

# A high-throughput assay for rapid and simultaneous analysis of perfect markers for important quality and agronomic traits in rice using multiplexed MALDI-TOF mass spectrometry

Ardashir K. Masouleh<sup>1</sup>, Daniel L. E. Waters<sup>1</sup>, Russel F. Reinke<sup>2</sup> and Robert J. Henry<sup>1,\*</sup>

<sup>1</sup>Centre for Plant Conservation Genetics, Southern Cross University, Lismore, NSW 2480, Australia

<sup>2</sup>Yanco Agricultural Institute, Yanco, NSW 2703, Australia

Received 19 December 2008;

revised 3 February 2009;

accepted 4 February 2009

\*Correspondence (fax +61 266222080;  
e-mail robert.henry@scu.edu.au)

## Summary

The application of single nucleotide polymorphisms (SNPs) in plant breeding involves the analysis of a large number of samples, and therefore requires rapid, inexpensive and highly automated multiplex methods to genotype the sequence variants. We have optimized a high-throughput multiplexed SNP assay for eight polymorphisms which explain two agronomic and three grain quality traits in rice. Gene fragments coding for the agronomic traits plant height (semi-dwarf, *sd-1*) and blast disease resistance (*Pi-ta*) and the quality traits amylose content (*waxy*), gelatinization temperature (*alk*) and fragrance (*fgr*) were amplified in a multiplex polymerase chain reaction. A single base extension reaction carried out at the polymorphism responsible for each of these phenotypes within these genes generated extension products which were quantified by a matrix-assisted laser desorption ionization-time of flight system. The assay detects both SNPs and indels and is co-dominant, simultaneously detecting both homozygous and heterozygous samples in a multiplex system. This assay analyses eight functional polymorphisms in one 5 µL reaction, demonstrating the high-throughput and cost-effective capability of this system. At this conservative level of multiplexing, 3072 assays can be performed in a single 384-well microtitre plate, allowing the rapid production of valuable information for selection in rice breeding.

**Keywords:** matrix-assisted laser desorption ionization-time of flight (MALDI-TOF), multiplex perfect markers, rice, single nucleotide polymorphism.

## Introduction

Single nucleotide polymorphisms (SNPs) are the most abundant class of sequence variation, and explain the occurrence of human genetic disease (Shastry, 2002) and many important traits in plants (Bryan *et al.*, 2000; Kennedy *et al.*, 2006). The high frequency of SNPs in many plant species, including rice, where comparison of data from *japonica* and *indica* cultivars identified one SNP every 170 bp and one indel every 540 bp (Yu *et al.*, 2002), in combination with their genome-wide distribution (Garg *et al.*, 1999; Drenkard *et al.*, 2000; Nasu *et al.*, 2002; Batley *et al.*, 2003), means that they have the capacity to generate high-resolution genetic maps (Bhatramakki *et al.*, 2002). The capacity for high resolution means that

SNP markers are an attractive tool for gene identification. When identified, causal SNPs are the perfect markers within marker-assisted selection programs (Gupta *et al.*, 2001; Rafalski, 2002; Batley *et al.*, 2003).

Several techniques have been developed to assay SNPs, including SNP microarray hybridization-based methods (Rapley and Harbron, 2004) and enzyme-based methods including those involving the use of DNA ligase, polymerase and nuclease (McGuigan and Ralston, 2002; Olivier, 2005; Costabile *et al.*, 2006; Gunderson *et al.*, 2006). Other methods, such as Pyrosequencing (Ahmadian *et al.*, 2000), *Taqman* (Livak, 1999) and polymerase chain reaction (PCR)-based approaches (Hayashi *et al.*, 2004), have been designed for SNP and indel detection; however, they are

**Table 1** MassARRAY markers for eight different functional polymorphisms

Polymorphism	Trait	Capture primers	Extension primer	Expected polymorphism
<i>sd-1</i> SNP	Semi-dwarf	F:* CGATGTTGATGACCATGGCG R:* CATCCTCCTCCAGGACGAC	AGGACGACGTCGGCGGC	[C/T]
<i>sd-1</i> Del	Semi-dwarf†	F: CACGCACGGGTTCTCCAG R: AGGAGTCCATGATCGTCAG	GCGACAGCTCCTTCATCTCCTCGC	[C/T/A]
<i>Pi-ta</i>	Blast resistance	F:* GCTTCTTTCTTCTGCGG R:* CAAACAATCATCAAGTCAGG	AAGTCAGGTTGAAGATGCATAG	[G/T]
<i>waxy</i> IN1	Amylose content	F:* GATCGATCTGAATAAGAGGG R:* CTGCTTGTGTTGTTCTGTTG	CAGGAAGAACATCTGCAAG	[G/T]
<i>waxy</i> EX6	Amylose content	F:* ACCTCAACAACAACCCATAC R:* GATCATCATGGATTCTTCG	CCCATACTCAAAGGAAGCTT	[C/A]
<i>alk3</i>	Gelatinization temperature	F:* TGTCTCGAACGGGTCGAAC R:* CTCAACCAGCTCTACGCCAT	CTTCTGCGGGCTGAGGGACACC	[A/G]
<i>alk4</i>	Gelatinization temperature	F:* TGACAAGGACCTCCTCGTAG R:* CGCAAGTACAAGGAGAGCTG	AAGGAGAGCTGGAGGGG	[G/C/TT]
<i>Fgr</i>	Fragrance	F:* ACCTCAACAACAACCCATAC R:* GTTAGGTTGCATTACTGGG	TGGGAGTTATGAACTGGTA	[TATAT/AAAAGATTATGGC]

\*A 10-mer tag, sequence 5'-ACGTTGGATG-3', was added to the 5' end of each amplification primer to avoid confusion in the mass spectrum and improve polymerase chain reaction performance.

†A modified method was applied to amplify this allele.

generally not cost- or time-effective per sample. PCR-based markers are preferable because they are efficient, cost-effective and require only a small quantity of genomic DNA for genotyping, and are thus suitable at all stages of plant growth, including early seedling stages.

An increasing number of genes controlling important traits in plants are being discovered, and the underlying polymorphisms can be converted into perfect molecular markers. Some recent examples of perfect markers for important traits in plants include rice fragrance (Bradbury *et al.*, 2005), wheat grain hardness (Morris, 2002), rice blast resistance (Kennedy *et al.*, 2006) and a range of other disease resistance genes (Jeong *et al.*, 2002), however, each of these have been single-trait, uniplex assays. Plant breeders often track and select for more than one trait within any one cross, and as the number of genes which control important traits expands, the need for rapid, simple, inexpensive, reliable multiplex genotyping methods will become more urgent (Hayashi *et al.*, 2004).

The objective of this study was to investigate the capability of the multiplex matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry system (Sequenom® MassARRAY®, San Diego, CA, USA) as a high-throughput platform for the rapid, simultaneous and robust multiplex assay of SNPs responsible for important agronomic and grain quality traits in rice. In this article, we report an assay for distinguishing between eight different important polymorphisms simultaneously in a single 5 µL reaction.

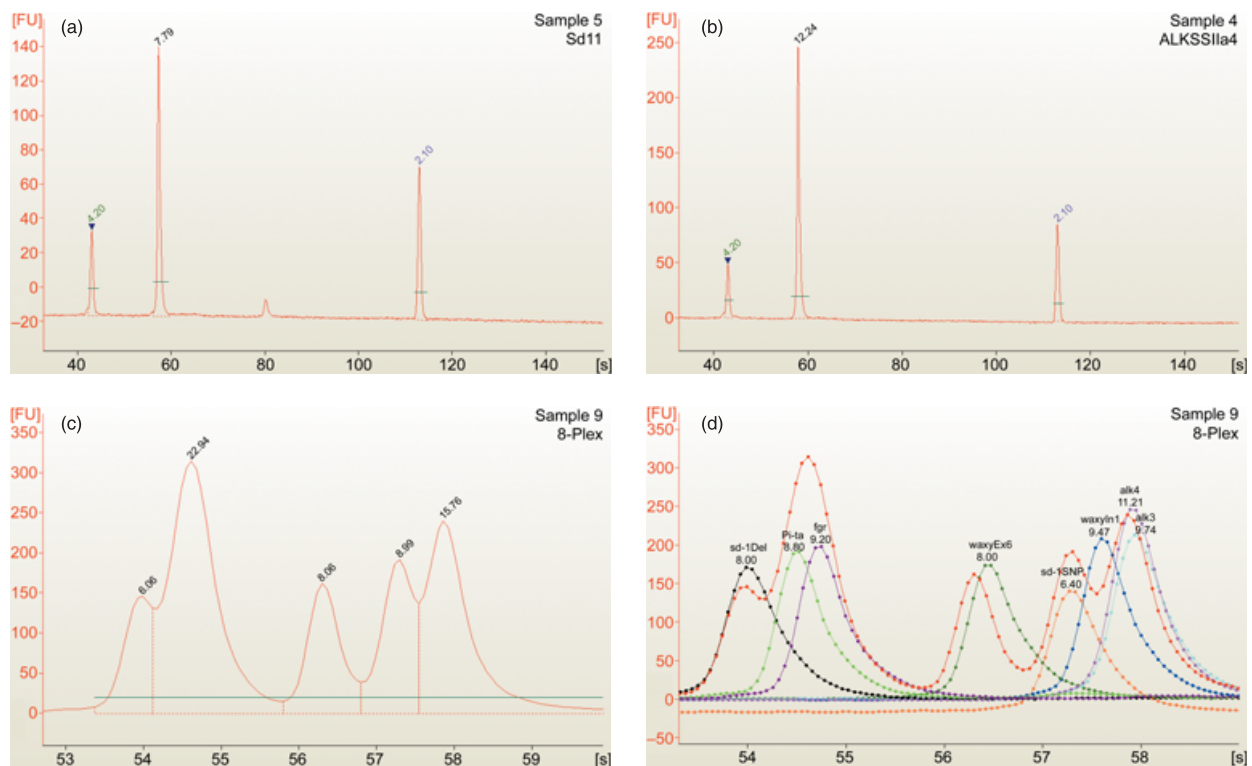
## Results

### Analysis of PCR products

Assays were constructed for eight polymorphisms defining each of the alleles of five genes controlling five important commercial traits. The traits and genes were semi-dwarf (*sd-1*, two alleles) (Sasaki *et al.*, 2002; Spielmeier *et al.*, 2002), blast disease resistance (*Pi-ta*, one allele) (Bryan *et al.*, 2000), amylose content (*waxy*, two SNPs) (Cai *et al.*, 1998; Larkin and Park, 1999; Chen *et al.*, 2008), gelatinization temperature (*alk*, two SNPs) (Umamoto, 2005; Waters *et al.*, 2006) and fragrance (*fgr*, one allele) (Bradbury *et al.*, 2005) (Table 1).

### Optimal capture primer concentration

The optimal primer concentration for the amplification of each target polymorphism in uniplex and eight-plex was 0.3 µM. Polymorphism detection at eight-plex was consistent with uniplex data. Increasing the uniplex primer concentration to 0.5 µM led to PCR products of higher concentration, except for *waxy*IN1, in which there was nonspecific amplification at this concentration. The concentrations of PCR products, as measured by a Bioanalyser 2100 (Agilent Technologies, Palo Alto, CA, USA) DNA 500 LabChip® Kit, ranged from 7.8 ng/µL (*sd-1* SNP) to 12.2 ng/µL (*alk4*) in uniplex (Figure 1a,b), and were relatively lower in eight-plex, ranging from



**Figure 1** Concentration of polymerase chain reaction (PCR) products in uniplex and eight-plex. (a) Concentration of *sd-1* SNP = 7.8 ng/μL (major peak); minor peaks correspond to size standard. (b) Concentration of *alk4* = 12.2 ng/μL (major peak); minor peaks correspond to size standard. (c) Concentration of PCR products in eight-plex. (d) Concentration of PCR products in eight-plex analysed individually (all in ng/μL).

6.40 ng/μL (*sd-1* SNP) to 11.21 ng/μL (*alk4*), which was sufficient to produce an excellent mass spectrum (Figure 1c,d).

### MgCl<sub>2</sub> concentration

The MgCl<sub>2</sub> concentration is one of the most important factors for accurate concurrent amplification of different loci in a multiplex system. The optimal concentration for the amplification of all loci in uniplex and multiplex was 3 mM. At this MgCl<sub>2</sub> concentration, all target loci were amplified free from nonspecific amplicons and primer dimers. At lower MgCl<sub>2</sub> concentrations of 2 and 2.5 mM, no target DNA was amplified and there were a surprising number of nonspecific bands and primer dimers. At concentrations higher than 3 mM, nonspecific bands were present in addition to the target loci. These results were consistent and reproducible in both uniplex and eight-plex.

### Identification of SNPs and polymorphisms in agronomic and quality loci

All eight loci were amplified in 25 cultivars and genotyped by multiplex MALDI-TOF analysis of single-base extension

products, and the polymorphisms were compared (Table 2). Of these, three were responsible for important agronomic traits and five for grain quality traits, including six nucleotide substitutions and two insertions/deletions (indels). Polymorphisms were distinguished at all agronomic and quality loci, as described below.

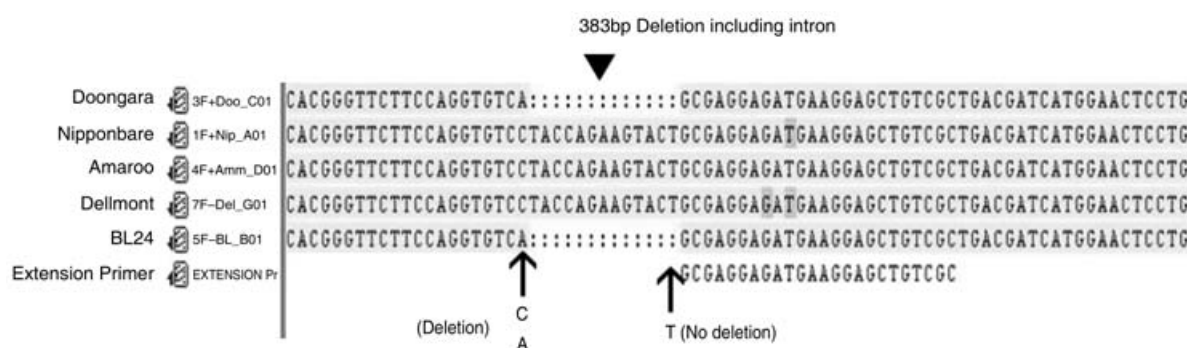
#### *sd-1*

The semi-dwarf phenotype is caused by a loss of function of the enzyme gibberellin 20-oxidase (GA 20-oxidase). A plant carrying the non-functional form of the gene, *sd-1*, which codes for this enzyme, has a diminished capacity to produce gibberellin, resulting in a reduced plant height and enhanced grain yield. Two alleles of *sd-1* were assayed. One *sd-1* allele, here called *sd-1*SNP, contains a C/T SNP in exon 2 of the gene (CTC = leucine/TTC = phenylalanine) (Spielmeyer *et al.*, 2002; Monna *et al.*, 2002). The other allele, here called *sd-1*Del, is characterized by a 280-bp (Spielmeyer *et al.*, 2002) or 278-bp (Sasaki *et al.*, 2002) deletion of part of exon 1 and exon 2 and 102–105 bp of the intron sequence, a 380–383-bp deletion in total (Figure 2).

Although a large deletion, such as *sd-1*Del, can be determined by the size difference of amplification products on a

**Table 2** Single nucleotide polymorphisms (SNPs) of 25 commercial rice cultivars at eight different functional loci

Cultivar	Polymorphism							
	<i>sd-1</i> SNP	<i>sd-1</i> Del	<i>Pi-ta</i>	<i>waxy</i> IN1	<i>waxy</i> EX6	<i>alk3</i>	<i>alk4</i>	<i>fgr</i>
Amaroo	T	A	T	T	A	G	TT	AAAGATT
Amber	C	A	T	G	C	G	GC	TATAT
Basmati 370	C	A	T	G	C	G	GC	TATAT
BL24	C	T	G	G	A	G	GC	AAAGATT
Calrose	C	A	T	T	A	G	TT	AAAGATT
Calmochi 202	T	A	T	T	A	G	TT	AAAGATT
Dawn	C	A	T	G	C	G	GC	AAAGATT
Della	C	A	T	G	C	G	GC	TATAT
Dellmont	C	T	T	G	C	G	GC	TATAT
Domsorkh	C	A	T	G	C	G	GC	TATAT
Doongara	C	T	T	G	C	G	GC	AAAGATT
Dragon Eye Ball	C	A	T	T	A	A	GC	TATAT
Goolarah	C	A	T	T	A	G	GC	TATAT
Jarra	T	A	T	T	A	G	TT	AAAGATT
Jasmin	T	T	T	T	A	G	TT	TATAT
Kyeema	C	A	T	T	A	G	GC	TATAT
Khao Dawk Mali 105	C	A	T	T	A	G	TT	TATA
L 202	C	T	T	T	A	G	GC	AAAGATT
Langi	T	A	T	T	A	G	GC	AAAGATT
Millin	T	A	T	T	A	A	GC	AAAGATT
M7	C	T	T	T	A	G	GC	AAAGATT
Nipponbare	C	A	T	T	A	A	GC	AAAGATT
Opus	T	A	T	T	A	A	GC	AAAGATT
Teqing	C	T	G	G	A	G	GC	AAAGATT
YRF 204	C	T	T	T	A	G	GC	TATAT

**Figure 2** Determination of *sd-1*Del gene by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry. There is a 383-bp deletion in semi-dwarf plants; therefore, the extended single base (mass-modified terminator) matches to 'C' or 'A', which is located just after deletion; otherwise, there is a T peak for tall plants.

simple 2% agarose gel (Figure 3), we assessed the suitability of MALDI-TOF for the identification of large indels. In theory, only one base (terminator) is added to the SNP site downstream of the extension primer. Therefore, accurate gene sequence information, particularly the flanking region just before and after the indel, is necessary because single-base extension either recognizes one base inside or outside of the indel. Theoretically, the indel can be determined by the ddNTP

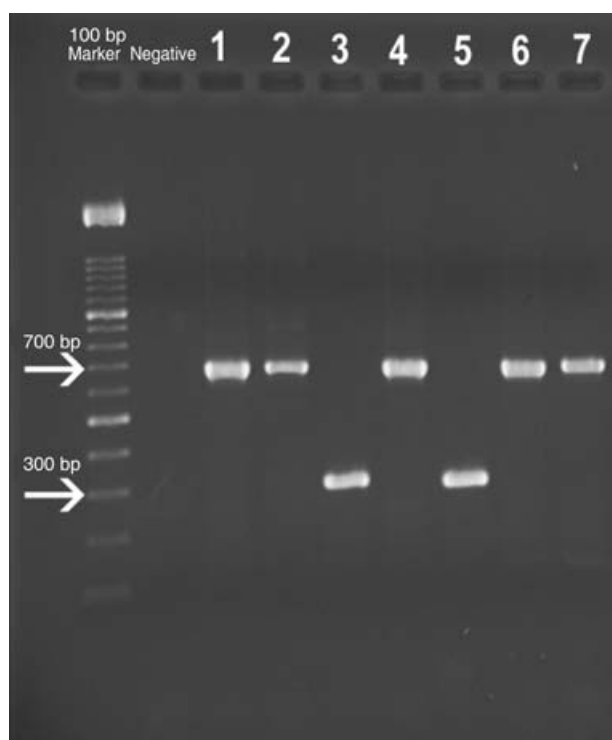
which terminates the extension reaction. However, when using the assay designed by Sequenom® (MassARRAY® Assay design 3.1) in both uniplex and eight-plex, no logic call was detected and all genotypes showed 'A', which corresponds to the *sd-1*Del allele. Modification of the method substantially improved the accuracy of analysis of this allele, from 43.7% to 5% (Table 3). The modification involved amplification of the region containing the deletion in uniplex using

**Table 3** Percentage of missing data in uniplex and eight-plex and apparent heterozygosity in eight-plex

Plex level	Assays/single nucleotide polymorphism							
	<i>sd-1</i> SNP	<i>sd-1</i> Del	<i>Pi-ta</i>	<i>waxy</i> IN1	<i>waxy</i> EX6	<i>alk3</i>	<i>alk4</i>	<i>fgr</i>
Missing data uniplex*	0%	8%	0%	0%	0%	1.1%	0%	0%
Missing data eight-plex*	4.5%	43.7%	0%	0%	4.2%	3.1%	0%	0%
Missing data modified eight-plex†	4.5%	5%	0%	0%	4.2%	3.1%	0%	0%
Apparent heterozygosity eight-plex and modified eight-plex	3.9%	0%	0%	0%	0%	0%	3.1%	0%

\*Assays designed with Sequenom® MassARRAY® Assay design 3.1.

†Assays designed with Sequenom® MassARRAY® Assay design 3.1, except for *sd-1*Del where polymerase chain reaction primers were designed by Primer 3, *sd-1*Del amplified in uniplex, and extended and analysed in eight-plex.



**Figure 3** Determination of *sd-1*Del gene on a 2% agarose gel. Fragments around 300 bp indicate 383-bp deletion in the *sd-1*Del gene, which is responsible for the semi-dwarf phenotype. Fragments of approximately 700 bp are the intact *sd-1*Del gene of tall plants. Lanes from left to right: 100-bp ladder; negative control; rice varieties Nipponbare; Kyeema; Doongara; Amaro; BL24; Della; Domsorkh.

PCR primers designed by Primer 3 (<http://frodo.wi.mit.edu>), followed by the addition of these uniplex amplicons to the other loci which had been amplified in seven-plex for all subsequent manipulations.

#### *Pi-ta*

*Pi-ta* is a major blast resistance gene in rice. *Pi-ta* encodes a 928-amino-acid polypeptide with a molecular mass of

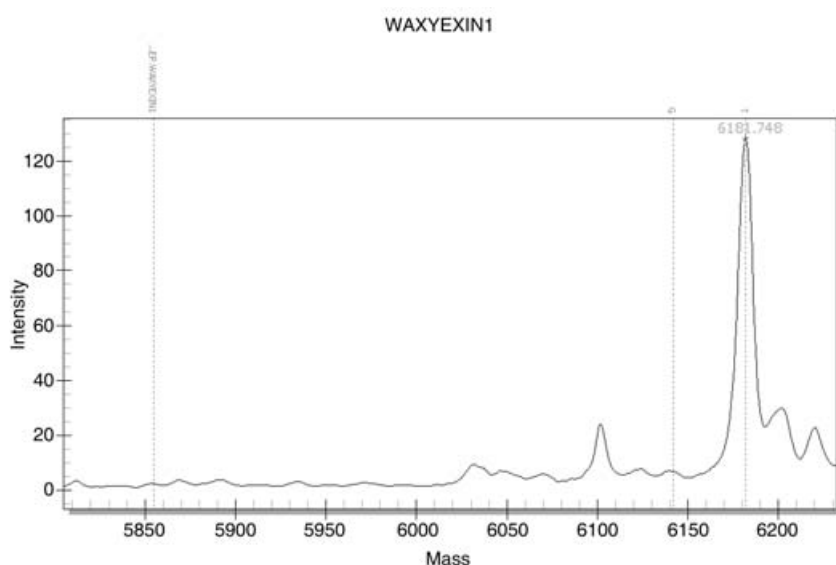
105 kDa. A [G/T] SNP distinguishes susceptible and resistant genotypes (Bryan *et al.*, 2000); amino acid 918 differs between resistant and susceptible genotypes: all susceptible genotypes have a serine (T) at this position, whereas resistant plants have alanine (G). Most of the cultivars in this study carried the 'T' allele that translates to serine (susceptible), whereas BL24 and Teqing contained the resistant 'G' allele (alanine).

#### *waxy*

The *waxy* gene encodes the enzyme granule-bound starch synthase, which is one of the key factors influencing rice starch quality by affecting apparent amylose content (Sano, 1984; Webb, 1991; Chen *et al.*, 2008). The [G/T] SNP at the intron 1/exon 1 splice site (*waxy*IN1) differentiates between varieties of high and low amylose content (Cai *et al.*, 1998) and, in combination with the exon 6 [C/A] SNP (*waxy*EX6), differentiates between varieties of high, intermediate and low amylose content in southern US germplasm (Chen *et al.*, 2008). Cultivars with 'T' in *waxy*IN1 and 'A' in *waxy*EX6 have the lowest amylose content or even glutinous starch. High polymorphism was found at *waxy*IN1 in the studied cultivars Amber, Basmati 370, BL24, Dawn, Della, Dellmont and Domsorkh, Doongara and Teqing contained the 'G' allele, and Jasmine, Nipponbare, Langi and M7 carried the 'T' allele (Figure 4). At *waxy*EX6, 18 of the 25 cultivars displayed 'A', which suggests low amylose content.

#### *alk*

The major gene regulating alkali disintegration in rice grains, *alk* (Gao *et al.*, 2003), encodes the enzyme starch synthase IIa (Umamoto *et al.*, 2004). Alkali disintegration is a convenient indirect measure of the gelatinization temperature of rice starch, which is, in turn, associated with rice cooking and eating quality. Two polymorphisms within exon 8 of *alk*, [A/G] (*alk3*) and [GC/TT] (*alk4*), are associated with gelatinization



**Figure 4** Sequenom® MassARRAY® waxyIN1 uniplex spectrum for cv. Langi which shows a peak for 'T'.

temperature class (Umemoto, 2005; Waters *et al.*, 2006). A combination of *alk3* 'G' and *alk4* 'GC' is found within varieties of high gelatinization temperature and low alkali spreading, whereas varieties with either *alk3* 'A' or *alk4* 'TT' are low gel temperature varieties. Both the [GC/TT] (*alk4*) and [A/G] (*alk3*) polymorphisms were determined in all cultivars.

#### *fgr*

A recessive gene (*fgr*) on chromosome 8 controls rice fragrance. The intact *Fgr* allele encodes a betaine aldehyde dehydrogenase (BADH2) in non-fragrant rice, whereas fragrant rice contains an 8-bp deletion and three SNPs which prematurely terminate the translation of BADH2. This changes the biosynthetic pathways in which BADH2 is active, resulting in the accumulation of 2-acetyl-pyrroline, which is responsible for fragrance (Bradbury *et al.*, 2005). The eight-plex assay identified 11 varieties with the fragrant allele *fgr*.

#### Missing data and heterozygosity

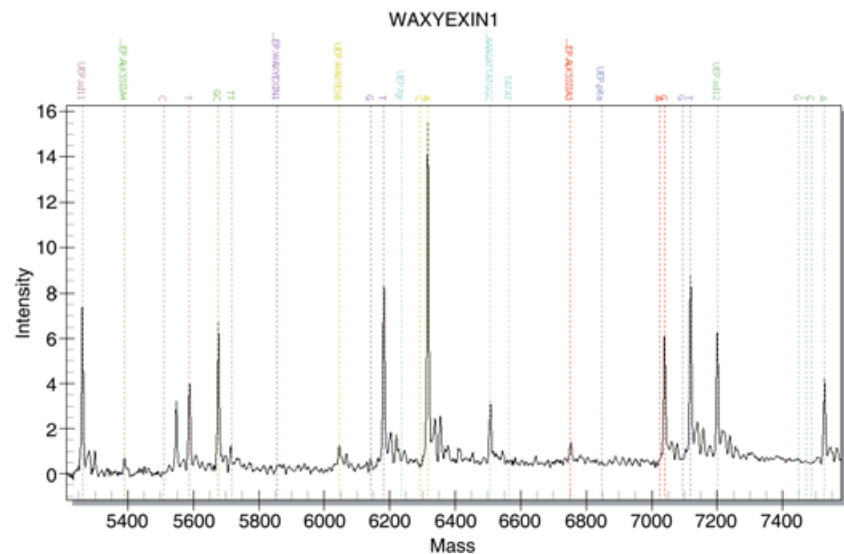
The highest rate of missing data belonged to *sd-1DEL* in eight-plex, which suggests that this allele is not compatible with the multiplex system (Table 3). No missing data were found in *waxyIN1*, *Pi-ta*, *alk4* and *fgr*. The apparent heterozygosity values were 3.9% and 3.1% in *sd-1SNP* and *alk4*, respectively.

## Discussion

We have demonstrated that DNA polymorphisms can be efficiently confirmed and analysed in rice using a MALDI-TOF mass spectrometry system (Ding and Cantor, 2003). These assays can be used as a marker-assisted selection tool in con-

ventional breeding programs. Rice has been at the forefront of the application of genomics and genomics tools to plant breeding and serves as a model for other crops. A whole rice genome sequence has been available for several years (Goff *et al.*, 2002; Yu *et al.*, 2002), and a comprehensive DNA polymorphism database has recently become available online (<http://irfgc.irri.org/index.php>). The availability of these resources has accelerated the rate at which gene function has been elucidated. Emerging DNA sequencing technologies are revolutionizing the field of genomics, bringing the reality of relatively inexpensive comparative genome sequencing of all the major crops much closer. MALDI-TOF mass spectrometry, in combination with comparative genome sequence data, will become increasingly useful in marker-assisted breeding as more genes that control important traits are identified.

An efficient PCR is the most important predictor for producing a reliable and consistent assay on this platform (Figure 5). The uniform simultaneous amplification of all loci will resolve the most commonly encountered problems (Siebert and Larrick, 1992). The number and intensity of correct SNP calls are increased with higher PCR product concentrations. The minimum concentration of PCR product is 4 ng/μL for loci which fall within the default size of 80–120 bp; however, longer PCR products require a higher concentration as measured by mass to maintain the molar concentration at acceptable levels for iPLEX extension reactions. The concentration of PCR products differs between uniplex and eight-plex systems, which may have an effect on peak height calls. These differences are a result of competition between each PCR in multiplex, and show 5.7%–17.8% reductions in the final eight-plex PCR assay compared with the uniplex assay.



**Figure 5** An eight-plex Sequenom® MassARRAY® spectrum for cv. Langi.

Even spectral peak heights (Figure 5) are critical for accurate genotype calls using MALDI-TOF mass spectrometry, and this is achieved by increasing the concentration of individual extension primers, not by modifying capture PCR conditions, because this does not have a significant effect on the final spectra. PCR yield is intrinsic to the PCR conditions and, when optimized, should be adhered to; increasing the concentration of template, primer and *Taq* enzyme above that recommended concentration may increase yield in uniplex; however, in multiplex, it may lead to the generation of dimers and spurious PCR products.

Accurate DNA sequence data for each polymorphism represent the most important prerequisite for accurate assay design. However, public domain databases and published papers can have conflicting data for each locus. For example, three different sequences for *sd-1Del* appear in the public domain: the deletion has been reported to be 382 bp (Spielmeyer *et al.*, 2002) or 383 bp (Monna *et al.*, 2002; Sasaki *et al.*, 2002), and differs by the length of intron and the exact location of the deletion. In cases such as this, re-sequencing the target region is necessary for accurate primer design, which ultimately leads to an accurate, consistent assay.

The capture PCR stage is important in uniplex reactions, but it is critical in the multiplex system because of the high rate of competition between primers consuming templates and enzyme. Some primers worked well in uniplex, but had missing calls in multiplex, suggesting that there were interactions between primers in eight-plex (Table 3). For example, interactions between *waxyIN1* and *fgr* increased the number of missing calls in eight-plex. There was, however, a high correlation of more than 98% between uniplex and eight-plex calls, and missing calls were around 0.15% and 1.68% (not

including *sd-1Del*) for uniplex and eight-plex respectively, which compares favourably with other sequencing methods (Jones *et al.*, 2007).

Multiplex MALDI-TOF is a powerful tool for the detection and confirmation of SNPs in rice. It has been suggested that this platform has the capability of determining more than 40 SNPs in multiplex (Sequenom, 2006) and, given that the platform can process ten 384-well plates per day, users can theoretically analyse in excess of 153 000 SNPs daily (Perkel, 2008). This technique can be applied to segregating populations in the early stages of breeding programs to positively select desired polymorphisms and traits, and is a co-dominant system, having the ability to detect alleles in hybrids, heterozygotes (Jones *et al.*, 2007) and polyploids (Henry *et al.*, 2008). The capacity of the system to accurately identify haplotypes at one or more loci, *alk* and *waxy* for example, allows for the efficient selection of target phenotypes within breeding programs.

## Experimental procedures

### Genotypes

All plant material was supplied by the Australian Plant DNA Bank (<http://www.biobank.com>). Twenty-five commercial rice cultivars were analysed: Amaroo, Amber, Basmati 370, BL24, Calrose, Calmochi 202, Dawn, Della, Dellmont, Domsorkh, Doongara, Dragon Eye Ball, Goolarah, Jarrah, Jasmine, Kyeema, Khao Dawk Mali 105, L202, Langi, Millin, M7, Nipponbare, Opus, Teqing and YRF204.

### DNA extraction

Total plant DNA was extracted from individual seedlings at 10 days after germination using a Qiagen (Valencia, CA, USA) DNeasy Plant Kit, according to the manufacturer's instructions.

### Primer design/generation of SNP markers

Capture and extension primers were designed by Sequenom® MassARRAY® Assay design 3.1 software, with the exception of the *sd-1DEL* primers which were designed by Primer 3 (<http://frodo.wi.mit.edu>). The optimal amplicon size containing the polymorphic site in the software was set to 80–120 bp. A 10-mer tag (5-ACGTTGGATG-3) was added to the 5' end of each amplification primer to avoid confusion in the mass spectrum and to improve PCR performance.

### Capture PCR protocol

Platinum® *Taq* DNA Polymerase (Invitrogen, Carlsbad, CA, USA) in a final volume of 5 µL was used for all capture PCRs. The eight-plex reaction was optimized by testing a number of capture primer and MgCl<sub>2</sub> concentrations in the ranges 0.2–1 µM and 1–3.5 µM, respectively. Uniplex assays using identical PCR conditions confirmed the results of all eight-plex experiments. The optimal eight-plex capture PCR consisted of 3–5 ng of template DNA, 0.5 µL 10 × PCR buffer (Invitrogen), 3 mM MgCl<sub>2</sub>, 2.5 mM of each deoxynucleoside triphosphate (dNTP), 5 µM of each primer and 1 unit of *Taq* polymerase (5 U/µL). The reactions were heated to 94 °C for 15 min, followed by 45 cycles of amplification at 94 °C for 20 s, 56 °C for 30 s and 72 °C for 1 min, followed by a final extension at 72 °C for 3 min.

As the *sd-1DEL* is relatively large, the amplification protocol was modified as follows: 3.75 µL of 10 × PCR buffer (50 mM), 2.25 µL of MgCl<sub>2</sub> (50 mM), 2.1 µL of 10 µM primers (each), 6 µL of dNTPs (2.5 mM), 12 µL of 2 × Enhancer [6% glycerol + 10% dimethyl sulphoxide (DMSO)], 0.3 µL of Platinum® *Taq* polymerase and 1.5 µL (5 U/µL) template. The thermocycling program was 94 °C for 5 min, followed by 45 cycles of amplification at 94 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min, followed by a final extension of 72 °C for 3 min. Finally, 1 µL of PCR product was added to the multiplex test tubes.

### Shrimp alkaline phosphatase (SAP) incubation

Unincorporated dNTPs were removed by SAP incubation according to the manufacturer's (Sequenom, San Diego, CA, USA) instructions.

### Primer extension and mass spectrometry

The remaining assay steps of primer extension, resin cleanup and mass spectrometry were undertaken according to the manufacturer's (Sequenom® MassARRAY®) instructions.

### Acknowledgements

The authors wish to thank Peter Bundock, Timothy Fitzgerald, Julie Pattemore and Timothy Sexton in the Centre for Plant Conservation Genetics for their valuable assistance in providing plant material and helpful discussions. We also gratefully acknowledge Stirling Bowen in Southern Cross Plant Genomics for his technical advice and assistance. The Australian Research Council Funded this work.

### References

- Ahmadian, A., Gharizadeh, B., Gustafsson, A.C., Sterky, F., Nyrén, P., Uhlén, M. and Lundeberg, J. (2000) Single-nucleotide polymorphism analysis by pyrosequencing. *Anal. Biochem.* **280**, 103–110.
- Batley, J., Mogg, R., Edwards, D., O'Sullivan, H. and Edwards, K.J. (2003) A high-throughput SNUPE assay for genotyping SNPs in the flanking regions of *Zea mays* sequence tagged simple sequence repeats. *Mol. Breed.* **11**, 111–120.
- Bhatramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register, J.C., Tingey, S.V. and Rafalski, A. (2002) Insertion–deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Mol. Biol.* **48**, 539–547.
- Bradbury, L.M.T., Fitzgerald, T.L., Henry, R.J., Jin, Q. and Waters, D.L.E. (2005) The gene for fragrance in rice. *Plant Biotechnol. J.* **3**, 363–370.
- Bryan, G.T., Wu, K.S., Farrall, L., Jia, Y., Hershey, H.P., McAdams, S.A., Faulk, K.N., Donaldson, G.K., Tarchini, R. and Valent, B. (2000) A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. *Plant Cell*, **12**, 2033–2046.
- Cai, X.L., Wang, Z.Y., Xing, Y.Y., Zhang, J.L. and Hong, M.M. (1998) Aberrant splicing of intron 1 leads to the heterogeneous 5' UTR and decreased expression of *waxy* gene in rice cultivars of intermediate amylose content. *Plant J.* **14**, 459–465.
- Chen, M.H., Bergman, C.J., Pinson, S.R.M. and Fjellstrom, R.G. (2008) *Waxy* gene haplotypes: associations with pasting properties in an international rice germplasm collection. *J. Cereal Sci.* **48**, 781–788.
- Costabile, M., Quach, A. and Ferrante, A. (2006) Molecular approaches in the diagnosis of primary immunodeficiency diseases. *Hum. Mutat.* **27**, 1163–1173.
- Ding, C. and Cantor, C.R. (2003) A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS. *Proc. Natl. Acad. Sci. USA*, **100**, 3059–3064.
- Drenkard, E., Richter, B.G., Rozen, S., Stutius, L.M., Angell, N.A., Mindrinos, M., Cho, R.J., Oefner, P.J., Davis, R.W. and Ausubel, F.M. (2000) A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. *Plant Physiol.* **124**, 1483–1492.
- Gao, Z.Y., Zeng, D.L., Cui, X., Zhou, Y.H., Yan, M.X., Huang, D.N., Li, J.Y. and Qian, Q. (2003) Map-based cloning of the *ALK* gene, which controls the gelatinization temperature of rice. *Sci. Chin. Ser. C, Life Sci.* **46**, 661–668.
- Garg, K., Green, P. and Nickerson, D.A. (1999) Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. *Genome Res.* **9**, 1087–1092.
- Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P. and Varma, H. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science*, **296**, 92–100.
- Gunderson, K.L., Steemers, F.J., Ren, H., Ng, P., Zhou, L., Tsan, C., Chang, W., Bullis, D., Musmacker, J. and King, C. (2006) Whole-genome genotyping. *Methods Enzymol.* **410**, 359–376.
- Gupta, P.K., Roy, J.K. and 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.
- Hayashi, K., Hashimoto, N., Daigen, M. and Ashikawa, I. (2004) Development of PCR-based SNP markers for rice blast resistance genes at the *Piz* locus. *Theor. Appl. Genet.* **108**, 1212–1220.
- Henry, R.J., Pattermore, J.A., Waters, D.L.E., Kharabian Masouleh, A., Bundock, P.C. and Elliott, F.G. (2008) Applications of the sequenom platform to SNP analysis in plants. In: *Plant and Animal Genomes XVI Conference*. San Diego, CA: Town and Country Convention Center.
- Jeong, S.C., Kristipati, S., Hayes, A.J., Maughan, P.J., Noffsinger, S.L., Gunduz, I., Buss, G.R. and Maroof, M.A.S. (2002) Genetic and sequence analysis of markers tightly linked to the soybean mosaic virus resistance gene, *Rsv 3*. *Crop Sci.* **42**, 265–270.
- Jones, E.S., Sullivan, H., Bhatramakki, D. and Smith, J.S.C. (2007) A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). *Theor. Appl. Genet.* **115**, 361–371.
- Kennedy, B.G., Waters, D.L.E. and Henry, R.J. (2006) Screening for the rice blast resistance gene *Pi-ta* using LNA displacement probes and real-time PCR. *Mol. Breed.* **18**, 185–193.
- Larkin, P.D. and Park, W.D. (1999) Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. *Plant Mol. Biol.* **40**, 719–727.
- Livak, K.J. (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genetic analysis. *Biomol. Eng.* **14**, 143–149.
- McGuigan, F.E. and Ralston, S.H. (2002) Single nucleotide polymorphism detection: allelic discrimination using TaqMan. *Psychiatr. Genet.* **12**, 133–136.
- Monna, L., Kitazawa, N., Yoshino, R., Suzuki, J., Masuda, H., Maehara, Y., Tanji, M., Sato, M., Nasu, S. and Minobe, Y. (2002) Positional cloning of rice semidwarfing gene, *sd-1*: 'Rice green revolution gene' encodes a mutant enzyme involved in gibberellin synthesis. *DNA Res.* **9**, 11–17.
- Morris, C.F. (2002) Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Mol. Biol.* **48**, 633–647.
- Nasu, S., Suzuki, J., Ohta, R., Hasegawa, K., Yui, R., Kitazawa, N., Monna, L. and Minobe, Y. (2002) Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. *DNA Res.* **9**, 163–171.
- Olivier, M. (2005) The Invader® assay for SNP genotyping. *Mutat Res.* **573**, 103–110.
- Perkel, J. (2008) SNP genotyping: six technologies that keyed a revolution. *Nat Methods*, **5**, 447–453.
- Rafalski, A. (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **5**, 94–100.
- Rapley, R. and Harbron, S.E. (2004) *Molecular Analysis and Genome Discovery*. Sussex, UK: Wiley.
- Sano, Y. (1984) Differential regulation of waxy gene expression in rice endosperm. *Theor. Appl. Genet.* **68**, 467–473.
- Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., Ishiyama, K., Saito, T., Kobayashi, M. and Khush, G.S. (2002) A mutant gibberellin-synthesis gene in rice. *Nature*, **416**, 701–702.
- Sequenom, I. (2006) iPLEX™ Gold Assay for SNP Genotyping. *Biotechniques. Protocol Guide*, **41**. San Diego, CA: Sequenom.
- Shastri, B.S. (2002) SNP alleles in human disease and evolution. *J. Hum. Genet.* **47**, 561–566.
- Siebert, P.D. and Larrick, J.W. (1992) Competitive PCR. *Nature*, **359**, 557–558.
- Spielmeier, W., Ellis, M.H. and Chandler, P.M. (2002) Semi-dwarf (*sd-1*), 'green revolution' rice, contains a defective gibberellin 20-oxidase gene. *Proc. Natl. Acad. Sci. USA*, **99**, 9043–9048.
- Umemoto, T. (2005) Single-nucleotide polymorphisms in rice starch synthase IIa that alter starch gelatinisation and starch association of the enzyme. *Funct. Plant Biol.* **32**, 763–768.
- Umemoto, T., Aoki, N., Lin, H., Nakamura, Y., Inouchi, N., Sato, Y., Yano, M., Hirabayashi, H. and Maruyama, S. (2004) Natural variation in rice starch synthase IIa affects enzyme and starch properties. *Funct. Plant Biol.* **31**, 671–684.
- Waters, D.L.E., Henry, R.J., Reinke, R.F. and Fitzgerald, M.A. (2006) Gelatinisation temperature of rice explained by polymorphisms in starch synthase. *Plant Biotechnol. J.* **4**, 115–122.
- Webb, B.D. (1991) Rice quality and grades. In: *Rice* (Bor S. Luh ed.), pp. 89–93. College Station, TX: USDA, Rice Quality Laboratory, Texas A&M University.
- Yu, J., Hu, S., Wang, J., Wong, G.K.S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y. and Zhang, X. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science*, **296**, 79–92.