DNA polymorphism in wild emmer wheat (*Triticum turgidum* subsp. *dicoccoides*) using barley expressed sequence tag-derived microsatellite markers

Emebiri L

Biosciences Research Division, Department of Primary Industries, Grains Innovation Park, Natimuk Road, Private Bag 260, Horsham Vic. 3401, Australia

INTRODUCTION

EST-SSRs have become the marker of choice for population genetic analyses due to the potential for analysis of functional diversity (Ayres *et al*. 1997; Saha *et al*. 2004) and a higher transferability across taxa than SSR markers generated from genomic DNA libraries (Ellis and Burke 2007). Furthermore, they are relatively easy and inexpensive to develop using publicly available EST databases and genomic softwares, and tend to produce substantially 'cleaner' data (that is, easier to analyse/interpret amplification profiles), compared to their anonymous counterparts (Pashley *et al*. 2006).

Although the genomes of grasses are very different in terms of size, ploidy level and chromosome number, gene order and sequences have been found to be well-conserved (Dubcovsky *et al*. 1996; Devos and Gale 2000), making transferable EST-SSR markers very attractive for population genetic analysis, as the function of many of the transcripts can be predicted through homology searches from the genomic databases. The present study sought to explore usefulness of newly developed barley EST-SSR markers for population genetic analysis in a sample collection of wild emmer wheat (*Triticum turgidum* subsp. *dicoccoides*), and to compare molecular diversity at these loci with those revealed by a set of wheat EST-SSR markers known to be transferable across the Triticeae.

MATERIALS AND METHODS

Plant materials

Plant materials used for the study comprised 85 accessions of wild emmer wheat obtained from the International Maize and Wheat Improvement Center (CIMMYT) through the CIMMYT-Australian Germplasm Evaluation (CAGE) suite of projects. The list is available at the CAGE site (http://mendel.lafs.uq.edu.au:8080/ICIS5/GWIS_LBY6.htm) For comparison, 9 durum (*Triticum turgidum durum;* AABB genome) cultivars and 2 accessions of *Aegilops tauschii* (DD genome) were included in the study.

Development of New Barley EST-SSR

Barley EST libraries (HVSME) developed at the Clemson University Genomics Institute (**CUGI**), were used to identify SSRs. A set of 17,269 sequences derived from the developing spike of the six-rowed cultivar, Morex, were screened for microsatellite-containing motifs using the Biome SSR Discovery Tool (Jewell *et al*. 2006), which integrates SPUTNIK, an SSR repeat finder, with Primer3, a PCR primer design program, into one pipeline tool. A total of 1445 primers pairs were designed, out of which 180 with the longest size of repeat motifs were selected for synthesis. Five of the polymorphic loci were selected for this study, based on their distribution across chromosomes in two barley maps. Summary details of the EST-SSR markers are provided in Table 1.

Wheat EST-SSR primers, developed from wheat ESTs available in the wEST database (http://wheat.pw.usda.gov/cgi-bin/ace/search/wEST), were retrieved from the published reports of Eujayl *et al*. (2002) and Peng *et al*. (2005). Only markers assigned to unique wheat chromosomes in these previous studies were selected.

Analysis of levels of variation

Alleles were recorded as co-dominant to avoid potential loss of information and allow accurate assessment of true genetic relationships. Basic statistics, such as gene diversity (or expected heterozygosity), allele per locus and polymorphism information content, were calculated using the Excel add-in MicroSatellite Toolkit (Park 2001). Genetic distance was calculated by the method introduced by Peakall *et al*. (1995) for codominant markers, and implemented in GenAlEX (Peakall and Smouse 2006). Mantel's non-parametric test (Mantel 1967) was carried out to determine whether the genetic similarities generated by the barley and wheat EST–SSRs provide similar measures of genetic relatedness.

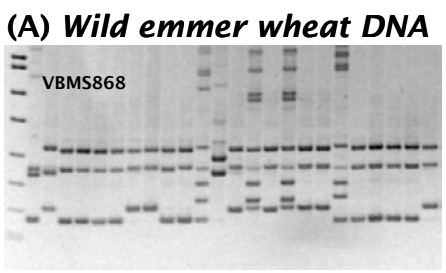
RESULTS AND DISCUSSIONS

Amplification of barley EST-SSR in wild emmer wheat

The Barley EST-SSR markers produced identical amplification patterns in both barley and wild emmer wheat, as demonstrated in Figure 1A,B. The number of alleles per locus ranged from 4 to 27 with an average of 12.3, which compared favourably with the average of 14.3 obtained with the wheat EST-SSR markers (Figure 1C). However, they were slightly less polymorphic, with PIC of 0.53 vs 0.66, and showed lower gene diversity, but two of the barley EST-SSR markers (VBMS868 and umb301) exhibited comparably high PIC values in the wild emmer wheat accessions (Figure 1C).

Table. 1. Primer sequence of barley EST-SSR markers and information on homology to wheat ESTs

Locus	Chr.	GeneBank accession No.	Primer 5'-3'	Homology to wheat EST	Best hit e-value
VBMS868	7H	BG368981	F: CTGCAAGAAGCCAAGAATAC R: ATTGGGAGTGCTAGGAGACT	BE499720	6e-68
VBMS5	7H	AF474373	F: TGAAGCTGACTACGACAATG R: GAACTTTCCCTTTGAGAGGT	AB029061	0.0
umb301	3H	AV944239	F: CTTCACATGTCTGGGAAAACA R: GACATGTTGGAAGGTGGCTT	AF085169	1e-63
VBMS221	5H	BE454337	F: GTCTTCGTACTCGCCTCTC R: CTCAGGGTGTAAGAGCTGTC	BE499481	7e-77
VBMS743	5H	BE194301	F: GTCTTCGTACTCGCCTCTC R: CTGAAGAAGGTGTTGAAAGC	BF482680	1e-97
VBMS310	6H	BE601955	F: ATCCAGTTTCAGCCACCA R: CGGTAGTAGTGGTACGTCG	BE605026	1e-82



(C)

Marker	No. of Alleles /locus	Nei's unbiased Gene Diversity	PIC
Barley EST-SSR			
VBMS868A	27.00	0.87	0.86
VBMS868B	27.00	0.94	0.93
VBMS868C	17.00	0.79	0.75
VBMS5	4.00	0.15	0.14
umb301	18.00	0.77	0.74
VBMS221B	4.00	0.42	0.35
VBMS743A	5.00	0.39	0.34
VBMS743B	9.00	0.60	0.55
VBMS743C	4.00	0.19	0.18
VBMS310	8.00	0.54	0.43
Average	*12.30*	*0.57*	*0.53*
Wheat EST-SSR			
cwem12C	22.00	0.70	0.68
cwem34g1	11.00	0.54	0.52
cwem34g2	7.00	0.75	0.71
cwem14B	10.00	0.60	0.57
cwem38D1	6.00	0.52	0.47
cwem38D2	5.00	0.66	0.61
DuPw004	29.00	0.82	0.81
DuPw038	20.00	0.87	0.86
DuPw023	19.00	0.71	0.68
Average	*14.33*	*0.69*	*0.66*

Note: PIC, polymorphism information content

Figure 1. Amplification pattern of the newly developed barley EST-SSR (VBMS868) in (A) wild emmer wheat and (B) barley; (C) summary statistics of population genetic parameters across EST-SSR markers in wild emmer wheat.

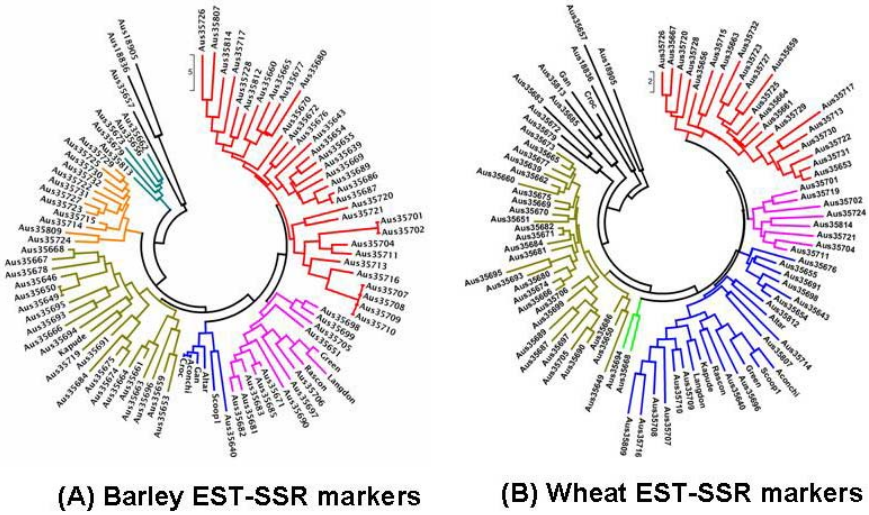


Figure 2. Near Neighbour phylogenetic trees of wheat germplasm accessions based on polymorphism at barley and wheat EST-SSR loci. The genotypes include 85 wild emmer wheat lines, 9 durum cultivars and 2 *Ae. tauschii* accessions.

Analysis of population structure

A neighbour joining dendrogram was constructed from the pair-wise genetic distance calculated using barley EST-SSR. The analysis divided the wheat accessions into 7 major groups of genetically related individuals, and correctly placed the *Ae*. *tauschii* and durum cultivars in separate unique groups from the wild emmer accessions (Figure 2A). In comparison, a similar analysis with wheat EST-SSR grouped the accessions into 6 groups (Figure 2B). Group memberships were different, and the genetic distance matrices showed no significant association, as established by Mantel test (Mantel 1967). This indicates that the barley EST-SSR markers explored a different spectrum of the diversity present in wild emmer wheat.

CONCLUSION

The major finding in this study is that EST-SSR markers derived from cDNA sequence of an elite malting barley cultivar (Morex) can be used to characterise genetic diversity in a wild, distantly related species, such as the wild emmer wheat. As the function of many of the transcripts can be predicted through homology searches from the genomic databases, the information has important implications for characterising the A genome of wild emmer wheat, identification of genes associated with domestication and understanding the functional basis of genetic diversity.

REFERENCES

Ayers NM, McClung AM, Larkin PD, Bligh HFJ, *et al*. (1997) Theor Appl Genet 94: 773-781.

Devos KM, Gale MD (2000) Plant Cell 12: 637-646.

Dubcovsky J, Luo MC, Zhong GY, Bransteitter R, *et al*. (1996) Genetics 143: 983-999.

Ellis JR, Burke JM (2007) Heredity 99: 125-132.

Eujayl I, Sorrells M, Baum M, Wolters P, Powell W (2002) Theor Appl Genet, 104: 399-407.

Jewell E, Robinson A, Savage D, Erwin T, *et al*. (2006) Nucleic Acids Research, 34, Web Server issue.

Mantel N (1967) Cancer Res. 27: 209-220.

Park SDE (2001) University of Dublin (http://www.animalgenomics.ucd.ie/sdepark/ms-toolkit/).

Pashley CH, Ellis JR, McCauley DE, Burke JM (2006) J Hered 97: 381-388.

Peakall R, Smouse PE & Huff DR (1995) Mol Ecol, 4: 135-147.

Peakall R, Smouse PE (2006) GENAlEX 6, Mol Ecol. Notes, 6: 288-295.

Peleg Z, Fahima T, Abbo S, Krugman T Saranga Y (2008) Genome, 5: 187-195.

Peng JH, Nora L, Lapitan V (2005) Funct Integr Genomics, 5: 80–96.

Saha MC, Mian MAR, Eujayl I, Zwonitzer JC, Wang L, *et al*. (2004) Theor Appl Genet 109: 783-791.