Proceedings of the 10th Australian Barley Technical Symposium, 2001

Single nucleotide primer extensions to type SNPs in barley

Maxime Paris[1], Reg Lance[2] and Michael G.K. Jones[1]

[1]WA State Agricultural Biotechnology Centre (SABC), Murdoch University, WA 6150
[2] Agriculture Western Australia, South Perth, WA 6151

Abstract

The progression in the discovery of barley Single Nucleotide Polymorphisms (SNPs) has necessitated the development of rapid and low-cost methods enabling their use as molecular markers for selection of traits of interest. Single Nucleotide Primer Extension (SNuPE) assays are used to genotype SNPs from minimal amounts of plant leaf tissue. Following PCR amplification of the target sequence, a genotyping primer with its 3' end directly flanking the SNP is extended by a single dideoxynucleoside triphosphate (ddNTP) complementary to the polymorphic base. Relying on the molecular weight difference between DNA bases, the incorporated nucleotide is identified by the mass increase of the extended primer. Matrix Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-ToF MS) enables absolute molecular weight determination of the extended primers and accurate identification of the targeted base. The method is non-gel based, amenable to automation and multiplexing and allows detection of heterozygotes. We report the application of SNuPE assays for co-dominant Marker-Assisted Selection (MAS) of barley.

Introduction

Single Nucleotide Polymorphisms (SNPs) are the most abundant sequence variations encountered in most genomes (Cho *et al.*, 1999; Griffin and Smith, 2000). Various large-scale discovery projects are currently aiming at identifying SNPs from a broad range of organisms, including crop plants. The abundance, ubiquity and interspersed nature of SNPs make them ideal candidates as molecular markers for marker-assisted plant breeding. While various SNP detection methods have been described (Landegren *et al.*, 1998), our objective was to identify a co-dominant and robust system amenable to multiplexing and automation to genotype SNPs from plants at the seedling stage. We report the development of Single Nucleotide Primer Extension (SNuPE) assays enabling co-dominant genotyping of SNPs from small amounts of barley tissue. The method was used to select barley seedlings carrying the *mlo*

powdery mildew resistance allele as well as superior alleles of β-amylase thermostability based on SNPs discovered within genes encoding these traits.

Materials and Methods

DNA extractions were performed from leaf tips collected from 4 to 10 day-old barley seedlings using the method of Paris and Carter (2000). Target regions containing candidate SNPs were amplified by PCR using flanking primers to generate amplicons of less than 300bp. Because residual primers and dNTPs interfere with the following reaction, these components were dephosphorylated using Shrimp Alkaline Phosphatase (SAP).

Single Nucleotide Primer Extension (SNuPE) assays: for each candidate SNP, an unlabeled genotyping primer with its 3' end directly flanking the SNP was synthesized. Following target amplification and SAP inactivation of residual primers and dNTPs, genotyping primers were annealed to the target sequence and extended by a single ddNTP complementary to the polymorphic nucleotide (e.g. Fig.1).

MALDI-ToF MS analysis: extension products were analysed using a Perseptive Biosystem Voyager Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometer (MALDI-ToF MS).

(a)
5'-AAAAGGGGTGCCGGCG**T**GGGAGCTGCCGGAGACGCCGTC-3'
CCCTCGACGGCCTCT-5'
SNP
Genotyping primer
5'-AAAAGGGGTGCCGGCG**C**GGGAGCTGCCGGAGACGCCGTC-3'
CCCTCGACGGCCTCT-5'

(b)
5'-AAAAGGGGTGCCGGCG**T**GGGAGCTGCCGGAGACGCCGTC-3'
ACCCTCGACGGCCTCT-5'
5'-AAAAGGGGTGCCGGCG**C**GGGAGCTGCCGGAGACGCCGTC-3'
GCCCTCGACGGCCTCT-5'

(c)
Unextended genotyping primer: 5'-TCTCCGGCAGCTCCC = 4465 Da
Genotyping primer + ddATP: 5'-TCTCCGGCAGCTCCC**A** = 4762 Da
Genotyping primer + ddGTP: 5'-TCTCCGGCAGCTCCC**G** = 4778 Da

Figure 1. SNuPE assay used to genotype a SNP identified in the coding region of the barley *mlo* gene. A genotyping primer with its 3' end directly flanking the SNP was annealed to the amplified target (a) and induced to extend by a single ddNTP complementary to the polymorphic base (b). Based on the molecular weight difference between ddNTPs, extension products varied in weight depending on the incorporated nucleotide (c).

Results and Discussion

Mlo: recessive alleles of the barley *Mlo* gene (*mlo*) confer resistance to the fungal pathogen powdery mildew. Following the identification of a SNP in the coding region of *mlo*, a SNuPE assay was designed to select progeny carrying *mlo* (Fig.1) (Paris *et al.*, 2001) thus allowing heterozygous progeny, which are disease susceptible, to be identified (Fig.2).

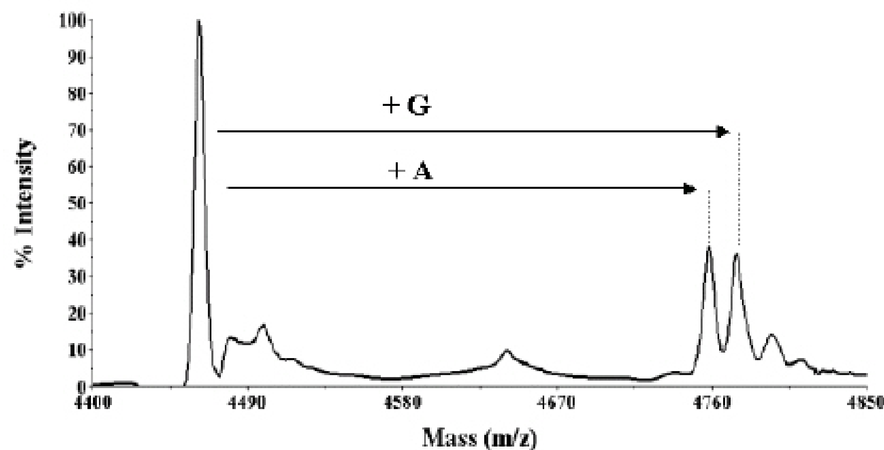

Figure 2. Heterozygous progeny for *Mlo* (*Mlo/mlo*) genotyped using the SNuPE assay illustrated in Figure 1. Samples were analysed by MALDI-ToF MS. Genotyping primers (4465 Da) are extended by a single ddATP (4762 Da) in powdery mildew resistant plants (*mlo*/*mlo*) and extended either by a single ddGTP (4778 Da) or both ddATP and ddGTP in susceptible plants (*Mlo*/*Mlo* or *Mlo/mlo*).

®-amylase: four alleles of β-amylase thermostability are found in barley (Sd2L, Sd1 and Sd2H and, from some *Hordeum spontaneum* lines, Sd3) (Eglinton *et al.*, 1998). Alignment of the amino acid sequences of the four barley β-amylase enzymes revealed a total of eight amino acid substitutions (Eglinton, personal communication) from which SNPs were subsequently deduced. The four alleles could be identified by

genotyping two of these SNPs using a duplex SNuPE assay (Fig.3) hence allowing selection of barley seedlings carrying superior alleles for β-amylase thermostability.

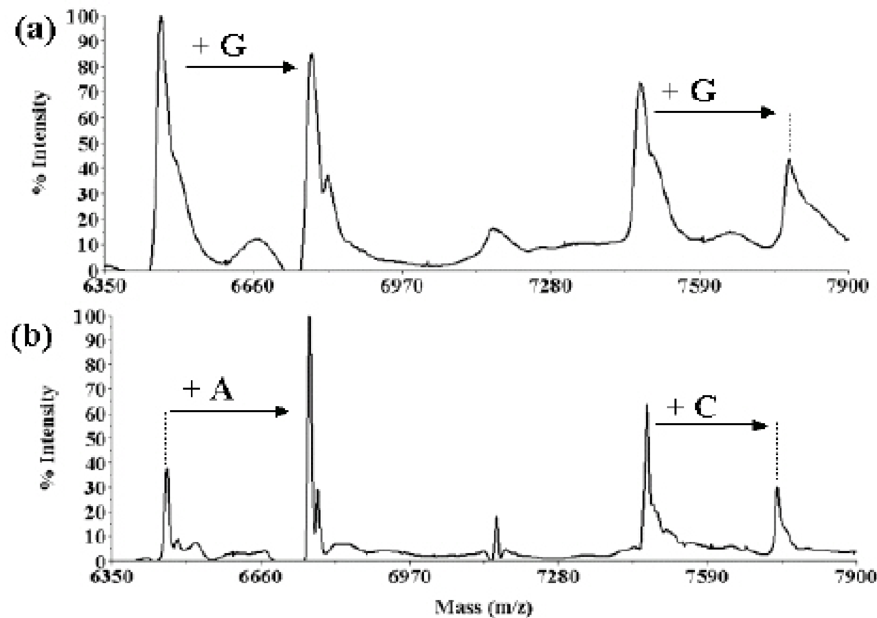

Figure 3. Duplex SNuPE assay used to score two SNPs involved in β-amylase thermostability. Samples were analysed by MALDI-ToF MS. *Hordeum spontaneum* line `AB75' carries the Sd3 allele (a), extending the genotyping primers (6468 Da and 7468 Da) by single ddGTPs to generate products of 6781 Da and 7781 Da. Cultivar Stirling carries the Sd2L allele (b), extending the first genotyping primer (6468 Da) with a single ddATP to generate a 6765 Da product and the second genotyping primer (7468 Da) by a single ddCTP to generate a 7741 Da product.

By July 2001, over 3 million human SNPs were available through public databases. Considerable effort is being devoted to SNP identification in plants, with SNP maps already available for some species (e.g. Cho *et al.*, 1999). Large numbers of SNPs are currently being identified in maize by database mining and alignment of EST sequences (Useche *et al.*, 2001). Analogous projects for other plant species, such as barley, have already been established.

Depending on the species and the genomic area under investigation, recent studies indicate SNP frequency to range from as high as 1 SNP per 10 to 15 bp in some non-coding regions of the human genome (Brookes, 1999) to a much lower 1 SNP per 4 kb in some highly conserved regions (Nickerson *et al.*, 1998). An average frequency of 1 SNP per 70 bp was estimated in maize (Rafalski *et al.*, 2001). With these figures in mind and the assumed average physical to genetic distance ratio of 4.4 Mb/cM in barley (Künzel *et al.*, 2000), SNPs could be estimated to occur, on average, once every 0.000016 cM in barley (70 bp/4.4 Mb/cM) - or, interpreted otherwise, 63000

SNPs could potentially be present in every cM. These estimates will most likely be adjusted significantly when detailed reports of barley SNP frequency become available and will always vary in function of genome position because of major differences in recombination rates across any given chromosome. They should, however, occur frequently in most regions.

Within coding exons the nucleotide diversity is some four-fold lower, with about half resulting in non-synonymous codon changes (Brookes, 1999; Nickerson *et al.*, 1998). Although not all SNPs are useful as genetic markers, many, such as SNPs identified by aligning cDNA sequences (cSNPs), and promoter region SNPs (pSNPs), can directly affect gene function and are `perfect' molecular markers, with no chance of recombination between the marker and the gene being selected for. When amino acid substitution occurs, protein function and morphology can be altered and, as a result, affect phenotype.

Marker-assisted backcrossing could be accelerated using a SNP based strategy by selecting backcross progeny for high recurrent content as well as for the donor trait(s). The recurrent fraction inherited can be quantified by genotyping selected SNPs interspersed throughout the genome while confirming the presence of the donor gene desired. The more detailed genetic analysis of the backcross progeny obtained using this strategy would enable accurate selection of superior genotypes and, perhaps, the elimination of breeding generations.

Conclusion

Both single and multiplex SNuPE assays analysed by MALDI-ToF MS were used to genotype SNPs from barley seedlings. The method allows co-dominant selection of barley seedlings and, because the markers are `perfect', there is no chance of recombination between the marker and the genes of interest. Further work is required to identify more useful SNPs in barley, directly related or at least closely related to traits of interest. Promoter region SNPs (pSNPs) and cSNPs should play substantial roles in the next generation of molecular markers used for marker-assisted breeding. Ultimately, sets of coding SNPs and/or pSNPs associated with most traits of interest will be unravelled. In the future, the availability of this information should lead to the development of strategies allowing complete genotyping of plant progeny using DNA chips or yet to be developed systems. Technology improvements and new assay concepts should facilitate SNP genotyping and eventually lead to the elimination of the costly and time-consuming PCR step.

Acknowledgment

Funding was kindly provided by the Australian Grains Research and Development Corporation (GRDC).

References

1. Brookes, A. J. (1999) The essence of SNPs, Gene, 234: 177-86.
2. Cho, R. J., Mindrinos, M., Richards, D. R., Sapolsky, R. J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T. L., Stammers, M., Federspiel, N., Theologis, A., Yang, W. H., Hubbell, E., Au, M., Chung, E. Y., Lashkari, D., Lemieux, B., Dean, C., Lipshutz, R. J., Ausubel, F. M., Davis, R. W., and Oefner, P. J. (1999) Genome-Wide Mapping with Biallelic Markers in *Arabidopsis thaliana*, Nature Genetics, 23: 203-7.
3. Eglinton, J. K., Langridge, P., and Evans, D. E. (1998) Thermostability Variation in Alleles of Barley Beta-Amylase, Journal of Cereal Science, 28: 301-309.
4. Griffin, T. J., and Smith, L. M. (2000) Single-Nucleotide Polymorphism Analysis by MALDI-TOF Mass Spectrometry, Trends in Biotechnology, 18: 77-84.
5. Künzel, G., Korzun, L., Meister, A., and Endo, T. R. (2000) High Resolution Physical Mapping of the Barley Genome, International Barley Genetics Symposium VIII, I: 293-298.
6. Landegren, U., Nilsson, M., and Kwok, P. Y. (1998) Reading Bits of Genetic Information - Methods for Single Nucleotide Polymorphism Analysis, Genome Research, 8: 769-776.
7. Nickerson, D. A., Taylor, S. L., Weiss, K. M., Clark, A. G., Hutchinson, R. G., Stengard, J., Salomaa, V., Vartiainen, E., Boerwinkle, E., and Sing, C. F. (1998) DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene, Nature Genetics, 19: 233-40.
8. Paris, M., and Carter, M. (2000) Cereal DNA: A Rapid High-Throughput Extraction Method for Marker Assisted Selection, Plant Molecular Biology Reporter, 18: 357-360.
9. Paris, M., Potter, R., and Jones, M. (2001) Typing Barley *Mlo* Alleles by Single Nucleotide Polymorphim Analysis using MALDI-ToF Mass Spectrometry, Plant and Animal Genome IX, P335.
10. Rafalski, A., Ching, A., Bhattramakki, D., Morgante, M., Dolan, M., Register, J. C., Smith, O. S., and Tingey, S. (2001) SNP Markers in Maize: Discovery and Applications, Plant and Animal Genome IX, W149.
11. Useche, F., Morgante, M., Hanafey, M., Tingey, S., Martins, W., Gao, G. R., and Rafalski, A. (2001) Computer Detection of Single Nucleotide Polymorphisms (SNPs) in Maize ESTs, Plant and Animal Genome IX, P333.