles.

Using two *E. phyllopogon* individuals, we identified 4710 SSR loci in 4132 contigs, and annotated 643 contigs (Table 52). Among these 643 contigs, 8631 annotations, potentially referring to 2155 unigenes, were searched (a given gene product can be associated with more than one annotation). Annotated *E. phyllopogon* sequences with SSR loci were functionally assigned and arranged into Gene

### Table 5

| Marker  | Motif          | Primer_F | Primer_R |
|---------|----------------|-----------|----------|
| EG_1    | TG             | getctggactgttgactatgcg | getctggactgttgactatgcg |
| EG_2    | TA             | cgtggactgttgactatgcg  | cgtggactgttgactatgcg  |
| EG_3    | TA             | tggactgttgactatgcg   | tggactgttgactatgcg   |
| EG_4    | TA             | gactgttgactatgcg     | gactgttgactatgcg     |
| EG_5    | TA             | tggactgttgactatgcg   | tggactgttgactatgcg   |
| EG_6    | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_7    | TA             | tggactgttgactatgcg   | tggactgttgactatgcg   |
| EG_8    | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_9    | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_10   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_11   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_12   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_13   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_14   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_15   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_16   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_17   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_18   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_19   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_20   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_21   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_22   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_23   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_24   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_25   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_26   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_27   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_28   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_29   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_30   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_31   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_32   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_33   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_34   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_35   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_36   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_37   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_38   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |
| EG_39   | TA             | cgtggactgttgactatgcg | cgtggactgttgactatgcg |

### Table 6

| Population | EP13 | EP14 | EP53 | Total |
|------------|------|------|------|-------|
| No. of alleles | 39   | 34   | 25   | 77    |
| No. of alleles per locus | 4.875 | 4.25 | 3.125 | 4.625 |
| No. of private alleles | 13   | 1    | 2    | 8     |
| Heterozygosity | 0.086 | 0.082 | 0.064 | 0.091 |
| Shannon’s information index | 0.381 | 0.21 | 0.087 | 0.222 |

### Table 7

| Source | df | SS | MS | Est. var. | P |
|--------|----|----|----|-----------|---|
| Among Pops | 3 | 135.563 | 45.188 | 3.328 | <0.01 |
| Within Pops | 44 | 231.250 | 5.256 | 5.256 | <0.001 |

df – degree of freedom, SS – sum of squares, MS mean squares, Est. var. – estimate of variance, % – percentage of total variation, P-value is based on 9999 permutations.

### 3.3. Annotation of contigs with SSR loci

Using two *E. phyllopogon* individuals, we identified 4710 SSR loci in 4132 contigs, and annotated 643 contigs (Table 52). Among these 643 contigs, 8631 annotations, potentially referring to 2155 unigenes, were searched (a given gene product can be associated with more than one annotation). Annotated *E. phyllopogon* sequences with SSR loci were functionally assigned and arranged into Gene
Fig. 2. Functional annotation of assembled sequences with SSR loci for the two samples of *E. phyllopon* based on gene ontology (GO) terms.
Ontology (GO) slim categories (Fig. 2). GO analyses suggested that contigs with SSR loci were mostly related to metabolic processes (12.1% of the total 2155 unigenes) and cellular processes (10.5%) among biological processes; cell (9.8%), cell part (9.7%) and organelle (8.6%) among cellular components; and binding (12.6%) and catalytic activity (8.7%) among molecular functions.

3.4. SNP discovery

In total, 49,179 SNPs were discovered between the two samples of *E. phyllopogon*. Table S3 shows the kind, sequence and location of 49,179 SNPs discovered between two samples of *E. phyllopogon*. Among these SNPs, transversions (67.1% of total SNPs) were much more frequent than transitions (Fig. 3).

4. Discussion

4.1. High GC content of *E. phyllopogon* genome

Higher GC content in plant genomes possibly contributes to an increased ability to adapt to various arable lands that are mainly maintained and regulated by human disturbance. Smarda et al. (2014) studied GC content in 239 different plant genomes, finding that the GC content of monocots varied between 33.6% and 48.9%, and increased GC content was documented in species able to grow in seasonally cold and/or dry climates, which possibly indicates GC-rich DNA may confer more stability during cell freezing and desiccation. The GC content of *E. phyllopogon* was higher than those of many monocots such as *Juncus inflexus* (33.7%), *Luzula badia* (33.6%), *Carex actutiformis* (35.6%), *Schoenoplectus lacustris* (35.8%), *Canna indica* (39.7%), *Oryza sativa* (43.6%) and *Triticum aestivum* (44.7%); and only lower than those of a few Poaceae species such as *Stipa calamagrostis* (47.5%) and *Zea mays* (47.4%) (Raats et al., 2013).

4.2. Characteristics on SSR motifs of *E. phyllopogon*

The majority of RAD SSR motifs were dinucleotide and with four motif-repeat. Gupta et al. (2015) identified SSR motifs in peanut (*Arachis hypogaea*) through RAD sequencing, and found that 67.6% of the motifs were dinucleotide, 14.6% were trinucleotide, 12.5% were tetranucleotide, 3.2% were pentanucleotide and 2.2% were hexanucleotide. Nevertheless, in eggplant (*Solanum melongena*), the percentages among total motifs with two to six nucleotides of dinucleotide, trinucleotide, tetranucleotide, pentanucleotide and hexanucleotide were 20.4%, 37.9%, 12.8%, 18.1% and 10.9% (Barchi et al., 2011). Using RAD sequencing in eggplant, Barchi et al. (2011) found that AAC was the most frequent kind of motif, accounting for 19.0% of the total SSRs, followed by AT (9.6%). Wang et al. (2015) analyzed the genomes of nine plant species from the Poaceae family, and found that among the genome SSRs of *O. sativa* ssp. indica, *O. sativa* ssp. japonica, *Phyllostachys heterocycla*, *Sorghum bicolor* and *Z. mays*, AT was the most frequent motif, and also very frequent in other Poaceae plants.

To test the validity of the SSRs identified by RAD sequencing here, we used eight SSRs to study the genetic diversity of four *E. phyllopogon* populations collected from rice fields in China. All eight loci were polymorphic, particularly when compared with the five SSRs that have been used for *Echinochloa* since 2002 (Danquah et al., 2002; Nozawa et al., 2006; Lee et al., 2015).

4.3. Potential usage of the SSRs and SNPs identified

A great number of *Echinochloa* species are aggressive invaders and managing crop lands requires unique strategies for each (Holm et al., 1979; Tabacchi et al., 2006). Thus, correctly identifying *Echinochloa* spp. is of agronomical and economic importance. The genus *Echinochloa* contains about 35 species that are widespread in both tropical and temperate regions and in dry or water-flooded soils (Flora of China, 2015). The taxonomy of this genus is complex, and *Echinochloa* species show wide variability in morphological, biological and physiological features (Danquah et al., 2002; Tabacchi et al., 2006; Vidotto et al., 2007). Conventionally, the identification of *Echinochloa* species has been attempted taxonomically using morphological assessment of plants, which has frequently been found to be difficult and uncertain (Tabacchi et al., 2006). Moreover, there are different taxonomic key systems for *Echinochloa* species, which may lead to misidentification (Flora of China, 2015; Tabacchi et al., 2006). Molecular identification of the *Echinochloa* species is not yet reliable and requires further study (Danquah et al., 2002; Kaye et al., 2014; Tabacchi et al., 2006). In addition, molecular markers may be very useful in studying the origin and distribution of herbicide-resistant populations (Okada et al., 2013; Osuna et al., 2011). SNPs and SSRs are ideal molecular tools for gene location and molecular breeding (Danquah et al., 2002; Gupta et al., 2015; Vandepitte et al., 2013; Zhang et al., 2011).

Acknowledgements

This research was supported by China Postdoctoral Science Foundation (2015M571763) and the Special Fund for Agroscientific Research in the Public Interest of China (201303022). We thank Kui Wu and Zhi-hui Yan (Nanjing Agricultural University, China) for providing helps on plant cultivation. Thanks are also due to the reviewers and editors for their helpful comments and English polish on earlier drafts of the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.pld.2017.08.004.

References

Baird, N.A., Etter, P.D., Atwood, T.S., et al., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3, e3376.
Barchi, L., Lanteri, S., Portis, E., et al., 2011. Identification of SNP and SSR markers in eggplant using RAD tag sequencing. BMC Genomics 12.
Catchen, J.M., Amores, A., Hohenlohe, P., et al., 2011. Stacks: building and genotyping loci De Novo from short-read sequences. G3 Genes Genomes Genet. 1, 171–182.
Danquah, E.Y., Hanley, S.J., Brookes, R.C., et al., 2002. Isolation and characterization of microsatellites in *Echinochloa* (L.) Beauv. spp. Mol. Ecol. Notes 2, 54–56.
Flora of China, 2015. Available from: www.efloras.org.
